Get Full Access and More at ExpertConsult.com

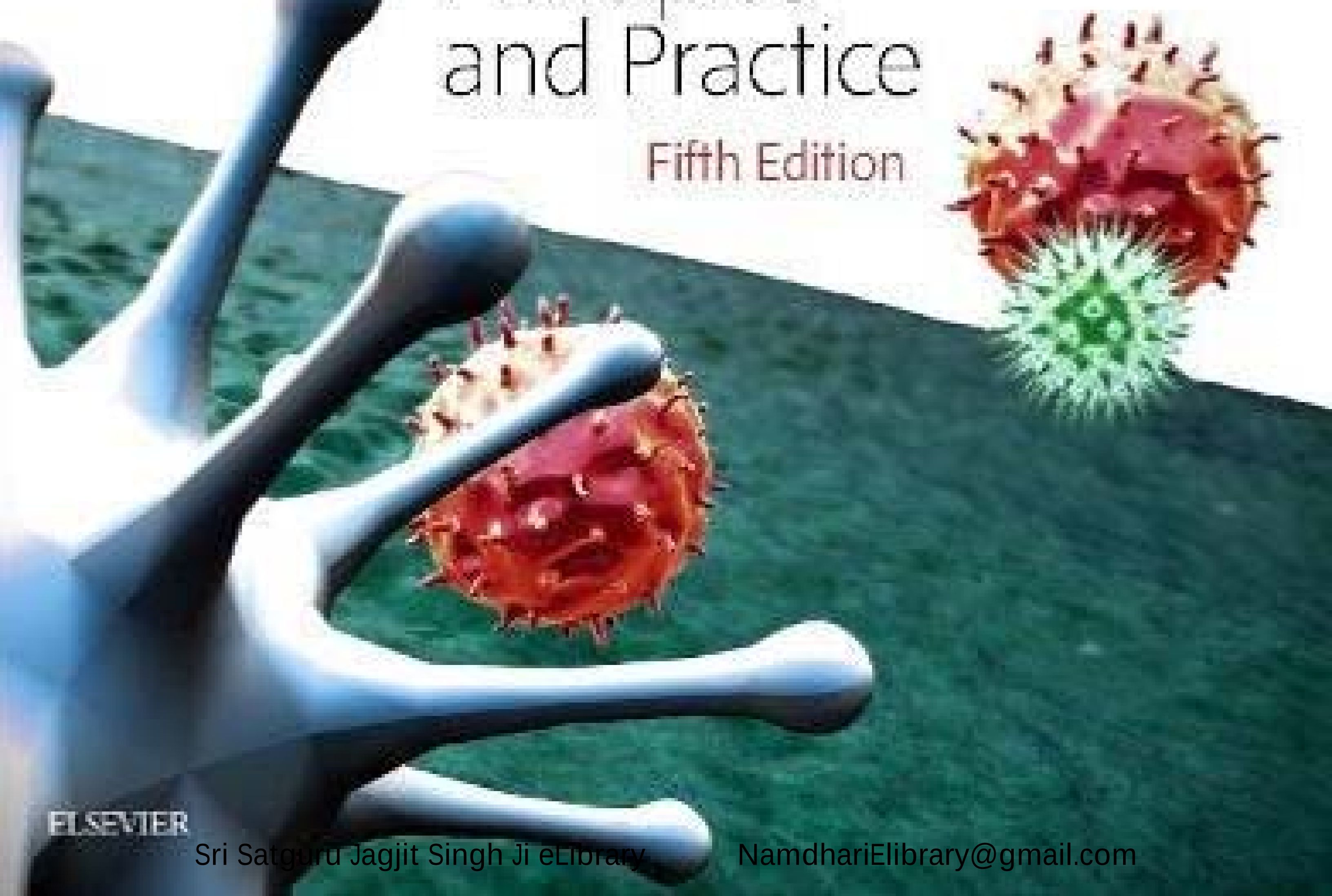
Edited by **Robert R. Rich**

Thomas A. Fleisher | William T. Shearer | Harry W. Schroeder Jr. | Anthony J. Frew | Cornelia M. Weyand

Clinical Immunology

Principles and Practice

Fifth Edition



ELSEVIER

Clinical Immunology: Principles and Practice

FIFTH EDITION

ROBERT R. RICH MD
Professor of Medicine and Dean Emeritus, University of Alabama at Birmingham, Birmingham, AL, USA

THOMAS A. FLEISHER MD
Executive Vice President, American Academy of Allergy, Asthma and Immunology, Milwaukee, WI; Scientist Emeritus, NIH Clinical Center, National Institutes of Health, Bethesda, MD, USA

WILLIAM T. SHEARER MD, PhD
Allergy and Immunology Service, Texas Children's Hospital, Professor of Pediatrics and Immunology, Section of Allergy and Immunology, Department of Pediatrics, Baylor College of Medicine, Houston, TX, USA

HARRY W. SCHROEDER, JR. MD, PhD
Professor of Medicine, Microbiology, and Genetics, Division of Clinical Immunology and Rheumatology, Director, UAB Program in Immunology, University of Alabama at Birmingham, Birmingham, AL, USA

ANTHONY J. FREW MD, FRCP
Professor of Allergy and Respiratory Medicine, Department of Respiratory Medicine, Royal Sussex County Hospital, Brighton, UK

CORNELIA M. WEYAND MD, PhD
Professor of Medicine, Stanford University, Stanford, CA, USA

ELSEVIER For additional online content visit ExpertConsult.com

ELSEVIER

© 2019, Elsevier Limited. All rights reserved.

First edition 1996
Second edition 2001
Third edition 2008
Fourth edition 2013
Fifth edition 2019

The right of Robert R. Rich, Thomas A. Fleisher, William T. Shearer, Harry W. Schroeder Jr., Anthony J. Frew, Cornelia M. Weyand to be identified as authors of this work has been asserted by them in accordance with the Copyright, Designs and Patents Act 1988.

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the publisher. Details on how to seek permission, further information about the Publisher's permissions policies and our arrangements with organizations such as the Copyright Clearance Center and the Copyright Licensing Agency, can be found at our website: www.elsevier.com/permissions.

This book and the individual contributions contained in it are protected under copyright by the Publisher (other than as may be noted herein).

Chapter 9, Cytokines and cytokine receptors is in the public domain.
Chapter 79, Lymphomas; Elaine S. Jaffe and Stefania Pittaluga contributions are the public domain.
Chapter 88, Protein kinase antagonists as therapeutic agents for immunological and inflammatory disorders; John J. O'Shea and Massimo Gadina contributions are in the public domain.
Chapter 92, Flow Cytometry is in the public domain.

Notices

Practitioners and researchers must always rely on their own experience and knowledge in evaluating and using any information, methods, compounds or experiments described herein. Because of rapid advances in the medical sciences, in particular, independent verification of diagnoses and drug dosages should be made. To the fullest extent of the law, no responsibility is assumed by Elsevier, authors, editors or contributors for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions, or ideas contained in the material herein.

ISBN: 978-0-7020-6896-6
E-ISBN: 978-0-7020-7039-6

Content Strategist: Belinda Kuhn/Kayla Wolfe
Content Development Specialist: Joanne Scott
Content Coordinator: Joshua Mearns
Project Manager: Andrew Riley
Design: Amy Buxton
Illustration Manager: Amy Faith Heyden
Illustrators: Martin Woodward, DNA Illustrations, Inc.
Marketing Manager: Kristin Koehler

ELSEVIER
Book Aid International
Working together to grow libraries in developing countries
www.elsevier.com • www.bookaid.org

Printed in China

Last digit is the print number: 9 8 7 6 5 4 3 2 1

PREFACE TO THE FIRST EDITION

https://archive.org/details/namdhari

Clinical immunology is a discipline with a distinguished history, rooted in the prevention and treatment of infectious diseases in the late nineteenth and early twentieth centuries. The conquest of historical scourges such as smallpox and (substantially) polio and relegation of several other diseases to the category of medical curiosities is often regarded as the most important achievement of medical science of the past fifty years. Nevertheless, the challenges facing immunologists in the efforts to control infectious diseases remain formidable; HIV infection, malaria and tuberculosis are but three examples of diseases of global import that elude control despite major commitments of monetary and intellectual resources.

Although firmly grounded in the study and application of defenses to microbial infection, since the 1960s clinical immunology has emerged as a far broader discipline. Dysfunction of the immune system has been increasingly recognized as a pathogenic mechanism that can lead to an array of specific diseases and failure of virtually every organ system. Pardoxically, although the importance of the immune system in disease pathogenesis is generally appreciated, the place of clinical immunology as a practice discipline has been less clear. As most of the noninfectious diseases if the human immune system lead eventually to failure of other organs, it has been organ-specific subspecialists who have usually dealt with their consequences. Recently, however, the outlook has begun to change as new diagnostic tools increasingly allow the theoretical possibility of intervention much earlier in disease processes, often before irreversible target organ destruction occurs. More importantly, this theoretical possibility is increasingly realized as clinical immunologists find themselves in the vanguard of translating molecular medicine from laboratory bench to patient bedside.

In many settings, clinical immunologists today function as primary care physicians in the management of patients with inmune-deficiency, allergic, and autoimmune diseases. Indeed many influential voices in the clinical disciplines of allergy and rheumatology support increasing coalescence of these traditional subspecialities around their intellectual core of immunology. In addition to his or her role as a primary care physician, the clinical immunologist is increasingly being looked to as a consultant, as scientific and clinical advances enhance his or her expertise. The immunologist with a ‘generalist’ perspective can be particularly helpful in the application of unifying principles of diagnosis and treatment across the broad spectrum of immunologic diseases.

Clinical Immunology: Principles and Practice has emerged from this concept of the clinical immunologist as both primary care physician and expert consultant in the management of patients with immunologic diseases. It opens in full appreciation of the critical role of fundamental immunology in this rapidly evolving clinical discipline. Authors of basic science chapters were asked, however, to cast their subjects in a context of clinical relevance. We believe the result is a well-balanced exposition of basic immunology for the clinician.

The initial two sections on basic principles of immunology are followed by two sections that focus in detail on the role of the immune system in defenses against infectious organisms. The approach is two-pronged. It begins first with a systematic survey of immune responses to pathogenic agents followed by a detailed treatment of immunologic deficiency syndromes. Pathogenic mechanisms of both congenital and acquired immune deficiency diseases are discussed, as are the infectious complications that characterize these diseases. Befitting its importance, the subject of HIV infection and AIDS receives particular attention, with separate chapters on the problem of infection in the immunocompromised host, HIV infection in children, anti-retroviral therapy and current progress in the development of HIV vaccines.

The classic allergic diseases are the most common immunologic diseases in the population, ranging from atopic disease to drug allergy to organ-specific allergic disease (e.g., of the lungs, eye and skin). They constitute a foundation for the practice of clinical immunology, particularly for those physicians with a practice orientation defined by formal subspecialty training in allergy and immunology. A major section is consequently devoted to these diseases, with an emphasis on pathophysiology as the basis for rational management.

The next two sections deal separately with systemic and organ-specific immunologic diseases. The diseases considered in the first of these sections are generally regarded as the core practice of the clinical immunologist with subdisciplinary emphasis in rheumatology. The second section considers diseases of specific organ failure as consequences of immunologically mediated processes that may involve virtually any organ system. These diseases include as typical examples the demyelinating diseases, insulin-dependent diabetes mellitus, the glomerulonephritides and inflammatory bowel diseases. It is in management of such diseases that the discipline of clinical immunology will have an increasing role as efforts focus on intervention early in the pathogenic process and involve diagnostic and therapeutic tools of ever-increasing sophistication.

One of the major clinical areas in which the expertise of a clinical immunologist is most frequently sought is that of allogeneic organ transplantation. A full section is devoted to the issue of transplantation of solid organs, with an introductory chapter on general principles of transplantation and management of transplantation rejection followed by separate chapters dealing with the special problems of transplantation of specific organs or organ systems.

Appreciation of both the molecular and clinical features of lymphoid malignancies is important to the clinical immunologist regardless of subspecialty background, notwithstanding the fact that primary responsibility for management of such patients will generally fall to the haematologist/oncologist. A separate section is consequently devoted to the lymphocytic leukemias and lymphomas that constitute the majority of malignancies seen in the context of a clinical immunology practice. The separate issues of immune responses to tumors and immunological strategies to treatment of malignant diseases are subjects of additional chapters.

Another important feature is the attention to therapy of immunologic diseases. This theme is constant throughout the chapters on the allergic and immunologic diseases, and because of the importance the editors attach to clinical immunology as a therapeutic discipline, an extensive section is also devoted specifically to this subject. Subsections are devoted to issues of immunologic reconstitution, with three chapters on treatment of immunodeficiences, malignancies and metabolic diseases by bonemarrow transplantation. Also included is a series of chapters on pharmaceutical agents currently available to clinical immunologists, both as anti-allergic and anti-inflammatory drugs, as well as newer agents with greater specificity for T cell-mediated immune responses. The section concludes with a series of chapters that address established and potential applications of therapeutic agents and approaches that are largely based on the new techniques of molecular medicine. In addition to pharmaceutical agents the section deals in detail with such subjects as apheresis, cytokines, monoclonal antibodies and immunotoxins, gene therapy and new experimental approaches to the treatment of autoimmunity. The book concludes with a section devoted to approaches and specific techniques involved in the diagnosis of immunologic diseases. Use of the diagnostic laboratory in evaluation of complex problems of immunopathogenesis has been a hallmark of the clinical immunologist since inception of the discipline and many clinical immunologists serve as directors of diagnostic immunology laboratories. Critical assessment of the utilization of techniques ranging from lymphocyte cloning to flow cytomeric phenotyping to molecular diagnostics are certain to continue as an important function of the clinical immunologist, particularly in his or her role as expert consultant.

In summary, we have intended to provide the reader with a comprehensive and authoritative treatise on the broad subject of clinical immunology, with particular emphasis on the diagnosis and treatment of immunological diseases. It is anticipated that the book will be used most frequently by the physician specialist practicing clinical immunology, both in his or her role as a primary physician and as a subsequent consultant. It is hoped, however, that the book will also be of considerable utility to the non-immunologist. Many of the diseases discussed authoritatively in the book are diseases commonly encountered by the generalist physician. Indeed, as noted, because clinical immunology involves diseases of virtually all organ systems, competence in the diagnosis and management of immunological diseases is important to virtually all clinicians. The editors would be particularly pleased to see the book among the references readily available to the practicing internist, pediatrician and family physician.

Robert R. Rich
Thomas A. Fleisher
Benjamin D. Schwartz
William T. Shearer
Warren Strober

PREFACE TO THE FIFTH EDITION

Each edition of *Clinical Immunology: Principles and Practice* has documented important changes in the discipline from the preceding one. This fifth edition is emphatically not an exception to that pattern. Indeed, advances in both the *Principles* and *Practice* of clinical immunology have been remarkable. The constant theme of every edition has been to emphasize that our discipline touches virtually all organ systems. The diseases that are covered range from too little to too much immunity; and from dysregulated, malignant or replaced immunological systems and functions. Fundamental concepts that are essential to a precise understanding of normal and disordered immune function and disease pathogenesis are again balanced by clinical descriptions, diagnostic approaches and therapeutic options.

Several examples of particularly notable advances are worth highlighting: Our increasing appreciation of the importance of the microbiota to normal immune system development and to the pathogenesis of immunologic and inflammatory diseases; dissection of relationships between the innate and adaptive immune systems that has served to further clarify the expression of inflammatory processes and their interaction in defenses against infectious agents; progress in rapid and cost-effective genomics that has led to the definition of numerous new primary immune deficiencies and provided new insights into the genetic aspects of many other immunologic diseases; understanding of immune deficiencies that reflect development of anti-cytokine auto-antibodies; detailed definition of cell signaling pathways and the structure of cell-surface molecules that have contributed enormously to the treatment of cancer and autoimmunity with a virtual explosion in novel therapeutics including check-point inhibitors and other recently developed immunomodulators; availability of many new humanized and human monoclonal antibodies and development of novel therapeutic approaches such as chimeric-antigen-receptor T cells; wide use of T cell excision circle receptor (TREC) assay to diagnose serious immune deficiencies of the newborn; and exploration of *in vivo* therapeutic editing of pathological mutations. With these new tools the practice of clinical immunology has become more interesting yet more complex, while offering important improvements in patient care.

Our goal with this edition is to enhance the interest of practitioners in the many specialties and subspecialties that the discipline impacts and to assist them in understanding this increasing complexity. With the increasing availability of powerful new therapeutic agents, the expert clinical immunologist today may function as a primary care physician or consultant in the management of patients with immune deficiencies, allergic, and autoimmune diseases involving multiple organ systems – a role reminiscent of that served by infectious disease physicians, particularly during the era of antibiotic proliferation in the last quarter of the previous century. It is consequently our hope that the book will find a place near the desk of most persons whose practice relies on the science or practice of clinical immunology as we have broadly defined it. We further trust that it will be especially useful to trainees and practitioners preparing for certification or recertification in an immunology-related subspecialty. In an effort to assist the latter group and as an aid to continuing education of all readers, we have added to the online version of the text, which all purchasers of the book can readily access, multiple choice questions relevant to every chapter. Additionally, we continue to believe that a comprehensive text on clinical immunology can be a valuable asset for generalists in any specialty, particularly internists, pediatricians and family physicians, who regularly care for patients across the broad spectrum of immunological disorders, offering an opportunity for physicians in all disciplines to upgrade their skills and education, and to benefit from the onward rush of science and practice improvement in modern clinical immunology.

The book continues features that have been well received in previous editions. Chapters are generously illustrated and all chapters contain summary Boxes (commonly in bulleted form) that provide Key Concepts and a Box labeled as *On the Horizon*, in which authors look to research opportunities for important advances over the next 5-10 years. Furthermore, due to the extraordinarily cross-disciplinary nature of clinical immunology, it is our hope that investigators working in one area might find new ideas and opportunities in the *On the Horizon* Boxes outside their primary area of focus. Other Boxes similarly summarize content with *Clinical Relevance*, *Clinical Pearls*, and *Therapeutic Principles*.

As always, we are immensely grateful to the hundreds of physicians and scientists whose contributions are the essence of the book. Finally, we recognize the diligence and commitment of our colleagues at Elsevier who have supported all aspects of the book's development and production, particularly Ms. Joanne Scott who has worked with both authors and editors from concept to birth to completion of every chapter.

Robert R. Rich
Thomas A. Fleisher
William T. Shearer
Harry W. Schroeder, Jr.
Anthony J. Frew
Cornelia M. Weyand

LIST OF CONTRIBUTORS

The editor(s) would like to acknowledge and offer grateful thanks for the input of all previous editions' contributors, without whom this new edition would not have been possible.

Roshini Sarah Abraham PhD, D(ABMLI)
Consultant
Department of Laboratory Medicine and Pathology
Mayo Clinic
Rochester, MN, USA
Professor of Medicine, Professor of Laboratory Medicine and Pathology

Cristina Albanesi BSc, PhD
Senior Investigator
Laboratory of Experimental Immunology
Fondazione "Luigi Maria Monti" (FLMM)—Istituto Dermopatico dell'Immacolata (IDI)-IRCCS
Rome, Italy

Ilias Alevizos DMD, MMSc
Assistant Clinical Investigator
National Institute of Dental and Craniofacial Research
National Institutes of Health
Bethesda, MD, USA

Juan Anguita PhD
Ikerbasque Professor
CIC bioGUNE
Derio, Bizkaia, Spain

Brendan Antiochos MD
Instructor
Division of Rheumatology
Johns Hopkins University School of Medicine
Baltimore, MD, USA

Cynthia Aranow, MD
Investigator, Clinical Research
Autoimmune and Musculoskeletal Diseases
The Feinstein Institute for Medical Research
Manhasset, NY, USA

John P. Atkinson, MD
Department of Medicine
Chief, Division of Rheumatology
Samuel B. Grant Professor of Medicine
Professor of Molecular Microbiology
Washington University School of Medicine
St. Louis, MO, USA

Howard A. Austin III MD
Senior Clinical Investigator
National Institute of Diabetes and Digestive and Kidney Diseases
National Institutes of Health, Bethesda, MD, USA

Subash Babu MBBS, PhD
Scientific Director
NIH-NIRT-International Center for Excellence in Research
National Institute for Research in Tuberculosis
Chennai, India

Mark C. Ballow MD
Professor, Department of Pediatrics
Division of Allergy, Immunology and Pediatric Rheumatology
Women and Children's Hospital of Buffalo
State University of New York at Buffalo
School of Medicine and Biomedical Sciences
Buffalo, NY, USA

James E. Balow MD
Clinical Director and Chief
Kidney Disease Section
National Institute of Diabetes and Digestive and Kidney Diseases
National Institutes of Health
Bethesda, MD, USA

John W. Belmont MD, PhD
Professor
Department of Molecular and Human Genetics
Baylor College of Medicine
Houston, TX, USA

Claudia Berek PhD
Group Leader, B Cell Immunology
Deutsches Rheuma-Forschungszentrum Berlin (DRFZ)
Berlin, Germany

Timothy Beukelman MD, MSCE
Associate Professor of Pediatric Rheumatology
Division of Rheumatology
The University of Alabama at Birmingham
Birmingham, AL, USA

Tapan Bhavsar MD, PhD
Clinical Fellow in Hematopathology
National Cancer Institute/National Institute of Health
Bethesda, MD, USA

J. Andrew Bird MD
Associate Professor
Department of Pediatrics
Division of Allergy and Immunology
University of Texas Southwestern Medical Center in Dallas
Dallas, TX, USA

Sarah E. Blutt PhD
Assistant Professor
Department of Molecular Virology and Microbiology and Department of Molecular and Cellular Biology
Baylor College of Medicine
Houston, TX, USA

Mark Boguniewicz MD
Professor, Division of Pediatric Allergy-Immunology
National Jewish Health
Denver, CO, USA

Rafael Bonamichi-Santos, MD
Division of Clinical Immunology and Allergy Division
University of São Paulo
São Paulo, SP, Brazil

Bertrand Boisson PhD
Laboratory of Human Genetics of Infectious Diseases
Necker Branch, Imagine Institute;
Paris Descartes University, France;
St. Giles Laboratory of Human Genetics of Infectious Diseases, Rockefeller Branch
The Rockefeller University
New York, NY, USA

Elena Borzova MD, PhD
Professor of Clinical Allergy
Department of Clinical Allergology
Russian Medical Academy of Postgraduate Education
Moscow, Russian Federation

Prosper N. Boyaka PhD
Professor
Department of Veterinary Biosciences
The Ohio State University
College of Veterinary Medicine
Columbus, OH, USA

Joshua Boyce MD
Professor of Medicine and Pediatrics
Director, Inflammation and Allergic Disease Research Section
Director, Jeff and Penny Vinik Center for Allergic Disease Research
Harvard Medical School, Brigham and Women's Hospital
Boston, MA, USA

Sarah K. Browne
Office of Vaccine Research and Review
Center for Biologics Evaluation and Research
Food and Drug Administration
Silver Spring, MD, USA

Wesley Burks MD
Curnen Distinguished Professor
Executive Dean
School of Medicine
The University of North Carolina
Chapel Hill, NC, USA

Jacinta Bustamante MD, PhD
Research Associate
Laboratory of Human Genetics of Infectious Diseases
Necker Branch, Imagine Institute
Paris Descartes University
Paris Sorbonne Cité
Associate Professor of Cellular Biology
Study Center for Primary Immunodeficiencies
Necker Hospital, Assistance Publique Hôpitaux de Paris
Paris, France
St. Giles Laboratory of Human Genetics of Infectious Diseases, Rockefeller Branch
The Rockefeller University
New York, NY, USA

Virginia L. Calder PhD
Senior Lecturer in Immunology
Department of Molecular Therapy and Genetics
UCL Institute of Ophthalmology
London, UK

Matthew Campbell, MD, MS
Assistant Professor
Department of Genitourinary Medical Oncology
Division of Cancer Medicine
The University of Texas MD Anderson Cancer Center
Houston, TX, USA

Adela Rambi G. Cardones, MD
Assistant Professor
Department of Dermatology
Duke University School of Medicine
Durham, NC

Jean-Laurent Casanova MD, PhD
Director
St Giles Laboratory of Human Genetics of Infectious Diseases
Rockefeller Branch, The Rockefeller University
Howard Hughes Medical Institute
New York, NY, USA;
Co-Director
Laboratory of Human Genetics of Infectious Diseases
Necker Branch, Imagine Institute
Paris Descartes University
Paris Sorbonne Cité;
Professor of Pediatrics
Pediatric Hematology-Immunology Unit
Necker Hospital, Assistance Publique Hôpitaux de Paris
Paris, France

Mariana Castells MD, PhD
Director, Drug Hypersensitivity and Desensitization Center
Director, Allergy Immunology Training Program
Associate Director Mastocytosis Center
Brigham and Women's Hospital
Harvard Medical School
Boston, MA, USA

Lisa A. Cavacini PhD
Associate Professor
Department of Medicine
University of Massachusetts Medical School
MassBiologics
Boston, MA, USA

Edwin S.L. Chan MD, FRCPC
Assistant Professor
Department of Medicine
New York University School of Medicine
New York, NY, USA

David D. Chaplin, MD, PhD
Professor of Microbiology and Medicine
University of Alabama at Birmingham
Birmingham, AL

W. Winn Chatham MD
Professor of Medicine
Louis W. Heck Clinical Scholar
Clinical Director, Division of Clinical Immunology and Rheumatology
University of Alabama at Birmingham
Birmingham, AL, USA

Edward S. Chen, MD
Assistant Professor
Division of Pulmonary and Critical Care Medicine
Johns Hopkins University School of Medicine
Baltimore, MD USA

Javier Chinen MD, PhD
Assistant Professor
Departments of Pediatrics
Baylor College of Medicine
Houston, TX, USA

Lisa Christopher-Stine MD, MPH
Assistant Professor of Medicine
Director
John Hopkins Myositis Center
Johns Hopkins University Bloomberg School of Public Health
Baltimore, MD, USA

Michael Ciancanelli PhD
Research Associate
St Giles Laboratory of Human Genetics of Infectious Diseases
Rockefeller Branch
The Rockefeller University
New York, NY, USA

Andrew P. Cope BSc, PhD, MBBS, FRCP, FHEA
Head, Academic Department of Rheumatology
Centre for Molecular and Cellular Biology of Inflammation
Division of Immunology, Infection and Inflammatory Disease
King's College School of Medicine
King's College London
London, UK
Arthritis Research UK Professor of Rheumatology

David B. Corry MD
Professor
Departments of Medicine, Pathology and Immunology
Baylor College of Medicine
Houston, TX, USA

Filippo Crea MD, FESC, FACC
Full Professor
Department of Cardiology
Catholic University of Sacred Heart
Rome, Italy

Randy Q. Cron MD, PhD
Professor of Pediatrics and Medicine
Children's Hospital of Alabama
University of Alabama at Birmingham
Birmingham, AL, USA

Jennifer M. Cuellar-Rodriguez MD
Staff Clinician
Laboratory of Clinical Infectious Diseases
National Institute of Allergy and Infectious Diseases
National Institutes of Health
Bethesda, MD, USA

Marinos C. Dalakas MD, FAAN
Professor of Neurology
Director, Neuromuscular Division
Thomas Jefferson University Philadelphia PA and
Chief, Neuroimmunology Unit, Department of Pathophysiology
National and Kapodistrian University of Athens Medical School
Athens, Greece

Sara M. Dann PhD
Assistant Professor
Departments of Internal Medicine, and Microbiology and Immunology
University of Texas Medical Branch
Galveston, TX, USA

Betty Diamond MD
Professor
Department of Microbiology and Immunology and Medicine (AECOM)
The Feinstein Institute for Medical Research
Director
Laboratory of Autoimmune Diseases and Musculoskeletal Disorders
Head, Center for Autoimmune Diseases and Musculoskeletal Disorders
Manhasset, NY, USA

Terry W. Du Clos MD, PhD
Professor of Medicine
School of Medicine
University of New Mexico;
Head of Rheumatology
VA Medical Center
Albuquerque, NM, USA

Stéphanie Dupuis-Boisson PhD
Research Associate
Laboratory of Human Genetics of Infectious Diseases
Necker Branch, Imagine Institute
Paris Descartes University
Paris Sorbonne Cité, Paris, France;
St Giles Laboratory of Human Genetics of Infectious Diseases
Rockefeller Branch
The Rockefeller University
New York, NY, USA

Todd N. Eagar PhD
Assistant Professor
Department of Pathology and Genomic Medicine
Houston Methodist Hospital
Houston, TX, USA

Craig A. Elmets MD
Director, UAB Skin Diseases Research Center
University of Alabama at Birmingham
Birmingham, AL, USA
Professor and Chair, Department of Dermatology

Doruk Erkan MD
Associate Physician-Scientist
Barbara Volcker Center for Women and Rheumatic Disease
New York, NY, USA;
Associate Professor of Medicine
Weill Cornell Medical College
Associate Attending Rheumatologist, Hospital for Special Surgery
New York, NY, USA

Laura Fanning, MD
Instructor of Medicine
Harvard Medical School, Brigham and Women's Hospital
Boston, MA, USA

Erol Fikrig MD
Section Chief, Division of Infectious Diseases
Yale University
Investigator, Howard Hughes Medical Institute
Professor of Epidemiology (Microbial Diseases) and Microbial Pathogenesis
Waldemar Von Zedtwitz Professor of Medicine (Infectious Diseases)
New Haven, CT, USA

Davide Flego, PhD
Catholic University of the Sacred Heart
UNICATT
Institute of Cardiology
Milan, Italy

Thomas A. Fleisher MD
Executive Vice President
American Academy of Allergy, Asthma and Immunology
Milwaukee, WI, USA
Scientist Emeritus
NIH Clinical Center
National Institutes of Health
Bethesda, MD, USA

Luz Fonacier MD
Section of Allergy and Immunology
NYU Winthrop Hospital
Mineola, NY, USA

Andrew P. Fontenot MD
Henry N. Claman Professor of Medicine
Division Head, Allergy and Clinical Immunology
Department of Medicine
University of Colorado Anschutz Medical Campus
Aurora, CO, USA

Alexandra F. Freeman MD
Staff Clinician
Laboratory of Clinical Infectious Diseases
National Institute of Allergy and Infectious Diseases
National Institutes of Health
Bethesda, MD, USA

Anthony J. Frew MD, FRCP
Professor of Allergy and Respiratory Medicine
Department of Respiratory Medicine
Royal Sussex County Hospital
Brighton, UK

Kohtaro Fujihashi DDS, PhD
Professor, Department of Pediatric Dentistry
Immunobiology Vaccine Center
The Institute for Oral Health Research, School of Dentistry
The University of Alabama at Birmingham
Birmingham, AL, USA

Massimo Gadina PhD
Director
Office of Science and Technology
National Institute of Arthritis Musculoskeletal and Skin Diseases
National Institutes of Health
Bethesda, MD, USA

Moshe E. Gatt MD
Resident, Department of Hematology
Hadassah-Hebrew University Medical Center
Jerusalem, Israel

M. Eric Gershwin MD
Chief, Division of Rheumatology
Allergy and Clinical Immunology
University of California Davis Health System
Distinguished Professor of Medicine, Jack and Donald Chi Professor of Medicine
Davis, CA, USA

Susan L. Gillespie MD, PhD
Associate Professor of Pediatrics
Baylor College of Medicine
Baylor International Pediatric AIDS Initiative (BIPAI)
Texas Children's Health Center for International Adoption
Houston, TX, USA

Jörg J. Goronzy MD, PhD
Professor of Medicine
Stanford University School of Medicine
Stanford, CA, USA

Sangeeta Goswami, MD, PhD
Clinical Specialist
Research Instructor, Department of Genitourinary Medical Oncology
Research, Division of Cancer Medicine The University of Texas MD Anderson Cancer Center
Houston, TX, USA

Clive E.H. Grattan MD, FRCP
Consultant Dermatologist
Dermatology Centre
Norfolk and Norwich University Hospital
Norwich, UK

Neil S. Greenspan MD, PhD
Professor of Pathology
Case Western Reserve University
Cleveland, OH, USA

Sarthak Gupta, MD
Systemic Autoimmunity Branch
National Institute of Arthritis and Musculoskeletal and Skin Diseases
National Institutes of Health
Bethesda, MD, USA

Claire E. Gustafson, PhD
Division of Immunology and Rheumatology
Department of Medicine
Stanford University
Stanford, CA, USA

Russell P. Hall III MD
J. Lamar Callaway Professor and Chair
Department of Dermatology
Duke University School of Medicine
Durham, NC, USA

Robert G. Hamilton PhD, D(ABMLI)
Professor of Medicine and Pathology
Johns Hopkins University School of Medicine and
Director
Johns Hopkins Dermatology, Allergy and Clinical Immunology Reference Laboratory
Johns Hopkins University School of Medicine
Baltimore, MD, USA

Laurie E. Harrington, PhD
Associate Professor
Department of Cell, Developmental and Integrative Biology
University of Alabama at Birmingham
Birmingham, AL, USA

Leonard C. Harrison MD DSc, DMedSci (hon. causa), FRACP, FRCPA, FAHMS
Professor, NHMRC Senior Principal Research Fellow
Population Health and Immunity Division
The Walter and Eliza Hall Institute of Medical Research
Victoria, Australia

Sarfaraz A. Hasni MD
Lawrence Schulman Clinical Research Scholar
National Institute of Arthritis, Musculoskeletal and Skin Diseases
National Institutes for Health
Bethesda, MD, USA

Arthur Helbling MD
Associate Professor of Allergology and Clinical Immunology
Division of Allergology
University Clinic for Rheumatology, Immunology and Allergology (RIA) Inselspital
Bern, Switzerland

Joanna Hester PhD
Kidney Research UK Senior Fellow
Nuffield Department of Surgical Sciences
University of Oxford
Oxford, UK

Steven M. Holland MD
Chief, Immunopathogenesis Section;
Tenured Investigator
National Institute of Allergy and Infectious Diseases
National Institutes of Health
Bethesda, MD, USA
Chief, Laboratory of Clinical Infectious Diseases

Dennis Hourcade, PhD
Professor of Medicine
Washington University
School of Medicine
St. Louis, MO, USA

Nicholas D. Huntington PhD
Molecular Immunology Division
The Walter and Eliza Hall Institute of Medical Research;
Department of Medical Biology
University of Melbourne
Parkville, Victoria, Australia

Tracy Hwangpo MD, PhD
Instructor
Division of Clinical Immunology and Rheumatology, Department of Medicine
University of Alabama at Birmingham School of Medicine
Birmingham, AL, USA

John B. Imboden MD
Alice Betts Endowed Chair for Research in Arthritis
Professor of Medicine
University of California
San Francisco, CA, USA

Fadi Issa D.Phil. BM BCh,
Academic Clinical Lecturer
Nuffield Department of Surgical Sciences,
University of Oxford
Oxford, UK

Shai Izraeli MD
Associate Professor of Pediatrics
Department of Pediatric Hemato-Oncology
Edmond and Lily Safra Children Hospital
Sheba Medical Center and University of Tel-Aviv School of Medicine
Tel-Aviv, Israel

Elaine S. Jaffe MD
Head, Hematopathology
Laboratory of Pathology
Center for Cancer Research
National Cancer Institute
Bethesda, MD, USA

Sirpa Jalkanen MD, PhD
Academy Professor
Center of Excellence
University of Turku
Turku, Finland

Stacie Jones MD
Division of Allergy/Immunology, Department of PediatricsUniversity of Arkansas for Medical Sciences and Arkansas Children's Hospital
Little Rock, AR, USA

Emmanuelle Jouanguy PhD
Research Associate
Laboratory of Human Genetics of Infectious Diseases
Necker Branch, Imagine Institute
Paris Descartes University
Paris Sorbonne Cité, Paris, France;
St Giles Laboratory of Human Genetics of Infectious Diseases
Rockefeller Branch
The Rockefeller University
New York, NY, USA

Sarah Kabbani, MD
Emory Vaccinology Training Grant Fellow
Division of Infectious Diseases, Department of Medicine
Emory University School of Medicine
Atlanta, GA, USA

Stefan H.E. Kaufmann PhD DR. DR H.C
Director and Professor of Microbiology and Immunology
Max-Planck-Institute for Infection-Biology
Charité Universitätsmedizin Berlin
Berlin, Germany

Farrah Kheradmand MD
Professor of Medicine, Pathology and Immunology
Department of Medicine
Baylor College of Medicine
Houston, TX, USA

Donald B. Kohn MD
Departments of M.I.M.G. and Pediatrics
University of California
Los Angeles, CA, USA

Robert Korngold PhD
Chairman and Senior Scientist
Department of Research
Hackensack University Medical Center
Hackensack, NJ, USA

Anna Kovalszki MD
Assistant Professor
University of Michigan Medical School
Division of Allergy and Clinical Immunology
Department of Medicine
University of Michigan
Ann Arbor, MI, USA

Douglas B. Kuhns PhD
Head
Neutrophil Monitoring Laboratory
Clinical Services Program
SAIC-Frederick, Inc.
NCI Frederick
Frederick, MD, USA

Hrishikesh Kulkarni, MD
Pulmonary and Critical Care Fellow
Washington University
School of Medicine
St. Louis, MO, USA

Caroline Y. Kuo MD
Allergy and Immunology
UCLA Medical Center
Santa Monica, CA, USA

Arash Lahouti M.D
Postdoctoral Research Fellow
Division of Rheumatology
Johns Hopkins University School of Medicine
Baltimore, MD, USA

C. Ola Landgren MD, PhD
Myeloma Service
Department of Medicine
Memorial Sloan-Kettering Cancer Center
New York, CA, USA

Arian Laurence PhD
Postdoctoral Fellow
Molecular Immunology and Inflammation Branch
National Institute of Arthritis and Musculoskeletal and Skin Diseases
National Institutes of Health
Bethesda, MD, USA

Joyce S. Lee MD
Assistant Professor of Medicine
Department of Medicine
University of Colorado Anschutz Medical Campus
Aurora, CO, USA

Catherine Lemière MD, MSc
CIUSS du Nord de l'île de Montréal
Université de Montréal
Montréal, Canada

Donald Y. M. Leung, MD, PhD
Division of Pediatric Allergy-Immunology
National Jewish Health
Denver, CO, USA

Arnold I. Levinson MD
Emeritus Professor of Medicine
Pulmonary, Allergy and Critical Care Division
Perelman School of Medicine
University of Pennsylvania
Philadelphia, PA, USA

Ofer Levy MD, PhD
Director, Precision Vaccines Program & Staff Physician
Division of Infectious Diseases, Department of Medicine
Boston Children's Hospital;
Associate Professor of Pediatrics
Harvard Medical School
Boston, MA, USA

Dorothy E. Lewis PhD
Professor of Internal Medicine
Division of Infectious Diseases
University of Texas Health Sciences Center
Houston, TX, USA

Phoebe Lin MD, PhD
Casey Eye Institute
Oregon Health and Science University
Portland, OR, USA

Andreas Linkermann, FASN
Division of Nephrology
Department of Internal Medicine III University Hospital Carl Gustav Carus at the Technische Universität Dresden
Dresden, Germany

Giovanna Liuzzo MD, PhD
Aggregate Professor
Department of Cardiology
Catholic University of Sacred Heart
Rome Italy

Michael D. Lockshin MD, MACR
Co-Director, Mary Kirkland Center for Lupus Research
Hospital for Special Surgery
Attending Physician, Hospital for Special Surgery
Professor of Medicine and Obstetrics-Gynecology
Joan and Sanford Weill College of Medicine of Cornell University
Director, Barbara Volcker Center for Women and Rheumatic Disease
New York, NY, USA

Allison K. Lord PhD
Department of Medicine
Division of Infectious Diseases
Massachusetts General Hospital
Boston, MA, USA

Jay N. Lozier, MD, PhD,
Senior Staff Clinician
Department of Laboratory Medicine
National Institutes of Health Clinical Center
Bethesda, MD, USA

Amber Luong MD, PhD
Assistant Professor of Otorhinolaryngology - Head and Neck, Surgery and Immunology and Autoimmune Diseases
The University of Texas Health Science Center at Houston
Houston, TX, USA

Raashid Luqmani DM, FRCP, FRCPE
Professor of Rheumatology/Consultant Rheumatologist
Rheumatology Department, NDORMS
University of Oxford
Oxford, UK

Meggan Mackay, MD
Associate Investigator Clinical Research
Autoimmune and Musculoskeletal Diseases
The Feinstein Institute for Medical Research
Manhasset, NY, USA

Jonathan S. Maltzman, MD, PhD
Associate Professor of Medicine
Division of Nephrology
Stanford University School of Medicine
Palo Alto, CA, USA

Peter J. Mannon MD, MPH
Professor of Medicine
Mucosal HIV and Immunobiology Center
University of Alabama at Birmingham
Birmingham, AL, US;
Director, Gastroenterology/Hepatology Clinical Research Program

Michael P. Manns MD
Professor and Chairman
Department of Gastroenterology
Hepatology and Endocrinology
Hannover Medical School
Hannover, Germany

James G. Martin MD DSc
McGill University Health Centre Research Institute and McGill University

Craig L. Maynard, PhD
Assistant Professor
Department of Pathology
University of Alabama at Birmingham
Birmingham, AL USA

Samual McCash, MD
Clinical Chemistry Service
Department of Laboratory Medicine
Memorial Sloan Kettering Cancer Center
New York, NY, USA

Douglas R. McDonald MD, PhD
Assistant Professor of Immunology
Departments of Pediatrics and Immunology
Children's Hospital, Boston and Harvard Medical School
Boston, MA, USA

Peter C. Melby MD
Professor of Medicine
Departments of Internal Medicine, Microbiology and Immunology, and Pathology
University of Texas Medical Branch
Galveston, Texas, USA

Stephen D. Miller PhD
Professor
Department of Microbiology-Immunology
Northwestern University Medical School
Chicago, IL, USA

Anna L. Mitchell MB, BS, BMednsci, MRCP (UK)
Clinical Research Training Fellow
Institute of Genetic Medicine, International Centre for Life
Newcastle University
Newcastle-upon-Tyne, UK

Amirah Mohd-Zaki
UCL Institute of Ophthalmology
London, UK

Carolyn Mold PhD
Professor
Department of Molecular Genetics and Microbiology
University of New Mexico School of Medicine
Albuquerque, NM, USA

David R. Moller MD
Professor of Medicine
Johns Hopkins University School of Medicine
Baltimore, MD, USA

Dimitrios S. Monos PhD
Director, Immunogenetics Laboratory
The Philadelphia Children's Hospital of Philadelphia
Professor of Pathology and Lab Medicine
Perelman School of Medicine, University of Pennsylvania
Philadelphia, PA, USA

Scott N. Mueller, PhD
Associate Professor
Department of Microbiology and Immunology
Peter Doherty Institute for Infection and Immunity
The University of Melbourne, at Melbourne;
Australian Research Council Queen Elizabeth II Research Fellow, Department of Microbiology and Immunology, University of Melbourne
Parkville, VIC, Australia

Catharina M. Mulders-Manders MD PhD
Resident internal medicine
Radboudumc Expertise Center for Immunodeficiency and Autoinflammation
Department of Internal Medicine
Radboud University Medical Centre
Nijmegen, The Netherlands.

Mark J. Mulligan, MD, FIDSA
Distinguished Professor of Medicine
Executive Director
The Hope Clinic of the Emory Vaccine Center
Division of Infectious Diseases, Department of Medicine
Emory University School of Medicine
Atlanta, GA, USA

Ulrich R. Müller MD
Professor, Consultant
Spital Ziegler, Spital Netz Bern
Bern, Switzerland

Pashna N. Munshi, MD
John Theurer Cancer Center
Hackensack University Medical Center
Hackensack, NJ, USA;
Lombardi Comprehensive Cancer Center
MedStar Georgetown University Hospital
Washington, DC, USA

Kazunori Murata PhD, DABCC
Clinical Chemist
Memorial Sloan Kettering Cancer Center
New York, NY, USA

Philip M. Murphy MD
Chief, Laboratory of Molecular Immunology
National Institute of Allergy and Infectious Diseases
National Institutes of Health
Bethesda, MD, USA

Nicolás Navasa
CIC bioGUNE
Derio, Bizkaia, Spain

Pierre Noel MD
Professor of Medicine
Division of Hematology-Oncology
Mayo College of Medicine
Scottsdale, AZ, USA

Luigi D. Notarangelo MD
Jeffrey Modell Chair of Pediatric Immunology Research
Division of Immunology
Children's Hospital Boston
Professor of Pediatrics and Pathology
Harvard Medical School
Boston, MA, USA

Robert L. Nussbaum MD, FACP, FACMG
Volunteer Clinical Faculty
UCSF
Chief Medical Officer, Invitae Corporation

Thomas B. Nutman MD
Head, Helminth Immunology Section
Clinical Parasitology Section
Laboratory of Parasitic Diseases
National Institutes of Health
Bethesda, MD, USA

Stephen L. Nutt PhD
Professor
Molecular Immunology
The Walter and Eliza Hall Institute
Melbourne, VIC, Australia

João B. Oliveira MD, PhD
Head
Human Disorders of Lymphocyte Homeostasis Unit
National Institutes of Health
Assistant Chief, Immunology Service
Department of Laboratory Medicine at National Institutes of Health
Bethesda, MD, USA

Thomas L. Ortel MD, PhD
Chief, Division of Hematology
Professor of Medicine and Pathology
Medical Director, Clinical Coagulation Laboratory
Duke University Medical Center
Durham, NC, USA

John J. O'Shea MD
Scientific Director
Molecular Immunology and Inflammation Branch
National Institute of Arthritis and Musculoskeletal and Skin Diseases
Bethesda, MD, USA

Sung-Yun Pai MD
Assistant Professor of Pediatrics
Division of Pediatric Hematology-Oncology
Children's Hospital Boston
Department of Pediatric Oncology
Dana-Farber Cancer Institute
Harvard Medical School
Boston, MA, USA

Lavannya Pandit MD
Assistant Professor of Medicine
Department of Medicine
Baylor College of Medicine
Houston, TX, USA

Mary E. Paul MD
Associate Professor
Department of Pediatrics
Texas Children's Hospital
Houston, TX, USA

Simon H.S. Pearce MD, FRCP
Professor of Endocrinology
Institute of Genetic Medicine, International Centre for Life
Newcastle University
Newcastle-upon-Tyne, UK

Daniela Pedicino, MD
Department of Cardiovascular and Thoracic Sciences
Catholic University of the Sacred Heart
Rome, Italy

Erik J. Peterson MD
Center for Immunology
Department of Medicine
University of Minnesota
Minneapolis, MN, USA

Capucine Picard MD, PhD
Research Associate
Laboratory of Human Genetics of Infectious Diseases
Necker Branch, Imagine Institute
Paris Descartes University
Paris Sorbonne Cité, Paris, France;
Professor of Immunology
Director
Study Center for Primary Immunodeficiencies
Necker Hospital, Assistance Publique Hôpitaux de Paris
Paris, France

Stefania Pittaluga MD, PhD
Staff Physician
Laboratory of Pathology
National Cancer Institute
National Institutes of Health
Hematopathology Section
Bethesda, MD, USA

Debra Long Priel, MS
Associate Scientist
Neutrophil Monitoring Lab
Clinical Services Program
Applied Developmental Research Directorate
Leidos Biomedical Research, Inc.
Frederick National Laboratory for Cancer Research,
Frederick, MD, USA

Jennifer Puck, MD
Department of Allergy and Immunology
UCSF Pediatric
San Francisco, CA, USA

Anne Puel PhD
Research Associate
Laboratory of Human Genetics of Infectious Diseases
Necker Branch, Imagine Institute
Paris Descartes University
Paris Sorbonne Cité, Paris, France
St Giles Laboratory of Human Genetics of Infectious Diseases
Rockefeller Branch
The Rockefeller University
New York, NY, USA

Andreas Radbruch PhD
Scientific Director
Deutsches Rheuma-Forschungszentrum Berlin (DRFZ)
Leibniz Institute
Berlin, Germany

Stephen T. Reece PhD
Senior Scientist
Department of Medicine
University of Cambridge School of Clinical Medicine
Cambridge, United Kingdom

John D. Reveille MD
Professor of Medicine
Division of Rheumatology
Department of Medicine
University of Texas Health Science Center at Houston
Houston, TX, USA

Robert R. Rich MD
Professor of Medicine and Dean Emeritus
University of Alabama at Birmingham
Birmingham, AL, USA

Chaim M. Roifman MD, FRCPC
Professor
Department of Paediatrics and Immunology
The Hospital for Sick Children
Toronto, ON, Canada

Antony Rosen MD
Director
Division of Rheumatology
Johns Hopkins University School of Medicine
Baltimore, MD, USA

James T. Rosenbaum MD
Casey Eye Institute, Oregon Health and Science University
Dever's Eye Institute
Portland, OR, USA

Sergio D. Rosenzweig, MD, PhD
Chief, Immunology Service
Department of Laboratory Medicine
NIH Clinical Center
National Institutes of Health
Bethesda, MD, USA

Barry T. Rouse DVM, PhD, DSc
Distinguished Professor
Department of Pathobiology
University of Tennessee
Knoxville, TN, USA

Scott D. Rowley MD
Chief, Blood and Marrow Transplantation
John Theurer Cancer Center
Hackensack University Medical Center
Hackensack, NJ, USA

Shimon Sakaguchi. MD, PhD
Distinguished Professor, Experimental Immunology
Immunology Frontier Research Center (IFReC)
Osaka University
Suita, Osaka, Japan

Marko Salmi MD, PhD
Department of Molecular Medicine
Department of Medical Biochemistry and Genetics
University of Turku
Turku, Finland

Andrea J. Sant, PhD
Professor of Microbiology and Immunology
University of Rochester Medical School
David H. Smith Center for Vaccine Biology and Immunology
Rochester, NY, USA

Sarah W. Satola, PhD
Assistant Professor of Medicine
Division of Infectious Diseases
Department of Medicine
Emory University School of Medicine
Atlanta, GA, USA

Valerie Saw, MD FRCOphth PhD
Consultant Ophthalmic Surgeon and Clinical Lecturer
UCL Institute of Ophthalmology
London, UK

Marcos C. Schechter, MD
Infectious Diseases Fellow
Division of Infectious Diseases
Department of Medicine
Emory University School of Medicine
Atlanta, GA, USA

Harry W. Schroeder, Jr, MD, PhD
Director
UAB Program in Immunology
University of Alabama at Birmingham
Birmingham, AL, USA;
Professor of Medicine, Microbiology, and Genetics
Division of Clinical Immunology and Rheumatology

Benjamin M. Segal, M.D.
Holtom-Garrett Family Professor of Neurology and Director of the Multiple Sclerosis Center
Department of Neurology
University of Michigan Medical School
Ann Arbor, MI, USA

Carlo Selmi MD
Department of Medicine and Hepatobiliary Immunopathology Unit
IRCCS Istituto Clinico Humanitas
Department of Translational Medicine
University of Milan, Milan, Italy;
Assistant Professor of Medicine
Division of Rheumatology, Allergy and Clinical Immunology
University of California at Davis
Davis, CA, USA

Sushma Shankar BM, BCh, Ba Physiol Sci (Hons), MRCS
Academic Clinical Fellow
Department of Surgery
John Radcliffe Hospital
Oxford, UK

Anu Sharma
Postdoctoral Researcher
Texas MD Anderson Cancer Center
Houston, TX, USA

Padmanee Sharma
Genitourinary Medical Oncology
Department of Dermatology and Cutaneous Surgery
The University of Miami Miller School of Medicine
Texas MD Anderson Cancer Center
Houston, TX, USA

William T. Shearer MD, PhD
Allergy and Immunology Service
Texas Children's Hospital;
Professor of Pediatrics and Immunology
Section of Allergy and Immunology, Department of Pediatrics
Baylor College of Medicine
Houston, TX, USA

Richard M. Siegel MD, PhD
Chief, Immunoregulation Section
Autoimmunity Branch
NIAMS National Institutes of Health
Bethesda, MD, USA

Anna Simon MD, PhD
Associate Professor of Immunodeficiency and Autoinflammation
Department of General Internal Medicine
N4i Centre for Immunodeficiency and Autoinflammation (NCIA)
Radboud University Nijmegen Medical Centre
Nijmegen, The Netherlands

Gideon P. Smith MD, PhD
Director Connective Tissue Diseases
Department of Dermatology
Massachusetts General Hospital of Harvard University
Boston, MA, USA

David S. Stephens MD
Vice President for Research
Robert W. Woodruff Health Sciences Center
Emory University
Atlanta, GA, USA
Professor of Medicine, Microbiology & Immunology, and Epidemiology
Emory University School of Medicine
Atlanta, GA, USA

Robin Stephens PhD
Associate Professor
Departments of Internal Medicine, and Microbiology and Immunology
University of Texas Medical Branch
Galveston, TX, USA

Alex Straumann MD
Chairman, Swiss EoE Clinic and EoE Research Network
Olten, Switzerland

Leyla Y. Teos PhD
Molecular Biologist
National Institute of Dental and Craniofacial Research
National Institutes of Health
Bethesda, MD, USA

Laura Timares PhD
Associate Professor of Dermatology
Department of Dermatology
University of Alabama at Birmingham School of Medicine
Birmingham, AL, USA

Wulf Tonnus
Division of Nephrology
Department of Internal Medicine III University Hospital Carl Gustav Carus at the Technische Universität Dresden
Dresden, Germany

Raul M. Torres PhD
Professor of Immunology and Microbiology
University of Colorado School of Medicine
Aurora, CO, USA

Gülbü Uzel MD
Staff Clinician
Laboratory of Clinical Infectious Diseases, NIAID, NIH
Allergy & Immunology - Clinical & Laboratory Immunology and Pediatric Rheumatology
National Institute of Allergy and Infectious Diseases
National Institutes of Health
Bethesda, MD, USA

Jeroen C.H. van der Hilst MD
Internal Medicine Resident
Department of General Internal Medicine
Radboud University Nijmegen Medical Center
Nijmegen, The Netherlands

Jos W.M. van der Meer MD, PhD
Professor of Medicine, Head
Department of General Internal Medicine
Nijmegen Institute for Infection, Inflammation and Immunity
N4i Centre for Immunodeficiency and Autoinflammation (NCIA)
Radboud University Nijmegen Medical Centre
Nijmegen, The Netherlands

John Varga MD
John and Nancy Hughes Professor of Medicine
Division of Rheumatology
Northwestern University
Feinberg School of Medicine
Chicago, IL, USA

Jatin M. Vyas MD PhD
Department of Medicine, Division of Infectious Diseases, Massachusetts General Hospital;
Department of Medicine, Harvard Medical School
Boston, MA, USA

Meryl Waldman MD
Staff Clinician
Kidney Disease Section
National Institute of Diabetes and Digestive and Kidney Diseases
National Institutes of Health
Bethesda, MD, USA

Peter Weiser MD
Associate Professor of Pediatric Rheumatology
Division of Rheumatology
Department of Pediatrics
The University of Alabama at Birmingham
Birmingham, AL, USA

Peter F. Weller MD
William Bosworth Castle Professor of Medicine
Harvard Medical School
Beth Israel Deaconess Medical Center
Allergy and Inflammation Division
Boston, MA, USA

Cornelia M. Weyand MD, PhD
Professor of Medicine
Stanford University
Stanford, CA, USA

Fredrick M. Wigley MD
Professor of Medicine
Department of Medicine
Division of Rheumatology
Johns Hopkins University School of Medicine
Baltimore, MD, USA

Robert J. Winchester MD
Professor of Medicine
College of Physicians and Surgeons
Columbia University
New York, NY, USA

James B. Wing, PhD
Assistant Professor, Experimental Immunology
Immunology Frontier Research Center (IFReC)
Osaka University
Suita, Osaka, Japan

Kathryn J. Wood DPhil
Professor of Immunology
Nuffield Department of Surgical Sciences
University of Oxford
John Radcliffe, Hospital
Oxford, UK

Xiaobo Wu, MD
Assistant Professor of Medicine
Washington University School of Medicine
St. Louis, MO, USA

Hui Xu MD, PhD
Professor of Dermatology
University of Alabama at Birmingham
Birmingham, AL, USA

Cassian Yee, MD
Professor, Department of Melanoma Medical Oncology
Division of Cancer Medicine
Texas MD Anderson Cancer Center
Houston, TX, USA

Shen-Ying Zhang MD, PhD
Research Associate
Laboratory of Human Genetics of Infectious Diseases
Necker Branch, Imagine Institute;
Paris Descartes University,
Paris Sorbonne Cité, Paris, France;
St. Giles Laboratory of Human Genetics of Infectious Diseases
Rockefeller Branch
The Rockefeller University
Research Associate St Giles Laboratory of Human Genetics of Infectious Diseases
New York, NY, USA

DEDICATION

To:
Susan Rich, Cathryn Rich and Darren Selement,
Phoebe and Kenneth Rich,
Lynn and Kenneth Todorov
Robert R. Rich

With thanks and gratitude to my family and my colleagues.
Thomas A. Fleisher

Lynn Des Prez and Christine, Mark, Christopher, Martin, John, Jesse and Melissa Shearer
William T. Shearer

Dixie Lee Schroeder; Harry W Schroeder III MD PhD, Maria Isabel, and Anabel Schroeder; Elena, Jeff and Liam Beck; Jeannette Schroeder and Antoni Bernard
Harry W. Schroeder, Jr

To my family for their support and encouragement and to my students for their interest in immunology and allergy
Anthony J. Frew

Jörg Goronzy and Dominic and Isabel Weyand Goronzy
Cornelia M. Weyand

1

The Human Immune Response

Robert R. Rich, David D. Chaplin

Clinical immunology is a medical subspecialty largely focused on a specific physiological process, inflammation, which is essential to good health, particularly in defense against pathogenic organisms, recovery from injury, and containment of neoplasms. But inflammation, mediated by the cells and soluble products of the immune system, is also a powerful contributor to the pathogenesis of diseases that affect virtually every organ system. A consequent challenge for clinical immunologists, both clinicians and basic scientists, is to reduce a dizzying array of disease descriptions to systematic understanding of pathogenic mechanisms, facilitating translation of fundamental concepts and new discoveries into more effective disease prevention or treatment.

This introductory chapter is directed to nonimmunologist clinicians and researchers. It is structured as an introduction to the interacting elements of the human immune system and their disordered functions in diseases. The subtleties, including immunological or molecular genetic jargon unavoidably used, are described in detail in the chapters that follow.

THE HOST–MICROBE INTERACTION

The vertebrate immune system is a product of eons of evolutionary relationships between rapidly evolving microbial organisms and their much less rapidly reproducing, and hence less adaptable, hosts.[1] In general, the relationship is mutually beneficial, each providing nutrients and other materials essential to the well-being of their partner—the host and its microbiome (Chapter 14). Occasionally, however, a normally beneficial relationship becomes pathological, as pathogenic microbes overwhelm the microbiome, invading host tissues and resulting in host morbidity or even death. Because the vertebrate host cannot win a battle with microbial invaders by rapid mutation and selection, the immune system employs a strategy of complexity and redundancy, which involves both the individual organism and its collective population. Reflecting plasticity of the response, specific defenses differ, depending on the nature of the infectious agent and its point of entry and distribution within the body. Regardless of the defense mechanism, an intended outcome is destruction or neutralization of the invading organism. A secondary consequence, however, can be collateral damage to host cells. These cells can be targeted for damage because they are sites of microbial residence and replication, or they can be damaged as "innocent bystanders." Depending on the site and severity of the host's defensive response, it may be accompanied by local and/or systemic symptoms and signs of inflammation, which may lead to long-lasting tissue dysfunction as a result of tissue remodeling and partial repair.

Adaptive and Innate Immunity

Immune responses are traditionally classified as *adaptive* (also termed *acquired* or *specific*) and *innate* (or *nonspecific*) (Table 1.1). The adaptive immune system, present uniquely in species of the phylum *Chordata*, is specialized for development of an inflammatory response based on recognition of specific "foreign" macromolecules that are predominantly, but not exclusively, proteins, peptides, and carbohydrates. The vast majority of chordate species are vertebrates, and this book addresses adaptive immunity of that subphylum. Its primary effectors are antibodies, B lymphocytes, T lymphocytes, and antigen-presenting cells (APCs). T and B lymphocytes express surface antigen receptors that are clonally specific as a consequence of receptor–gene rearrangements. Expansion of clones of lymphocytes specific for any particular antigen is induced by antigen encounter and consequent activation and proliferation, thereby constituting the basis of immunological memory.

Innate immune responses are phylogenetically far more ancient, being widely represented in multicellular phyla.[2] Rather than being based on exquisitely specific recognition of a diverse array of macromolecules (*i.e.*, antigens), they are focused on recognition of common molecular signatures of microbial organisms that are not present in vertebrates.[3] Among these structures, which are termed *pathogen-associated molecular patterns* (PAMPs) or *danger-associated molecular patterns* (DAMPs), are bacterial cell wall constituents, such as mannose-rich oligosaccharides, lipopolysaccharides, peptidoglycans, and several nucleic acid variants, including double-stranded RNA and unmethylated CpG DNA. For both innate and adaptive immunity responses, defense effector mechanisms can require either direct cell-to-cell contact or the activity of cytokines and chemokines, which are hormone-like soluble molecules that act in the cellular microenvironment (cell-mediated immunity). Most immune responses include participation of both modes of response.[2-4]

The elements of innate immunity are diverse (Chapter 3). They include physical barriers to pathogen invasion (*e.g.*, skin, mucous membranes, cilia, and mucus), as well as an array of cellular and soluble factors that can be activated by secreted or cell surface products of the pathogen, including PAMPs. Recognition of PAMPs by cells in innate immunity, which also commonly function as APCs to the lymphocytes of adaptive immunity, is via cell membrane or cytoplasmic receptors known as pattern recognition receptors (PRRs). PRRs can be either membrane bound or cytoplasmic. Membrane-bound PRRs include Toll-like receptors (TLRs) and C-type lectin receptors (CLRs). Humans express 10 distinct TLRs, which recognize (among others) specific bacterial glycolipids, lipopolysaccharide; viral single-stranded

TABLE 1.1 Features of Innate and Adaptive Immune Systems

Distinguishing Features

Innate Immunity	Adaptive Immunity
Germ line-encoded receptors targeting pathogen molecular patterns	Clonally variable receptors generated somatically by rearrangement of gene elements
Does not require immunization	Consequence of B- and/or T-cell activation
Limited memory	Immunological memory well developed
Includes physical barriers to pathogen	Antibody and cytotoxic T cells

Common Features

Cytokines and chemokines
Complement cascade
Phagocytic cells
Natural killer (NK) cells
"Natural" antibodies

RNA; and bacterial and viral unmethylated CpG DNA. CLR are particularly important in antifungal innate immunity but also have important roles in defenses against bacteria, viruses, and parasites. They comprise a large family that commonly recognizes microbe-specific carbohydrate ligands or structurally similar lectin-like domains. Cytoplasmic PRRs include RIG-1–like receptors (RLRs) and nucleotide oligomerization domain (NOD)-like receptors (NLRs). RLRs are involved in recognition of viruses through interaction with intracytoplasmic viral double-stranded RNA (dsRNA), and NLRs recognize bacterial peptidoglycan motifs.[4]

Cells of the innate immune system are commonly triggered through activation of the NF-κB transcription factor via the MyD88 signaling pathway, thereby inducing an inflammatory response using mechanisms that are broadly shared with those of the adaptive immune system. These include activation of various types of innate lymphoid cells (*e.g.,* natural killer [NK] cells), which are characterized by absence of clonally expressed receptors for specific antigen (see below), activation of granulocytes and other phagocytes, the secretion of inflammatory cytokines and chemokines, and interactions of the many participants in the complement cascade. Additionally, activation of cells of innate immunity that also act as APCs for the adaptive immune system results in upregulation of membrane molecules (*e.g.*, CD80, CD86) that provide the second signal, along with the T-cell receptor (TCR) for antigen, necessary for induction of antigen-specific T cells.[5]

Finally, because recognition of pathogens by the innate immune system relies on germline encoded, nonrearranged receptors held in common by the specific cell type, innate immunity is more rapidly responsive. It can initiate in minutes to hours and generally precedes development of a primary adaptive immune response by at least several days.

CELLS OF THE IMMUNE SYSTEM

The major cellular constituents of both innate and adaptive immunity originate in bone marrow, where they differentiate from multipotent hematopoietic stem cells (HSCs) along several pathways to become granulocytes, lymphocytes, and APCs (Chapter 2).

Granulocytes

Polymorphonuclear leukocytes (granulocytes) are classified by light microscopy into four types. By far the most abundant in the peripheral circulation are neutrophils, which are principal effector cells linking the innate and adaptive responses by virtue of their expression of surface receptors for antibody and complement (Chapter 21). They are phagocytic cells that ingest, kill, and degrade microbes and other targets of an immune attack within specialized cytoplasmic vacuoles that contain potent antimicrobial enzymes and oxidative pathways. The phagocytic activity of neutrophils is promoted by their surface display of receptors for antibody molecules (specifically the Fc portion of immunoglobulin G [IgG] molecules) (Chapter 15) and activated complement proteins (particularly the C3b component) (Chapter 21). Neutrophils are the predominant cell type in acute inflammatory infiltrates and are the primary effector cells in immune responses to pyogenic bacteria (Chapter 27).

Eosinophils (Chapter 24) and basophils (Chapter 23) are the other circulating forms of granulocytes. A close relative of the basophil, but derived from distinct bone marrow precursors, is the tissue mast cell, which does not circulate in blood. Eosinophils, basophils, and mast cells are important in defenses against multicellular pathogens, particularly helminths (Chapter 31). Their defensive functions are not based on phagocytic capabilities but, rather, on their ability to discharge potent biological mediators from their storage granules into the cellular microenvironment. This process, termed *degranulation*, can be triggered by antigen-specific IgE molecules that bind to basophils and mast cells via high-affinity receptors for the Fc portion of IgE (FcεR) on their surfaces. In addition to providing a mechanism for anthelmintic host defenses and certain antibacterial responses, this is also the principal mechanism involved in acute (IgE-mediated) allergic reactions (Chapters 41–49).

Lymphocytes

Three broad categories of lymphocytes are identified on the basis of display of particular surface molecules: B cells, T cells, and innate lymphoid cells; each of these categories can be further subdivided according to specific function and display of distinguishing cell surface molecules (Chapter 2). All lymphocytes differentiate from common lymphoid stem cells in bone marrow. B cells create their immunoglobulin receptors in bone marrow and differentiate into antibody-producing cells in the periphery (Chapter 7). T-cell precursors move from bone marrow to the thymus (or, in some cases, to extrathymic tissue compartments), where they complete their differentiation and selection (Chapter 8).

T cells and B cells are the heart of immune recognition, a property reflecting their clonally specific cell surface receptors for antigen (Chapter 4). The TCR is a heterodimeric integral membrane molecule expressed exclusively by T lymphocytes. B-cell receptors for antigen (BCRs) are membrane immunoglobulin (mIg) molecules of the same antigenic specificity that the cell and its terminally differentiated progeny, plasma cells, will secrete as soluble antibodies. Memory B cells and nondividing, long-lived plasma cells may account substantially for persistence of antibody responses (including production of autoantibodies) over many years.[6]

Receptors for "antigen" on the third class of lymphocytes, innate lymphoid cells (ILCs), are not clonally expressed. ILCs are subdivided into three major groups according to the cytokines that they produce (*e.g.,* group 1 ILCs, including NK cells, produce interferon-γ [IFN-γ] and tumor necrosis factor [TNF]).[7] ILCs express receptors for PAMPs and, as such, serve as major effectors of innate immunity. They also recognize target cells that might otherwise elude the immune system (Chapters 2, 17). Thus recognition of NK cell targets is based substantially on what their targets lack rather than on what they express. NK cells express receptors of several types for major histocompatibility complex (MHC) class I molecules via killer immunoglobulin-like receptors (KIRs).[8] KIRs are expressed on the plasma membrane of NK cells (and some T cells), which interact with class I molecules to alter NK-cell cytotoxic function. Most KIRs express in their intracellular domain a tyrosine-based inhibitory-motif (ITIM) that suppresses NK activity, thereby preventing NK cell activity directed against normal self-cells. In contrast, some KIRs express a tyrosine-based activation motif (ITAM), which amplifies their activity. NK cells will kill target cells unless they receive an inhibitory signal transmitted by an ITIM receptor. Virus-infected cells and tumor cells that attempt to escape T-cell recognition by downregulating their expression of class I molecules become susceptible to NK cell–mediated killing because the NK cells receive an activation signal and/or fail to receive an inhibitory signal through the ITAM- and ITIM-containing MHC class I receptors. The balance between ITIM and ITAM is regulated by the microenvironmental milieu, increasing expression of ITAM in the presence of viral-infected or cancer cells and of ITIM as necessary to maintain self-tolerance and prevent autoimmunity. A high frequency of ITAM-expressing cells has been reported in some patients with autoimmune diseases.[9]

Although NK cell–mediated innate immunity has been long considered to lack immunological memory, recent studies suggest that NK cells can exhibit memory of previous encounters with microbes or other antigens, the molecular basis of which remains to be fully elucidated.[10] NK cells can also participate in antigen-specific immune responses by virtue of their surface display of the activating ITAM receptor CD16, which binds the constant (Fc) region of IgG molecules. This enables them to function as effectors of a cytolytic process termed *antibody-dependent cellular cytotoxicity* (ADCC), a mechanism exploited clinically with monoclonal antibody (mAb) therapeutic agents.[11]

In general, pathways leading to differentiation of T cells, B cells, and ILCs are mutually exclusive, representing a permanent lineage commitment. No lymphocytes express both mIg and TCRs. However, a subset of T cells, termed *NKT cells*, exhibit both NK-like cytotoxicity and αβTCR with limited receptor diversity.

Antigen-Presenting Cells

KEY CONCEPTS

Features of Antigen-Presenting Cells

- Capacity for uptake and partial degradation of protein antigens
- Expression of major histocompatibility complex (MHC) molecules for binding antigenic peptides
- Chemokine receptors to allow colocalization with T cells
- Expression of accessory molecules for interaction with T cells
- Receptors for pathogen- or danger-associated molecular patterns
- Secretion of cytokines that program T helper (Th) cell responses

A morphologically and functionally diverse group of cells, all of which are derived from bone marrow precursors, is specialized for presentation of antigen to lymphocytes, particularly T cells (Chapter 6). Included among such cells are dendritic cells, monocytes (present in the peripheral circulation), macrophages (solid tissue derivatives of monocytes), cutaneous Langerhans cells (Chapter 19), and constituents of the reticular endothelial system within solid organs. B lymphocytes that specifically capture antigen via their clonally expressed mIg can also function efficiently in antigen presentation to T cells.

Cardinal features of APCs include their expression of both class I and class II Major Histocompatibility Complex (MHC; Chapter 5) molecules as well as requisite accessory molecules for T-cell activation (*e.g.*, B7-1, B7-2/CD80, CD86).[12] Upon activation, APC elaborate cytokines that induce specific responses in cells to which they are presenting antigen. In addition to processing and presenting antigen, APCs can regulate activation of the immune system via innate cell surface receptors, which contribute to determination of whether the antigen is pathogen associated.

APCs differ substantially among themselves with respect to mechanisms of antigen uptake and effector functions. Immature dendritic cells show high phagocytic and pathogen-killing activity but low ability to present antigen and activate T cells.[12] Dendritic cells (DCs) that have ingested a pathogen or foreign antigen can be induced to mature by inflammatory stimuli, especially via cells of the innate immune system and by direct activation through receptors for PAMPs or DAMPs.[12,13] Monocytes and macrophages are actively phagocytic, particularly for antibody and/or complement-coated (opsonized) antigens that bind to their surface receptors for IgG and C3b. These cells are also important effectors of immune responses, especially in sites of chronic inflammation. Upon further activation by T-cell cytokines, they can kill ingested microorganisms by oxidative pathways similar to those employed by polymorphonuclear leukocytes.

The interaction between B cells acting as APCs and T lymphocytes is notable as the cells are involved in a mutually amplifying circuitry of antigen presentation and response.[12] The process is initiated by antigen capture through B cell mIg and ingestion by receptor-mediated endocytosis. This is followed by proteolytic antigen degradation and then display to T cells as oligopeptides bound to MHC molecules. Like other APCs, B cells display CD80, which provides a requisite second signal to the antigen-responsive T cell via CD28, its accessory molecule for activation (Fig. 1.1). As a result of T-cell activation, T-cell cytokines that regulate B-cell differentiation and antibody production are produced, and T cells are stimulated to display the surface ligand CD40L (CD154), which can serve as the second signal for B-cell activation through its inducible surface receptor

BASIS OF ADAPTIVE IMMUNITY

The essence of adaptive immunity is molecular distinction between self constituents and potential pathogens (for simplicity, self/nonself discrimination, but perhaps more precisely discrimination between molecular species perceived as signaling potential "danger" and those that do not). This discrimination is a major responsibility of both T lymphocytes and cells of the innate immune system. It reflects the selection of thymocytes that have generated specific antigen receptors, which, upon later encounter, can bind nonself antigenic peptides bound to self-MHC molecules. The consequence of this selection process is that foreign proteins are recognized as antigens, whereas self-proteins are tolerated (*i.e.*, are not perceived as antigens). Additionally, the cells of innate immunity contribute importantly, through PAMPs/DAMPs and several mechanisms still being defined, to the essential

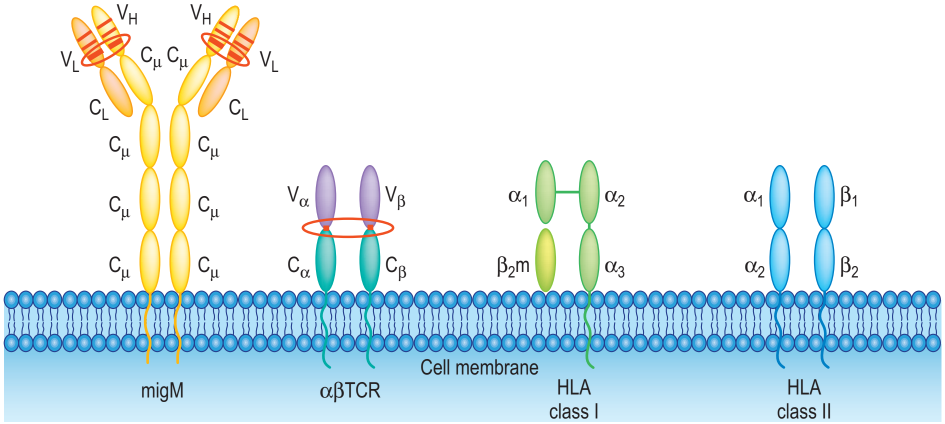


FIG 1.1 Antigen-Binding Molecules. Antigen-binding pockets of immunoglobulin (Ig) and T-cell receptor (TCR) comprise variable (V) segments of two chains translated from transcripts that represent rearranged V(D)J or VJ gene elements. Thin red bars designate two of the complementarity determining regions (CDRs) that form portions of the Ig antigen-binding site. The red ovals with thick red bars designate the regions of very high sequence variability in both Igs and TCRs that are generated by recombination of the 3′-end of the V gene element with the D and J gene elements or with the J gene element. In the Ig molecule this is designated CDR3. Antigen-binding pockets of Ig molecules are formed by the three-dimensional folding of the heavy and light chains that juxtapose the CDRs of one heavy chain and one light chain. Antigen-binding grooves of MHC molecules are formed with contributions from α_1 and β_1 domains of class II molecules and from α_1 and α_2 domains of class I molecules. All of these molecules are members of the immunoglobulin superfamily. C, constant-region domain; β_2, beta-2 microglobulin; mIgM, membrane immunoglobulin M; HLA, human leukocyte antigen; MHC, major histocompatibility complex.

distinction between commensal (not dangerous) and potentially pathogenic (dangerous) microbes.

T lymphocytes generally recognize antigens as a complex of short linear peptides bound to self-MHC molecules on the surfaces of APCs (Chapter 6). The source of these peptides can be either extracellular or intracellular proteins and derived from either self or foreign (*e.g.,* microbial) molecules. With the exception of superantigens (see below), T cells neither bind antigen in native conformation nor recognize free antigen in solution. The vast majority of antigens for T cells are oligopeptides. However, the antigen receptors of NKT cells can recognize lipid and glycolipid antigens that are presented to them by MHC-like CD1 molecules.[14]

Antigen recognition by T cells differs fundamentally from that by antibodies, which are produced by B lymphocytes and their derivatives. Antibodies are oriented toward recognition of extracellular threats and, unlike T cells, can bind complex macromolecules and can bind them in their native conformation either at cell surfaces or in solution. Moreover, antibodies show less preference for recognition of proteins; antibodies against carbohydrates, nucleic acids, lipids, and simple chemical moieties can be readily produced. Although B cells can also be rendered unresponsive by exposure to self-antigens, particularly during differentiation in bone marrow, this process does not define foreignness within the context of self-MHC recognition.

Clonal Basis of Immunological Memory

An essential element of self/nonself discrimination is the clonal nature of antigen recognition. Although the immune system can recognize a vast array of distinct antigens, all of the receptors of a single T cell or B cell (and their clonal progeny) have identical antigen-binding sites and hence a particular specificity (Chapter 4). A direct consequence is the capacity for antigen-driven immunological memory. This phenomenon derives from the fact that after an initial encounter with antigen clones of lymphocytes that can recognize the antigen proliferate and differentiate into effector cells, most of which ultimately are consumed or undergo apoptosis, and a smaller population of long-lived memory cells. These memory cells constitute a pool of cells larger than the initial naïve responders. They can elicit a greater and more rapid response upon subsequent antigen encounter. These two hallmarks of adaptive immunity, clonal specificity, and immunological memory provide a conceptual foundation for the use of vaccines in prevention of infectious diseases (Chapter 90). Immunological memory involves not only the T cells charged with antigen recognition but also the T cells and B cells that mediate the efferent limb of an inflammatory response. In its attack on foreign targets, the immune system can exhibit exquisite specificity for the inducing antigen, as is seen in the epitope-specific lysis of virus-infected target cells by cytolytic T cells.

ANTIGEN-BINDING MOLECULES

KEY CONCEPTS

Features of the Immunoglobulin (Ig) Superfamily

- Large family of ancestrally related genes (more than 100 members)
- Most products involved in immune system function or other cell–cell interactions
- Ig superfamily members have one or more domains of ~100 amino acids, each usually translated from a single exon
- Each Ig domain consists of a pair of β-pleated sheets usually held together by an intrachain disulfide bond

Three sets of molecules are responsible for the specificity of adaptive immune responses by virtue of their capacity to bind foreign antigen. These molecules are Igs, TCRs, and MHC molecules (see Fig. 1.1) (Chapters 4, 5). All are products of a very large family of ancestrally related genes, the immunoglobulin superfamily, which includes many other molecules essential to induction and regulation of immune responses.[15,16] Members of the Ig superfamily exhibit characteristic structural features. The most notable of these is organization into homologous domains of approximately 110 amino acids that are usually encoded by a single exon with an intradomain disulfide bond, characteristically configured as antiparallel strands, forming two opposing β-pleated sheets.

Immunoglobulins and T-Cell Receptors

The remarkable specificity of Ig and TCR molecules for antigen is achieved by a mechanism of genetic recombination that is unique to Ig and TCR genes (Chapter 4). The antigen-binding site of both types of molecules comprises a groove formed by contributions from each of two constituent polypeptides. In the case of immunoglobulins, these are a heavy (H) chain and one of two alternative types of light (L) chains, κ or λ. In the case of TCRs, either of two alternative heterodimers can constitute the antigen-binding molecule, one comprised of α and β chains, and the other of γ and δ chains. The polypeptides contributing to both Igs and TCRs can be divided into an antigen-binding amino-terminal variable (V) domain and one or more carboxy-terminal constant (*i.e.,* nonvariable) domains. Ig constant region domains generally include specific sites responsible for the biological effector functions of the antibody molecule (Chapter 15).

KEY CONCEPTS

Comparison of T-Cell and B-Cell Receptors for Antigen

Similarities

- Members of the immunoglobulin (Ig) superfamily
- Two polypeptide chains contribute to antigen-binding site
- Each chain divided into variable and constant regions
- Variable regions constructed by V(D)J rearrangements
- Nongenomic N-nucleotide additions at V(D)J junctions
- Exhibit allelic exclusion
- Negative selection against receptors with self-antigen specificity
- Transmembrane signaling involving coreceptor molecules

Differences

- Ig can be secreted; T-cell receptor (TCR) is not
- Ig recognizes conformational antigen (Ag) determinants; TCR recognizes linear determinants
- Ig can bind antigen in solution; TCR binds antigen when presented by major histocompatibility complex (MHC) molecule on antigen-presenting cell (APC)
- Somatic hypermutation of Ig genes can enhance antigen-binding affinity
- Ig genes can undergo isotype switching
- Inflammatory effector functions by the Ig constant domains
- Positive selection of TCR for self-MHC recognition

The most noteworthy feature of the vertebrate immune system is the process of genetic recombination that generates a virtually limitless array of specific antigen receptors from a rather limited genomic investment. This phenomenon is accomplished by the recombination of genomic segments that encode the variable domains of Ig and TCR polypeptides[17] (Chapter 4). The products of these rearranged gene elements provide a specific B or T cell with its unique antigen receptor. The variable domain of the mature receptor is created by the rearrangement of two or three separate gene segments. These are designated V (variable) and J (joining), for IgL chains and TCR α and γ chains, and V, D (diversity) and J, for IgH and TCR β and δ chains. In addition to rearrangement, N-nucleotide addition also contributes substantially to receptor diversity. N-nucleotide addition results in the insertion, at the time of rearrangement, of one or more nongenomic nucleotides at the junctions between V, D, and J segments through the action of terminal deoxynucleotidyl transferase (TdT).[17] This permits receptor diversity to extend beyond germline constraints. Analysis of the linear sequences of many Ig V regions domains has shown that they contain three sites of high sequence variability that have been designated *complementarity determining regions 1–3 (CDR1–3)* to indicate that they are the sites that contact antigen (see Fig. 1.1).

DNA rearrangement involved in generating T- and B-cell receptors is controlled by recombinases that are active in early thymocytes and in B precursor cells in bone marrow. The process is sequential and carefully regulated, generally leading to translation of one receptor of unique specificity for any given T or B lymphocyte. This result is achieved through a process termed *allelic exclusion*, wherein only one member of a pair of allelic genes potentially contributing to an Ig or TCR molecule is rearranged at a time.[18]

The process of allelic exclusion is not absolute, and a small number of lymphocytes will express dual functional Ig or TCR transcripts and, in some cases, two distinct surface receptors.[19] But B cells exclusively rearrange Ig genes, not TCR genes, and vice-versa for T cells. Moreover, after producing a functional heavy chain, B cells sequentially rearrange L chain genes, typically κ before λ. Thus B cells express either κ or λ chains, but not both. Similarly, thymocytes express α and β genes or γ and δ genes, and only rarely T cells with αδ or γβ receptors.

There is one feature of V region construction that is essentially reserved to B cells. This is somatic hypermutation (SHM), a process that can continue at discrete times throughout the life of a mature B cell at both the $V_HD_HJ_H$ and V_LJ_L gene exons.[20] Because these rearranged gene exons encode the binding groove that contains the specific points of contact with antigen, on occasion the random process of SHM will result in cells expressing mIg with increased affinity for the antigen they recognize. Typically, cells with increased affinity for antigen are activated preferentially, particularly at limiting doses of antigen. Thus the average affinity of antibodies produced during the course of an immune response tends to increase, a process termed *affinity maturation*.

TCRs do not show evidence of SHM. This absence may be related to the focus on selection in the thymus involving corecognition of a self-MHC molecule and self-peptides,[21] (Chapter 8) rather than the continuous process of antigen-driven selection in the periphery by B cells after SHM. Thymic selection results in deletion by apoptosis of the vast majority of differentiating thymocytes by mechanisms that place stringent boundaries around the viability of a thymocyte with a newly expressed TCR specificity. Once a T cell is fully mature and ready for emigration from the thymus, its TCR is essentially fixed, reducing the likelihood of emergent autoimmune T-cell clones in the periphery.

Receptor Selection

The receptor expressed by a developing thymocyte must be capable of binding with low-level affinity to some particular MHC self-molecule, either class I or class II, expressed by a resident thymic epithelial cell or APC. Because their receptors are generated by a process of semirandom joining of rearranging exon segments coupled with N-nucleotide additions, most thymocytes fail this test. They are consequently deleted as not being useful to an immune system that requires T cells to recognize antigen that is bound to self-MHC molecules. Thymocytes surviving this hurdle are said to have been "positively selected"[21] (Fig. 1.2A). Conversely, a small number of thymocytes bind with an unallowably high affinity for a combination of MHC molecule plus antigenic peptide expressed by a thymic APC. Because the peptides available for MHC binding at this site are derived almost entirely from self-proteins, differentiating thymocytes with such receptors are intrinsically dangerous as potentially autoimmune. This deletion of thymocytes with high-affinity receptors for self-MHC plus (presumptively) self-peptide is termed "negative selection"[21] (Fig. 1.2B), a process that may also involve activity of regulatory T cells (Tregs).[22,23]

Another feature that distinguishes B cells from T cells is that the cell surface antigen receptors of the former can be secreted in large quantities as antibody molecules, the effector functions of which are carried out in solution or at the surfaces of other cells. Secretion is accomplished by alternative splicing of Ig messenger RNAs (mRNAs) to include or exclude a transmembrane segment that is encoded by the Ig heavy-chain genes.

Immunoglobulin Class Switching

In addition to synthesizing both membrane and secreted forms of Igs, B cells also undergo class switching. Antibody molecules are comprised of five major classes (isotypes). In order of abundance in serum, these are IgG, IgM, IgA, IgD, and IgE (Chapters 4 and 15). In humans the IgG class is further subdivided into four subclasses and the IgA class into two subclasses. The class of Ig is determined by the sequence of the constant region of its heavy chain (C_H). The H chain constant region gene locus is organized with exons that encode each of the Ig isotypes and subclasses located downstream (3') of the variable (V_H) genes. Thus an antibody-producing cell with a successfully rearranged $V_HD_HJ_H$ exon can change the class of antibody molecule that it synthesizes by utilization of different C_H genes without changing its unique antibody specificity. This process, termed *class switch recombination*, is regulated by cytokines and is accomplished through the action of activation-induced cytidine deaminase.[24]

There is no process comparable with class switch recombination in T cells. The two types of TCRs are products of four

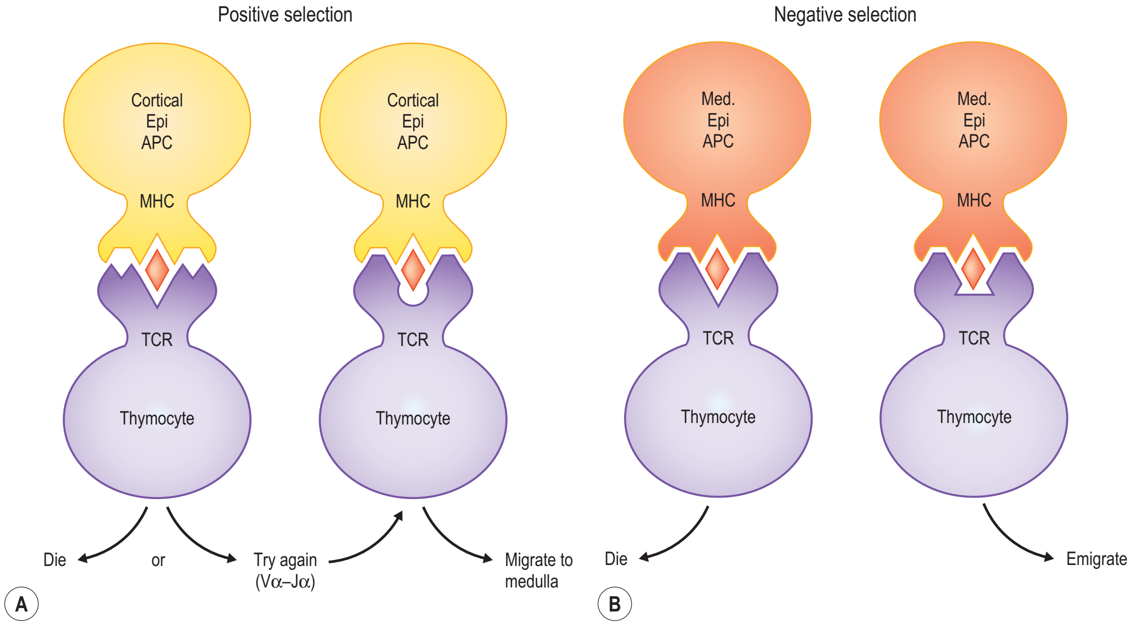


FIG 1.2 Two-Stage Selection of Thymocytes Based on Binding Characteristics of Randomly Generated T-Cell Receptors (TCRs). (A) *Positive selection.* "Double-positive" ($CD4^+$, $CD8^+$) thymocytes with TCRs capable of *low* avidity binding to some specific self-MHC molecule (either class I or class II) expressed by thymic cortical epithelial cells are positively selected. This process may involve sequential attempts at α gene rearrangement in order to express an αβ TCR of appropriate self-MHC specificity. If binding is to a class I molecule, the positively selected thymocyte becomes CD8 single-positive, and if to a class II molecule, a CD4 single-positive. Thymocytes that are unsuccessful in achieving a receptor with avidity for either a class I or a class II self-MHC molecule die by apoptosis. The solid diamond represents a self-peptide derived from hydrolysis of an autologous protein present in the thymic microenvironment or synthesized within the thymic epithelium itself. (B) *Negative selection.* "Single-positive" ($CD4^+$ *or* $CD8^+$) thymocytes, positively selected in the thymic cortex, that display TCRs with *high* avidity for the combination of self-MHC plus some self (autologous) peptide present in the thymus are negatively selected (*i.e.,* die) as potentially "autoimmune." Those few thymocytes that have survived both positive and negative selection emigrate to the periphery as mature T cells.

independent sets of V-region and C-region genes. A large majority of peripheral blood T cells express αβ TCRs, with a small fraction expressing γδ TCRs (usually ≤5% in peripheral blood). There is a higher representation of γδ T cells in certain tissues, particularly those lining mucous membranes, where they may be specialized for recognition of heavily glycosylated peptides or nonpeptide antigens that are commonly encountered in these tissue compartments. Thymocytes are committed to the expression of either αβ or γδ TCR, and their differentiated progeny (T cells) never change their TCR type in the periphery.

Major Histocompatibility Complex

MHC molecules constitute a third class of antigen-binding molecules. When an MHC class I molecule was initially crystallized, an unknown peptide was found in a binding groove formed by the first two (α_1 and α_2) domains of the molecule. This binding groove has since been established as a general feature of MHC molecules.[25] It is now known that the function of MHC molecules is to present antigen to T cells in the form of oligopeptides that reside within this antigen-binding groove (Chapter 6). The most important difference between the nature of the binding groove of MHC molecules and those of Ig and TCR is that the former does not represent a consequence of gene rearrangement. Rather, all the available MHC molecules in an individual are encoded in a linked array, which in humans is located on chromosome 6 and designated the human leukocyte antigen (HLA) complex.

MHC molecules are of two basic types, class I and class II. Class I molecules are found on the surface of almost all somatic cells, whereas cell surface expression of class II genes is restricted primarily to cells specialized for APC function. Class I molecules have a single heavy chain that is an integral membrane protein comprised of three external domains (see Fig. 1.1). The heavy chain is noncovalently associated with β_2 microglobulin, a nonpolymorphic, non–membrane-bound, single-domain Ig superfamily molecule that is encoded in humans on chromosome 15, not linked to the MHC. Class II MHC molecules, in contrast, comprise two polypeptide chains, α and β (or A and B), of approximately equal size, each of which consists of two external domains connected to a transmembrane region and cytoplasmic tail. Both chains of class II molecules are anchored on the cell by a transmembrane domain, and both are encoded within the MHC. Class I and class II molecules have a high degree of structural homology, and both fold to form a peptide-binding groove on their exterior face, with contribution from the α_1 and α_2 domains for class I molecules and from α_1 and β_1 domains for class II.[25]

There are three class I loci (HLA-A, -B, and -C) and three class II subregions (HLA-DR, -DQ, and -DP) that are principally involved in antigen presentation to T cells (Chapter 5). The functions of other class I and class II genes within this complex are less clear. Some, at least, appear to be specialized for binding (presentation) of peptide antigens of restricted type, source, or function (*e.g.*, HLA-E),[26] and others (*e.g.*, HLA-DM and HLA-DO) are clearly involved in antigen processing and loading of antigenic peptides into the binding cleft of the HLA-DR, -DQ, and -DP molecules [27] (Chapter 6). Additionally, members of a family of "nonclassic" class Ib molecules, $CD1_{a-d}$, which are encoded on chromosome 1, outside the MHC, are specialized for binding and presentation of lipid and lipid-conjugate antigens to T cells.[14,28]

The HLA complex represents an exceedingly polymorphic set of genes (Chapter 5). Consequently, most individuals are heterozygous at each major locus. In contrast to TCRs and Igs, the genes of the MHC are codominantly expressed. Thus, at a minimum, an APC can express six class I molecules and six class II molecules (the products of the two alternative alleles of three class I and three class II loci). This number is, in fact, usually an underestimate, as a consequence of additional complexity in the organization of the class II region (Chapter 5).

ANTIGEN PRESENTATION

Because MHC genes do not undergo recombination, the number of distinct antigen-binding grooves that they can form is many orders of magnitude less than that for either TCRs or Igs. Oligopeptides that bind to MHC molecules are the products of self or foreign proteins. They are derived by hydrolytic cleavage within APCs and are loaded into MHC molecules before expression at the cell surface (Chapter 6). Indeed, stability of MHC molecules at the cell surface requires the presence of a peptide in the antigen-binding groove; cells mutant for the loading of peptide fragments into MHC molecules fail to express MHC molecules on their cell surfaces.[29] Since in the absence of infection most hydrolyzed proteins are of self-origin, the binding groove of most MHC molecules contains a self-peptide. Class I and class II molecules differ from one another in the length of peptides that they bind, usually 8–9 amino acids for class I and 14–22 amino acids for class II.[25] Although important exceptions are clearly demonstrable, they also generally differ with respect to the source of peptide. Those peptides binding to class I molecules usually derive from proteins synthesized intracellularly (*e.g.*, autologous proteins, tumor antigens, virus proteins, and proteins from other intracellular microbes), whereas class II molecules commonly bind peptides derived from proteins synthesized extracellularly (*e.g.*, extracellular bacteria, nonreplicating vaccines, toxins/allergens). Endogenous peptides are generated by the immunoproteasome and then are loaded into newly synthesized class I molecules in the endoplasmic reticulum following active transport from the cytosol. Proteolytic breakdown and loading of exogenous peptides into class II molecules, in contrast, occurs in acidic endosomal vacuoles. As a consequence of proteolytic processing and binding into an MHC molecule, T cells see linear peptide epitopes. In contrast, because B-cell antigen recognition requires neither proteolytic processing nor binding into an MHC molecule, B cells recognize native, three-dimensional epitopes.

In addition to the recognition of lipids and lipid-conjugates presented by CD1 molecules, there are other exceptions to the generalization that MHC molecules only present (and T cells only recognize) oligopeptides. It has been known for many years that T cells can recognize haptens, presumably covalently or noncovalently complexed with peptides residing in the antigen-binding groove. This phenomenon is familiar to physicians as contact dermatitis to nonpeptide antigens, such as urushiol (from poison ivy) and nickel ion. Additionally, a newly recognized subset of T cells designated mucosal-associated (semi-)invariant T (MAIT) cells recognize vitamin B_2 (riboflavin) and vitamin B_9 (folate) derivatives bound to MR1, a nonpolymorphic MHC class I–like molecule; these vitamin derivatives are expressed by many strains of bacteria and yeast.[30] As MAIT cells constitute ~5% of human T cells and up to 25% of CD8 cells, their binding specificity suggests a role for these cells in host defenses. Additionally, certain human γδ T cells can recognize a variety of nonpeptide phosphoantigens, such as phosphorylated nucleotides, other phosphorylated small molecules, and alkylamines. The role of

APCs and MHC-like molecules in presentation of phosphoantigens to γδ T cells remains a subject of investigation.[31]

Another exception to the generalization of T-cell recognition of oligopeptides is represented by a group of proteins termed *superantigens* (SAgs).[31] SAgs, of which the staphylococcal enterotoxin A represents a prototype, are produced by a broad spectrum of microbes, ranging from retroviruses to bacteria. They differ from conventional peptide antigens in their mode of contact both with MHC class II molecules and TCRs (Chapter 6). They do not undergo processing to oligopeptide fragments but, rather, bind as intact proteins to class II molecules and TCRs outside the antigen-binding grooves. Their interaction with TCRs is predominantly determined by variable residues of the TCR Vβ region. Because SAgs bind more or less independently of the TCRs α chain and the other variable segments of the β chain, they are capable of activating much larger numbers of T cells compared with conventional peptide antigens, and hence their name. SAgs cause a wave of T-cell activation, proliferation, and production of proinflammatory molecules that can have profound clinical consequences, leading to development of such diseases as toxic shock syndrome.[31]

LYMPHOCYTE ADHESION AND TRAFFICKING

The capacity to survey continuously the antigenic environment is an essential element of immune function. APCs and lymphocytes must be able to find antigen wherever it occurs. Surveillance is accomplished through an elaborate interdigitated circulatory system of blood and lymphatic vessels that establish connections between the solid organs of the peripheral immune system (*e.g.*, spleen, lymph nodes, and lymphoid structures in mucosal tissues) in which the interactions between immune cells predominantly occur (Chapter 2).

The trafficking and distribution of circulating cells of the immune system is largely regulated by interactions between molecules on leukocyte surfaces and ligands on vascular endothelial cells[32] (Chapter 11). Leukocyte-specific cellular adhesion molecules can be expressed constitutively or can be induced by cytokines (*e.g.*, as a consequence of an inflammatory process). Several families of molecules are involved in the regulation of lymphocyte trafficking. Particularly important are selectins and integrins, which ensure that mobile cells home to appropriate locations within lymphoid organs and other tissues. Selectins are proteins characterized by a distal carbohydrate-binding (lectin) domain. They bind to a family of mucin-like molecules, the endothelial vascular addressins. Integrins are heterodimers essential for the emigration of leukocytes from blood vessels into tissues. Members of the selectin and integrin families not only are involved in lymphocyte circulation and homing but are also important in interactions between APCs, T cells, and B cells in induction and expression of immune responses. Certain endothelial adhesion molecules, mostly members of the Ig superfamily, are similarly involved in promoting interactions between T cells and APCs, as well as in leukocyte transmigration from the vasculature. Additionally, receptors for chemokines are important determinants of lymphocyte migration, particularly in guiding tissue-selective cell trafficking.[33]

LYMPHOCYTE ACTIVATION

For both B cells and T cells, initial activation is a two-signal event[34] (Chapter 12). This generalization is particularly true for immunologically naïve cells that have not been previously exposed to antigen. The first signal is provided by antigen. Most commonly, antigens for B cells are proteins with distinct sites, termed *epitopes*, which are bound by membrane Ig. Such epitopes can be linear, defined by a contiguous amino acid sequence or (more frequently) can be conformationally defined by the three-dimensional structure of the antigen. Epitopes can also be simple chemical moieties (haptens) that have been attached, usually covalently, to amino acid side chains (Chapter 6). In addition to proteins, some B cells have receptors with specificity for carbohydrates, and less commonly lipids or nucleic acids. Antigens that stimulate B cells can be either in solution or fixed to a solid matrix (*e.g.*, a cell membrane). As previously noted, the nature of antigens that stimulate T cells is more limited. TCRs do not bind antigen in solution but are usually stimulated only by small molecules, primarily oligopeptides, which reside within the antigen-binding cleft of a self-MHC molecule.

The second signal requisite for lymphocyte activation is provided by an accessory molecule expressed on the surface of the APC (*e.g.*, B7/CD80) for stimulation of T cells or on the surface of a helper T cell (*e.g.*, CD40L/CD154) for activation of B lymphocytes. The cell surface receptors for this particular second signal on T cells is CD28 and on B cells is CD40 (Fig. 1.3). Other cell surface ligand-receptor pairs may similarly provide the second signal (Chapters 8, 12). The growth and differentiation of both T cells and B cells additionally requires stimulation with one or more cytokines, which are peptide hormones secreted in small quantities by activated leukocytes and APCs for function in the cellular microenvironment.[35] In the absence of a second signal, cells stimulated only by antigen become unresponsive to subsequent antigen stimulation (anergic)[36] (Chapter 12).

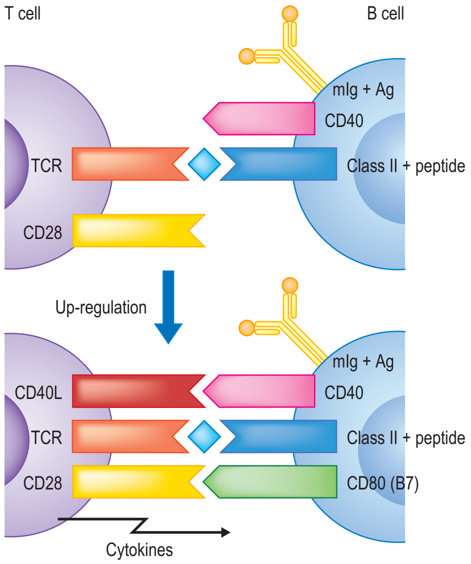


FIG 1.3 Reciprocal Activation Events Involved in Mutual Simulation of T Cells and B Cells. T cells constitutively express T-cell receptors (TCRs) and CD28. B cells constitutively express mIg and major histocompatibility complex (MHC) class II. Antigen bound to mIg is endocytosed and processed to peptide fragments that bind to MHC class II molecules for presentation to TCRs. Activation of B cells by antigen (Ag) upregulates their expression of CD80 that interacts with CD28 to activate T cells. This upregulates CD40L (CD154) on the T cell and induces cytokine synthesis. Costimulation of B cells by antigen, CD40L, and cytokines leads to Ig production.

Signal transduction through the antigen receptor is a complex process involving interactions between the specific receptor and molecules coexpressed in the cell membrane.[37] For B cells, this is a heterodimer, Igα/Igβ; and for T cells it is a macromolecular complex, CD3, usually comprising γ, δ, εε and ς chains.

Within the cell membrane, antigen receptor stimulation induces phosphorylation of Igα/Igβ or CD3 and hydrolysis of phosphatidylinositol 4,5-bisphosphate by phospholipase C, leading to generation of diacylglycerol (DAG) and inositol 1,4,5-trisphosphate (IP_3). As a consequence of signal transduction and secondarily of DAG and IP_3 generation, tyrosine and serine/threonine protein kinases are activated. In turn, these kinases catalyze phosphorylation of a number of signal transducing proteins, leading to activation of cytoplasmic transcription factors NF-AT in T cells and NF-κB in both T cells and B cells. These transcription factors then translocate to the nucleus, where they bind to 5' regulatory regions of genes that are critical to generalized lymphocyte activation[37,38] (Chapter 12).

CELL-MEDIATED IMMUNE RESPONSES

T-Cell Subsets

T lymphocytes expressing an αβ TCR can be divided into two major subpopulations based on the class of MHC molecule that their TCR recognizes and the consequent expression of one of a pair of TCR accessory molecules, CD4 or CD8 (Chapters 4, 8). Binding of CD4 to class II MHC molecules or CD8 to class I MHC molecules on APCs contributes to the overall strength of intercellular molecular interactions. The ratio of CD4 to CD8 cells in peripheral blood is usually about 2 : 1.

CD4 T Cells, Cytokines, and Chemokines

The activities of CD4 T lymphocytes, commonly referred to as *T helper (Th) cells*, are mediated predominantly through the secretion of cytokines (Chapter 9). Cytokine activity can include autostimulation (autocrine function) if the cell producing the cytokine also expresses a surface receptor for it, or stimulation of other cells in the microenvironment of the cytokine-secreting cell (paracrine function) including B cells, APCs, and other T cells. Although it is now recognized that their biological effects are broader than implied by their name, many of the principal cytokines active in the immune system are known as *interleukins* (ILs), implying that they are produced by a leukocyte to act on other leukocytes.

The specific profile of cytokines produced by CD4 T cells allows further functional subdivision (Chapters 16, 18).[35,39] CD4 T cells elaborating the "inflammatory" cytokines involved in effector functions of cell-mediated immunity, such as IL-2 and IFN-γ, are designated Th1 cells. Th2 CD4 T cells synthesize cytokines, such as IL-4 and IL-13, which control and regulate antibody responses and activate cells that are involved in host defense against parasites. Differentiation of Th1 versus Th2 subsets is a process substantially controlled by positive feedback loops, being promoted particularly by IL-12 in the case of Th1 cells and IL-4 in the case of Th2 cells. It is important to note that generalizations regarding cytokine activity are usually oversimplifications, reflecting a substantial overlap and multiplicity of functions (Chapter 9). For example, although IL-2 was initially identified as a T-cell growth factor, it also significantly affects B-cell differentiation. The prototypic inflammatory cytokine, IFN-γ, which promotes differentiation of effector function of cytotoxic lymphocytes (CTLs) and macrophages, is also involved in regulation of Ig isotype switching. And IL-4, although known primarily as a B-cell growth and differentiation factor, can also stimulate proliferation of T cells.

A distinct subset of cytokines is a large group of highly conserved cytokine-like molecules, smaller than typical cytokines (~7–12 kilodaltons [kDa]), termed *chemo*tactic cyto*kines*, or *chemokines*[33] (Chapter 10). Chemokines are classified on the basis of the number and spacing of specific cysteine residues. They regulate and coordinate trafficking and activation of leukocytes, functioning importantly in host defenses, and also broadly in a variety of nonimmunological processes, including organ development and angiogenesis. They are characterized by binding to seven-transmembrane-domain G protein–coupled receptors. Of particular interest to clinical immunologists, the chemokine receptors CCR5 and CXCR4 together with CD4 are utilized by human immunodeficiency virus (HIV) as coreceptors to gain entry into target cells.[40]

Cytokines produced by activated T cells can downregulate as well as initiate or amplify immune responses.[41] Downregulating cytokines include IL-10 (produced by both T cells and B cells) and transforming growth factor-β (TGF-β). The functions of IL-10 *in vivo* are thought to include both suppression of proinflammatory cytokine production and enhancement of IgM and IgA synthesis. TGF-β, produced by virtually all cells, expresses a broad array of biological activities, including promotion of wound healing and suppression of both humoral and cell-mediated immune responses.

In addition to their central role in initiation and regulation of immune responses, CD4 T cells are important effectors of cell-mediated immunity (Chapter 16). Through the elaboration of inflammatory cytokines, particularly IFN-γ, they are essential contributors to the generation of chronic inflammation, characterized histologically by mononuclear cell infiltrates, where their principal role is thought to be the activation of macrophages. Additionally, in some circumstances CD4 T cells can function as cytotoxic effectors, either directly as CTLs (in which case the killing is "restricted" for recognition of antigen-bound self-MHC class II) or through the elaboration of cytotoxic cytokines, such as lymphotoxin and TNF-α.

A third subset of Th cells, designated Th17, has been recognized more recently. With differentiation driven particularly by TGF-β and IL-23 and characterized by the production of the proinflammatory cytokine IL-17, Th17 cells are important in the induction and exacerbation of autoimmunity in a variety of disease models, as well as in host defenses against a broad spectrum of extracellular bacteria, fungi and other pathogens.[42] Research continues to identify additional examples of CD4 T cells, which may become recognized as distinct subsets, whose function is governed by other predominantly expressed cytokines to achieve specialized effector responses.

One final category of CD4 cells, Tregs, plays a crucial role suppressing the functions of other lymphocytes. Tregs can differentiate either in the thymus (tTregs) or in the periphery (pTregs). A third category of Tregs are induced *in vitro* (iTregs)[43] (Chapter 18). These are commonly characterized by surface expression of CD4 and CD25 and by nuclear expression of the transcription factor Foxp3. Peripheral activation of $CD25^+$ Tregs is via the TCRs; the cells are IL-2 dependent and apparently require cell-to-cell contact for suppressive function. They can suppress functions of both CD4 and CD8 T cells, as well as B cells, NK cells, and NKT cells. In contrast to activation, suppressor effects are independent of the antigen specificity of the target cells. Other Tregs are noted for production of inhibitory cytokines, including IL-10- and TGF-β–secreting Th3 cells, and IL-10–producing Tr1 cells.[44,45]

CD8 T Cells

The best understood function of CD8 T cells is that of CTL effectors.[46] CTLs are particularly important in host defenses against virus-infected cells, where they can kill target cells expressing viral peptides bound to self-MHC class I molecules (Chapter 17). This process is highly specific and requires direct apposition of CTLs and target cell membranes. Bystander cells, expressing MHC molecules that have bound peptides that the CD8 T cell does not recognize, are not affected. The killing is unidirectional; the CTL itself is not harmed, and after transmission of a "lethal hit," it can detach from one target to seek another. Killing occurs via two mechanisms: a death-receptor-induced apoptotic mechanism and a mechanism involving insertion of perforins into the target cell membrane to create a pore through which granzymes and other cytotoxic enzymes can be transferred from the CTL into the target cell. CTL activity is enhanced by IFN-γ. As CTL function is dependent on target cell surface display of MHC class I molecules, a principle mechanism of immune evasion by viruses and tumors is elaboration of factors that downregulate class I expression (Chapter 25). However, this increases susceptibility of such cells to cytolysis by NK cells that are activated to attack cells expressing low levels of class I MHC molecules.

ANTIBODY-MEDIATED IMMUNE RESPONSES

The structure of antibodies permits a virtually limitless binding specificity of its antigen-binding groove. Antigen binding can then be translated into biological effector functions based on properties of the larger nonvariable (constant) region of its heavy chains (Fc fragment) (Chapter 15). Moreover, in response to cytokines in the cellular microenvironment, through the mechanism of isotype switching an antibody-producing cell can alter the exons that are used to encode its heavy-chain constant region and thereby the biological effects of its secreted product without affecting its antigen-binding specificity. With functional heterogeneity determined by isotype, the antibody molecules provide a broad-based and efficient defense system against extracellular microbes or foreign macromolecules (*e.g.,* toxins and venoms) (Chapters 15, 27, 90).

Each antibody class contributes differently to an integrated defense system.[47] IgM is the predominant class formed on initial contact with antigen (primary immune response). As a monomeric structure comprises two light (κ or λ) and two heavy (μ) chains, it is initially expressed as a membrane bound antigen receptor on the surface of B lymphocytes. The avidity of serum IgM for antigen binding is increased by its organization into a pentamer of five of the monomeric subunits held together by a joining (J) chain. IgM is essentially confined to the intravascular compartment. As a multivalent antigen binder that can efficiently activate ("fix") complement, it is an important contributor to immune responses early after the initial encounter with antigen. The synthesis of IgM, compared with other isotypes, is much less dependent on the activity of T lymphocytes.

IgG is the most abundant immunoglobulin in serum and the principal antibody class of a secondary (anamnestic or memory) immune response. IgG molecules are heterodimeric monomers with two light (κ or λ) and two heavy (γ) chains joined by interchain disulfide bridges. Because of its abundance, its capacity to fix complement, and the expression on phagocytes of Fcγ receptors, IgG is the most important antibody for systemic secondary immune responses. IgG is the only isotype that is actively transported across the placenta. These transported maternal IgG antibodies provide the neonate with an important level of antibody protection during the early months, when its own antigen-driven antibody responses are first developing (Chapter 38).

IgA is the principal antibody in the body's secretions (Chapter 20). It is found in serum in monomeric form of two light and two heavy (α) chains or as a dimer joined by J chain. In secretions, it is usually present in dimeric form and is actively secreted across mucous membranes by attachment of a specialized secretory component (SC) that is recognized by the polyIg receptor on mucosal epithelial cells. Dimeric IgA is found in high concentration in tears, saliva, and secretions of the respiratory, gastrointestinal, and genitourinary systems; it is relatively resistant to enzymatic digestion. It is particularly abundant in colostrum, where its concentration may be >50 times that in serum, providing passive immunity to the gastrointestinal system of a nursing neonate. IgA does not fix complement by the antibody-dependent pathway and hence does not promote phagocytosis. Its role in host defenses lies in preventing a breach of the mucous membrane surface by microbes or their toxic products.

IgD and IgE are present in serum at concentrations much lower than that of IgG. The biological role of IgD remains controversial. B cells can express both membrane IgM and IgD by alternative splicing of the Ig heavy chain gene, or can secrete only IgD via an apparently atypical form of class switch recombination.[48] These mechanisms do not require T-cell help.

Although IgE is the least abundant isotype in serum, it has dramatic biological effects because it is responsible for immediate-type hypersensitivity reactions, including systemic anaphylaxis (Chapter 42). Such reactions reflect expression of high-affinity receptors for Fcε on the surfaces of mast cells and basophils. Cross-linking of IgE molecules on such cells by antigen induces their degranulation, with the immediate release of preformed potent biological mediators and *de novo* synthesis and secretion

KEY CONCEPTS

Biological Properties of Immunoglobulin (Ig) Classes

- IgM
  - Principal Ig of the primary immune response
  - Generally restricted to the vascular compartment
  - Antigen receptor (monomer) for most naïve B cells
  - Fixes complement potently
- IgG
  - Principal Ig of secondary immune responses
  - Binds to Fcγ receptors on neutrophils, monocytes/macrophages, natural killer (NK) cells
  - Four subclasses with different effector functions
  - Fixes complement (except IgG_4 subclass)
- IgA
  - Principal Ig of mucosal immunity
  - Two subclasses
- IgD
  - Antigen receptor for mature B cells
  - Typically coexpressed with membrane IgM
- IgE
  - Binds to Fcε receptors on mast cells and basophils
  - Antibody of immediate hypersensitivity
  - Important in defenses against helminths

of additional proinflammatory molecules. The protective role of IgE is in host defenses against parasitical infestation, particularly with helminths (Chapter 31).

Complement and Immune Complexes

As noted, the biological functions of IgG and IgM are importantly reflections of their capacities to activate the complement system. Through a cascade of sequential substrate–enzyme interactions, the 11 principal components of the antibody-dependent complement cascade (C1q, C1r, C1s, and C2–C9) cause many of the principal consequences of an antigen-antibody interaction (Chapter 21). These include the establishment of pores in a target cell membrane by the terminal components (C5–C9) leading to osmotic lysis; the production of factors (principally C5a) with chemotactic activity for phagocytic myeloid cells; opsonization by C3b, promoting phagocytosis; and the ability to induce degranulation of mast cells (C3a, C4a, and C5a). There are three distinct pathways to complement activation.[49] The pathway mediated by the binding of the first component (specifically C1q) to IgG or IgM has been termed the "classical" pathway (CP). The lectin pathway is similar to the CP but is activated by selected carbohydrate-binding proteins, the mannose (or mannan)-binding lectin (MBL), and ficolins, which recognize certain carbohydrate repeating structures on microorganisms. MBL and ficolins are plasma proteins that are homologous to C1q and contribute to innate immunity through their capacity to induce antibody- and C1q-independent activation of the CP. Finally, a large number of substances, including certain bacterial, fungal, and viral products, can directly activate the cascade through a distinct series of proteins also leading to activation of the central C3 component. Although bypassing C1, C4, and C2, this distinct pathway can achieve all the biological consequences of C3–C9 activation. Non–antibody-induced activation of C3 is referred to as the "alternative" pathway (AP) or "properdin" pathway. Additionally, the central components of the cascade (*e.g.,* C5a) can be directly produced by the action of serine proteases of the coagulation system.[49] Reflecting these separate pathways to activation and the fact that many types of leukocytes express receptors for activated complement components, the complement system is a major contributor to the efferent limbs of both innate and acquired immune systems.

In addition to their roles in pathogen/antigen elimination, constituents of the complement system, together with antigen–antibody (immune) complexes, act at leukocyte surfaces to regulate immune functions. For example, interaction of immune complexes via FcγR on B cells decreases their responsiveness to stimulation. In contrast, complement activation on B-cell surfaces coligates their receptors with B-cell receptors for antigen, rendering the cells more readily activated and resistant to apoptosis.

Essential for the proper function of the complement system is a series of downregulatory mechanisms that prevent unwanted activation of the system and that extinguish its activity when no longer needed. The regulatory pathways are mediated by a combination of both soluble complement-binding and digesting molecules and cell surface binding proteins.

APOPTOSIS AND IMMUNE HOMEOSTASIS

An immune response is commonly first viewed in a "positive" sense—that is, lymphocytes are activated, proliferate, differentiate, and carry out effector functions. It is equally important, however, that this positive response be tightly regulated by mechanisms that operate to turn off the response and to eliminate cells no longer required.[50,51] Under physiological circumstances, once an immune response fades, commonly as a consequence of antigen depletion, two pathways to terminal lymphocyte differentiation become available: apoptosis or differentiation into memory cells. Memory cells are, of course, a key to the effectiveness of the adaptive immune system, since a second activating encounter with antigen (*e.g.,* pathogen) is both more rapid and more productive. Isotype-switched high-affinity antibodies are rapidly produced, and/or clones of CTL effector cells proliferate. But the majority of lymphocytes in an active response are not required for maintenance of immunological memory, and the necessity for homeostasis leads to apoptosis of cells no longer required.

Apoptosis (or Regulated Cell Death; RCD) is a unique process of cellular death, widely conserved phylogenetically, and distinguished from death by necrosis by cellular shrinking, DNA fragmentation, and breakdown of cells into "apoptotic bodies" containing nuclear fragments and intact organelles that can be eliminated by phagocytosis without release into the extracellular space of the majority of intracellular, especially nuclear, components. Necrosis can be genetically determined (Regulated Necrosis; RN) or unregulated, reflecting some accidental or otherwise inevitable process (Chapter 13). Apoptosis depends on the activation of cysteinyl proteases, termed *caspases*, which cleave proteins that regulate DNA repair and the establishment/maintenance of cellular architecture. In the absence of these apoptotic mechanisms, massive proliferation of cells in lymphoid tissues results and is seen clinically as autoimmune lymphoproliferative syndrome (ALPS), which is characterized by lymphocytosis with lymphadenopathy and splenomegaly as well as autoimmunity and hypergammaglobulinemia.[52]

MECHANISMS OF IMMUNOLOGICAL DISEASES

Immunological diseases can be classified on the basis of our understanding of normal immune physiology and its perturbations in disease states (Table 1.2). One type of immunological disease results from failure or deficiency of a component of the immune system leading to failure of normal immune function (Chapters 32–40). Such disorders are usually identified by increased susceptibility to infection (Chapter 37). Failure of host defense can be congenital (*e.g.,* X-linked agammaglobulinemia; Chapter 34) or acquired (*e.g.,* acquired immunodeficiency syndrome [AIDS]; Chapter 39). It can be global (*e.g.,* severe combined immunodeficiency [SCID]; Chapter 35) or, quite specific, involving only a single component of the immune system (*e.g.,* selective IgA deficiency; Chapter 34).

A second type of immunological disease is malignant transformation of immunological cells (Chapters 77–80). Manifestations of leukocyte malignancies are protean, most commonly

TABLE 1.2 Mechanisms of Immunological Diseases

1. Functional deficiency of key immune system components
    a. Congenital
    b. Acquired
2. Malignant transformation of immune system cells
3. Immunological dysregulation
4. Autoimmunity
5. Untoward consequences of physiological immune function

reflecting the secondary consequences of solid organ or bone marrow infiltration or replacement of normal cells by tumor cells, with resulting anemia and immunological deficiency.

The remaining types of immunopathogenesis are more specific to the immune system. Dysregulation of an essentially intact immune system constitutes a third general type of immune disorder. Features of an optimal immune response include antigen recognition and elimination, with little adverse effect on the host. Both initiation and termination of the response, however, involve regulatory interactions that can go awry when the host is challenged by antigens of a particular structure or presented in a particular fashion. Diseases of immune dysregulation can result from genetic and environmental factors that act together to produce a pathological immune response, such as acute allergic diseases (Chapters 41–49). Some forms of allergic disease are thought to be a consequence of insufficient exposure to nonpathogenic microbes and other potential allergens in early childhood, resulting in an increased susceptibility to allergy, atopy, and asthma once the immune system has matured. The so-called "hygiene hypothesis" suggests that mucosal tissue-colonizing organisms play key roles in the initial establishment of immune homeostasis.[53] The importance of establishing immune homeostasis early in life is also supported by studies demonstrating reduction in the likelihood of food allergy associated with feeding of the allergenic foods to infants at high risk for allergy[54-56] (Chapter 45).

A fourth type of immunological disorder is the result of failure of a key feature of normal immune recognition, the molecular discrimination between self and nonself. Ambiguity in this discrimination can lead to autoimmune tissue damage (Chapters 50–76). Although such damage can be mediated by either antibodies or T cells, the common association of specific autoimmune diseases with inheritance of particular HLA alleles (Chapter 5) suggests that the pathogenesis of autoimmune diseases usually represents a failure of regulation of the anti-self inflammatory response by T cells. The immunologic attack on self-tissues can be general, leading to systemic autoimmunity, such as systemic lupus erythematosus; or it can be localized, as in organ-specific autoimmune diseases. In the latter instances, the immune system attacks specific types of cells and usually particular cell surface molecules. In most cases, pathology is a consequence of target tissue destruction (*e.g.*, multiple sclerosis, rheumatoid arthritis, or insulin-dependent diabetes mellitus). However, depending on the antigenic specificity of the abnormal immune response, autoimmunity can lead to receptor blockade (*e.g.*, myasthenia gravis or insulin-resistant diabetes) or hormone receptor stimulation (*e.g.*, Graves disease). It is thought by many immunologists that ambiguity in self/nonself discrimination is commonly triggered by an unresolved encounter with an infectious organism or other environmental agent that shares some structural features with self-tissue structures, although this remains a subject of controversy[57,58] (Chapter 50). Insight into mechanisms whereby specific HLA alleles predispose to development of autoimmunity and others may be protective are suggested by studies in HLA-transgenic mice, which suggest that alleles that predispose animals to a particular autoimmune disease may reflect a T cell phenotype associated with secretion of pro-inflammatory cytokines. In contrast protective alleles were associated with elaboration of trolerogenic cytokines by regulatory T cells.[59]

A fifth form of immunological disease occurs as a result of physiological, rather than pathological, immune functions. Inflammatory lesions in such diseases are the result of the normal function of the immune system. A typical example is contact dermatitis to such potent skin sensitizers as urushiol, the causative agent of poison ivy dermatitis (Chapter 44). These diseases can also have an iatrogenic etiology that can range from mild and self-limited (*e.g.*, delayed hypersensitivity skin test reactions) to life-threatening (*e.g.*, graft-versus-host disease, organ graft rejection).

HOST IMMUNE DEFENSES SUMMARIZED

The first response upon initial contact with an invading pathogen depends on components of the innate immune system (Chapter 3). This response begins with recognition of PAMPs expressed by cells of the pathogen. These include lipoproteins, lipopolysaccharide, unmethylated CpG-DNA, and bacterial flagellin, among others. PAMPs bind to PRRs on or within effector cells of the host's innate immune system, including DCs, granulocytes, and ILCs.[3] The best characterized PRRs are the TLRs, first recognized as determinants of embryonic patterning in *Drosophila* and subsequently appreciated as components of host defenses in both insects and vertebrates. TLR subfamilies can be distinguished by expression either on the cell surface or in intracellular compartments. A second major family of PRRs comprises NLRs, which detect intracellular microbial products. Binding of TLRs or NLRs by PAMP ligands triggers intracellular signaling pathways via multiple "adapters," leading to a vigorous inflammatory response.

The innate immune response also includes the capacity of NK lymphocytes to identify and destroy, by direct cytotoxic mechanisms, cells lacking surface expression of MHC class I molecules, which marks them as potentially pathogenic.[4] Additionally, an innate immune response involves elements of the humoral immune system that function independently of antibody, especially the activation of the complement cascade through the lectin pathway and the AP, with consequent opsonization of particles and microbes to promote their phagocytosis and destruction.

The nature of the adaptive immune response to any particular pathogenic agent is determined largely by the context in which the pathogen is encountered. Regardless, effectiveness depends on the four principal properties of adaptive immunity: *(i)* a virtually unlimited capacity to bind macromolecules, particularly proteins, with exquisite specificity, reflecting generation of antigen-binding receptors by genetic recombination and, in the case of B cells, somatic hypermutation; *(ii)* the capacity for self/nonself discrimination, consequences of a rigorous process involving positive and negative selection during thymocyte differentiation, as well as negative selection during B-cell differentiation; *(iii)* the property of immunological memory, reflecting antigen-driven clonal proliferation of T cells and B cells, which results in increasingly rapid and effective responses on second and subsequent encounters with a particular antigen or pathogen; and *(iv)* mechanisms for pathogen destruction, including direct cellular cytotoxicity, release of inflammatory cytokines, opsonization with antibody and complement, and neutralization in solution by antigen precipitation or conformational alteration plus phagocytosis and intracellular digestion.

Although most acquired immune responses include multiple defense mechanisms, several generalizations may be conceptually useful. T cell–mediated (and NK cell–mediated) effector functions are particularly important in defenses against pathogens encountered intracellularly or at cell surfaces, such as intracellular viruses, intracellular bacteria, and tumor cells. These responses involve the production of inflammatory cytokines by CD4 Th1 cells, as well as the direct cytolytic activity of CD8 CTLs. In contrast, host defenses to most antigens encountered primarily in the extracellular milieu are largely dependent on humoral mechanisms (antibody and complement) for antigen neutralization, precipitation, or opsonization and subsequent destruction by phagocytes. Targets of antibody-mediated immunity include extracellular virus particles, bacteria, and toxins (or other foreign proteins). It is worth reiterating, however, that induction of an effective antibody response (including isotype switching) and development of immunological memory (resulting from B-cell clonal expansion and B memory cell differentiation) require antigen activation not only of specific B cells but also CD4 T cells, particularly of the Th2 type. Additionally, antibacterial and antifungal responses that involve prominent responses by neutrophils require CD4 T cells of the Th17 type.

KEY CONCEPTS

Characteristic Infections Associated With Immune Deficiency Syndromes

Deficiencies of T Cell–Mediated Immunity

- Mucocutaneous fungal infections, especially *Candida albicans*
- Systemic (deep) fungal infections
- Systemic infection with attenuated viruses (*e.g.*, live viral vaccines)
- Infection with viruses of usually low pathogenicity (*e.g.*, cytomegalovirus)
- *Pneumocystis jiroveci* pneumonia

Antibody Deficiencies

- Infections by encapsulated bacteria (*e.g.*, *Streptococcus* spp., *Haemophilus influenza*)
- Recurrent pneumonia, bronchitis, sinusitis, otitis media
- *Giardia lamblia* enteritis

Phagocyte Deficiencies

- Infection by gram-positive bacteria (*e.g.*, staphylococci, streptococci)
- Gram-negative sepsis
- Systemic fungal infections (*e.g.*, *Candida* spp., *Aspergillus* spp.)

Adhesion Molecule Deficiencies

- Infections with pyogenic bacteria (especially staphylococci)
- Cutaneous and subcutaneous abscesses

Complement Component Deficiencies

- C3 deficiency: Infections with encapsulated bacteria
- Deficiency of terminal components: infections with gram-negative bacteria, especially *Neisseria* spp.

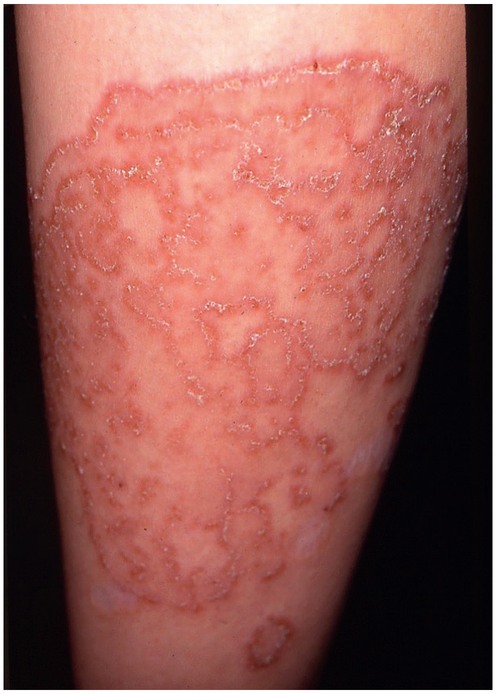


FIG 1.4 Leg of a 16-year-old patient with chronic mucocutaneous candidiasis as a consequence of a congenital T-cell deficiency associated with hypoparathyroidism and adrenal insufficiency.

Finally, clinical "experiments of nature" have proven particularly instructive in our efforts to understand the role of specific components of the immune system in overall host defenses[60] (Chapter 37). The importance of T cell–mediated immunity in host defenses to intracellular parasites, fungi (Fig. 1.4), and viruses is emphasized by the remarkable susceptibility of patients with T cell–deficiency to pathogenic organisms, such as *Pneumocystis jiroveci* and *Candida albicans*, and by the fact that utilizing attenuated live virus vaccines in such patients can lead to devastating disseminated infections. Indeed, the relationship between susceptibility to particular potential pathogens and specific immunological deficiencies is nicely illustrated by recent demonstrations that the pathogenesis of various familial forms of chronic mucocutaneous candidiasis reflects deficiency of IL-17–mediated immunity.[61]

However, patients with defects in antibody synthesis or phagocytic cell function are characteristically afflicted with recurrent infections with pyogenic bacteria, particularly gram-positive organisms. And patients with inherited defects in synthesis of terminal complement components have increased susceptibility to infection with species of *Neisseria*.

In recent years, immunology has entered the lay lexicon, largely as a result of the HIV pandemic. People throughout the world are now aware of the tragic consequences of immune deficiency. The remarkable progress in understanding this disease, however, depended substantially on earlier studies of relatively rare patients with primary immunodeficiency syndromes, more recent acceleration due to progress in genomic definition of their molecular basis. Similarly, cure of patients with primary immunodeficiencies by cellular reconstitution, particularly bone marrow or stem cell transplantation[62] (Chapter 82), presaged recent progress in correction of such diseases by gene replacement therapy[63] (Chapter 85). The "present" of clinical immunology is, indeed, bright. But its future potential to impact prevention and treatment of many challenging diseases, including cancer (Chapter 77), through specific analysis of genetic mutations and enhancement or suppression of global or antigen-specific immune responses and check-point inhibition is even more exciting to contemplate.[64] A few approaches are broadly hinted at here, and it is hoped that readers will enjoy considering such "perspectives" throughout the book and, given the nature of the immune system, it is also hoped will challenge themselves to transform a particular author's views to new and different clinical settings.

Studies in experimental animals, especially murine studies, have been critical to our understanding of molecular aspects of

ON THE HORIZON

Enhancement of Immune Responses

- Use of CRISPR/Cas9 technology in correction of gene defect in monogenic immunodeficiency diseases
- Enhancement of molecular-pathway-specific cancer immunotherapy
- Prediction of molecular targets and improvements in adjuvants for vaccines

Suppression of Immune Responses

- Further development of protocols for prevention of childhood allergic diseases by appropriate environmental exposures in infancy
- Improved allergen-specific immunotherapy of allergic diseases
- Prevention of graft-versus-host disease in allogeneic bone marrow transplantation cases
- Induction of antigen-specific immune tolerance in humans
- Pharmacological development of specific inhibitors of cytokines, chemokines, and their receptors

Immunodiagnostics and Immunopathogenesis

- Routine use of genomic analysis for development of personalized medicine applicable to immunological diseases
- Development of novel diagnostic tools based on nanotechnology arrays
- Application of gastrointestinal microbiome analysis in understanding the pathogenesis of inflammatory bowel diseases, leading to novel therapeutic options
- Understanding the role of inflammation in the pathogenesis of leading causes of morbidity and mortality, including myocardial infarction, stroke, and Alzheimer disease

immune system function and have contributed importantly to our appreciation of how aberrations of such functions are involved in the pathogenesis of disease. Insights gained from use of transgenic mice (including murine expression of human genes) and constitutive or conditional gene-knockout mice are essential to a comprehensive view of the immune system at the advancing edge of its clinical application, implying that future progress in clinical immunology will equally depend on detailed analysis in such systems. It has become apparent, however, that there are important differences in aspects of the human and rodent immune systems. Consequently, the carefully studied patient, particularly when coupled with the power of increasingly feasible genome and transcriptome sequencing, will remain the ultimate crucible for understanding human immunity and the roles of the immune system in the pathogenesis of and protection from disease.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Flajnik MF, Kasahara M. Origin and evolution of the adaptive immune system: genetic events and selective pressures. Nat Rev Genet 2010;11:47–59.
2. Buchmann K. Evolution of innate immunity: clues from invertebrates via fish to mammals. Front Immunol 2014;5:459.
3. Vivier E, Raulet DH, Moretta A, et al. Innate or adaptive immunity? The example of natural killer cells. Science 2011;331:44–9.
4. Liu J, Cao X. Cellular and molecular regulation in innate inflammatory responses. Cell Mol Immunol 2016;13:711–21.
5. Tukhvatulin AL, Dzharullaeva AS, Shcheblyakov DV, et al. Powerful complex immunoadjuvant based on synergistic effect of combined TLR4 and NOD2 activation significantly enhances magnitude of humoral and cellular adaptive immune responses. PLoS ONE 2016;11:e0155650.
6. Tiburzy B, Kulkarni U, Hauser AE, et al. Plasma cells in immunopathology: concepts and therapeutic strategies. Semin Immunopathol 2014;36:277–88.
7. Juelke K, Romagnani C. Differentiation of human innate lymphoid cells (ILCs). Curr Opin Immunol 2016;38:75–85.
8. Saunders PM, Vivian JP, O'Connor GM, et al. A bird's eye view of NK cell receptor interactions with their MHC class I ligands. Immunol Rev 2015;267:148–66.
9. Liu H, Li L, Voss C, et al. A comprehensive immunoreceptor phosphotyrosine-based signaling network revealed by reciprocal protein-peptide array screening. Mol Cell Proteomics 2015;14:1846–58.
10. O'Sullivan TE, Sun JC, Lanier LL. Natural killer cell memory. Immunity 2015;43:634–45.
11. Rezvani K, Rouce RH. The application of natural killer cell immunotherapy for the treatment of cancer. Front Immunol 2015;6:578.
12. Den Haan JM, Arens R, van Zelm MC. The activation of the adaptive immune system: cross-talk between antigen-presenting cells, T cells and B cells. Immunol Lett 2014;162:103–12.
13. Hammer GE, Ma A. Molecular control of steady-state dendritic cell maturation and immune homeostasis. Annu Rev Immunol 2013;31:743–91.
14. Rossjohn J, Pellici DG, Patel O, et al. Recognition of CD1d-restricted antigens by natural killer T cells. Nat Rev Immunol 2012;12:845–57.
15. Hirano M, Das S, Guo P, et al. The evolution of adaptive immunity in vertebrates. Adv Immunol 2011;109:125–57.
16. Xu Z, Jin B. A novel interface consisting of homologous immunoglobulin superfamily members with multiple functions. Cell Mol Immunol 2010;7:11–19.
17. Dudley DD, Chaudhuri J, Bassing CH, et al. Mechanism and control of V(D)J recombination versus class switch recombination: similarities and differences. Adv Immunol 2005;86:43–112.
18. Corcoran AE. Immunoglobulin locus silencing and allelic exclusion. Semin Immunol 2005;17:141–54.
19. Maul RW, Gearhart PJ. AID and somatic hypermutation. Adv Immunol 2010;105:159–91.
20. Mechtcheriakova D, Svoboda M, Meshcheryakova A, et al. Activation-induced cytidine deaminase (AID) linking immunity, chronic inflammation, and cancer. Cancer Immunol Immunother 2012;61:1591–8.
21. von Boehmer H. Selection of the T-cell repertoire: receptor-controlled checkpoints in T-cell development. Adv Immunol 2004;84:201–38.
22. Coutinho A, Caramalho I, Seixas E, et al. Thymic commitment of regulatory T cells is a pathway of TCR-dependent selection that isolates repertoires undergoing positive or negative selection. Curr Top Microbiol Immunol 2005;293:43–71.
23. Simons DM, Picca CC, Oh S, et al. How specificity for self-peptides shapes the development and function of regulatory T cells. J Leukoc Biol 2010;88:1099–107.
24. Hwang JK, Alt FW, Yeap LS. Related mechanisms of antibody somatic hypermutation and class switch recombination. Microbiol Spectr 2015;3:MDNA3-0037-2014.
25. Madden DR. The three-dimensional structure of peptide-MHC complexes. Annu Rev Immunol 1995;13:587–622.
26. Hansen SG, Wu HL, Burwitz BJ, et al. Broadly targeted CD8^{+} T cell responses restricted by major histocompatibility complex E. Science 2016;351:714–20.
27. Mellins ED, Stern LJ. HLA-DM and HLA-DO, key regulators of MHC-II processing and presentation. Curr Opin Immunol 2014;26:115–22.
28. Cohen NR, Garg S, Brenner MB. Antigen presentation by CD1: lipids, T cells, and NKT cells in microbial immunity. Adv Immunol 2009;102:1–94.
29. Blum JS, Wearsch PA, Cresswell P. Pathways of antigen processing. Annu Rev Immunol 2013;31:443–73.
30. Godfrey DI, Uldrich AP, McCluskey J, et al. The burgeoning family of unconventional T cells. Nat Immunol 2015;16:1114–23.
31. Spaulding AR, Salgado-Pabon W, Kohler PL, et al. Staphylococcal and streptococcal superantigen exotoxins. Clin Microbiol Rev 2013;26:422–47.
32. Smith CW. Adhesion molecules and receptors. J Allergy Clin Immunol 2008;121:S375–9.

33. Griffith JW, Sokol CL, Luster AD. Chemokines and chemokine receptors: positioning cells for host defense and immunity. Annu Rev Immunol 2014;32:659–702.
34. Esensten JH, Helou YA, Chopra G, et al. CD28 costimulation: from mechanism to therapy. Immunity 2016;44:973–88.
35. Commins SP, Borish L, Steinke JW. Immunologic messenger molecules: cytokines, interferons, and chemokines. J Allergy Clin Immunol 2010;125:S53–72.
36. Chappert P, Schwartz RH. Induction of T cell anergy: integration of environmental cues and infectious tolerance. Curr Opin Immunol 2010;22:552–9.
37. Chisolm DA, Weinmann AS. TCR-signaling events in cellular metabolism and specialization. Front Immunol 2015;6:292.
38. De Obaldia ME, Bhandoola A. Transcriptional regulation of innate and adaptive lymphocyte lineages. Annu Rev Immunol 2015;33: 607–42.
39. Zhu J, Paul WE. Peripheral CD4$^+$ T-cell differentiation regulated by networks of cytokines and transcription factors. Immunol Rev 2010;238:247–62.
40. Ajamkitsumrit B, Dampier W, Martin-Garcia J, et al. Defining differential genetic signatures in CXCR4- and the CCR5-utilizing HIV-1 co-linear sequences. PLoS ONE 2014;9:e107389.
41. Taylor A, Verhagen J, Blaser K, et al. Mechanisms of immune suppression by interleukin-10 and transforming growth factor-beta: the role of T regulatory cells. Immunology 2006;117:433–42.
42. Patel DD, Kuchroo VK. Th17 cell pathway in human immunity: lessons from genetics and therapeutic interventions. Immunity 2015;43:1040–51.
43. Abbas AK, Benoist C, Bluestone JA, et al. Regulatory T cells: recommendations to simplify the nomenclature. Nat Immunol 2013;14:307–8.
44. Campbell DJ. Control of regulatory T cell migration, function, and homeostasis. J Immunol 2015;195:2507–13.
45. Plitas G, Rudensky AY. Regulatory T cells: differentiation and function. Cancer Immunol Res 2016;4:721–5.
46. Tscharke DC, Croft NP, Doherty PC, et al. Sizing up the key determinants of the CD8$^+$ T cell response. Nat Rev Immunol 2015;15:705–16.
47. Schroeder HW Jr, Cavacini L. Structure and function of immunoglobulins. J Allergy Clin Immunol 2010;125:S41–52.
48. Chen K, Cerutti A. The function and regulation of immunoglobulin D. Curr Opin Immunol 2011;23:345–52.
49. Bajic G, Degn SE, Thiel S, Andersen GR. Complement activation, regulation, and molecular basis for complement-related diseases. EMBO J 2015;34:2735–57.
50. Zheng l, Lenardo M. Restimulation-induced cell death: new medical and research perspectives. Immunol Rev 2017;277:44–60.
51. Ruland J. Return to homeostasis: downregulation of NF-κB responses. Nat Immunol 2011;12:709–14.
52. Oliveira JB, Bleesing JJ, Dianzani U, et al. Revised diagnostic criteria and classification for the autoimmune lymphoproliferative syndrome (ALPS): report from the 2009 NIH international workshop. Blood 2010;116:e35–40.
53. Gensollen T, Iyer SS, Kasper DL, et al. How colonization by microbiota in early life shapes the immune system. Science 2016;352:539–44.
54. Du Toit G, Roberts G, Sayre PH, et al. Randomized trial of peanut consumption in infants at risk for peanut allergy. N Engl J Med 2015;372:803–13.
55. Perkin MR, Logan K, Marrs T, et al. Enquiring about tolerance (EAT) study: feasibility of an early allergenic food introduction regimen. J Allergy Clin Immunol 2016;137:1477–86.
56. Ierodiakonou D, Garcia-Larsen V, Logan A, et al. Timing of allergenic food introduction to the infant diet and risk of allergic or autoimmune disease. JAMA 2016;316:1181–92.
57. Kivity S, Agmon-Levin N, Blank M, et al. Infections and autoimmunity—friends or foes? Trends Immunol 2009;30:409–14.
58. Steed AL, Stappenbeck TS. Role of viruses and bacteria-virus interactions in autoimmunity. Curr Opin Immunol 2014;31:102–7.
59. Ooi JD, Petersen J, Tan YH, et al. Dominant protection from HLA-linked autoimmunity by antigen-specific regulatory T cells. Nature 2017;545:243–7.
60. Raje N, Dinakar C. Overview of immunodeficiency disorders. Immunol Allergy Clin North Am 2015;35:696–726.
61. Lanternier F, Cypowyj S, Picard C, et al. Primary immunodeficiencies underlying fungal infections. Curr Opin Pediatr 2013;25: 736–47.
62. Booth C, Silva J, Veys P. Stem cell transplantation for the treatment of immunodeficiency in children: current status and hopes for the future. Expert Rev Clin Immunol 2016;12:713–23.
63. Cicalese MP, Aiuti A. Clinical applications of gene therapy for primary immunodeficiencies. Hum Gene Ther 2015;26:210–19.
64. Chen DS, Mellman I. Elements of cancer immunity and the cancer-immune set point. Nature 2017;541:321–30.

MULTIPLE-CHOICE QUESTIONS

1. Which of the following statements regarding B cells is correct?
   A. B cells can function as antigen-presenting cells (APCs).
   B. Activation for antibody production requires antigen presentation by APCs.
   C. Activation requires association of B-cell receptors for antigen with CD3.
   D. Expression of major histocompatibility complex (MHC) class II molecules on B cells requires cytokines produced by APCs.

2. Natural killer (NK) cells of the innate immune system:
   A. Express clonally specific antigen receptors
   B. Exhibit allelic exclusion of antigen receptors
   C. Express receptors for immunoglobulin molecules
   D. Upregulate cytotoxic activity on virus-infected target cells that express increased levels of major histocompatibility complex (MHC) class I molecules

3. Expression of human leukocyte antigen (HLA) class I molecules:
   A. Exhibits allelic exclusion
   B. Is associated with activation-induced cytidine deaminase
   C. Produces an antigen-binding site formed by folding of a single polypeptide chain
   D. Comprises two polypeptide chains that are both encoded within the HLA gene complex

4. A 2-year-old child is referred for evaluation of massive lymphadenopathy, splenomegaly, and apparent autoimmune anemia of 6 months' duration without fever or weight loss. You suspect the child's primary problem is:
   A. Acute myelogenous leukemia
   B. Excess production of IL-2 by CD4$^+$ T helper (Th) cells
   C. Dysregulated proliferation of B cells
   D. Defective lymphocyte apoptosis

5. A 6-year-old boy is referred for evaluation of a presumptive diagnosis of a genetic deficiency of intercellular adhesion molecules. The referring pediatrician's basis for this referral was:
   A. Recurrent viral infections
   B. Unusual susceptibility to infection with gram-negative bacteria
   C. Recurrent cutaneous abscesses
   D. Presence of *Pneumocystis jiroveci* pneumonia

6. A 35-year old female presents with exophthalmos, tremor, and weight loss. You suspect that the cause of her autoimmune disease is:
   A. Destruction of target gland hormone-producing cells
   B. Destruction of target gland stromal cells
   C. Blocking of target cell hormone receptors
   D. Stimulation of target cell hormone receptors

Organization of the Immune System

Dorothy E. Lewis, Sarah E. Blutt

The human immune system consists of organs and movable cells. This design provides central locations for the initial production and differentiation of committed cells from naïve precursors, such as the fetal liver, the bone marrow, and the thymus; and more dispersed sites for the selection and further differentiation of cells into mature effector cells, such as the spleen, lymph nodes, and intestinal Peyer patches. This arrangement also allows regulation of immune responses at locations peripheral to primary lymphoid organs to provide local control of infectious processes. This chapter covers the basic features and the ontogeny of cells involved in the immune response, as well as the essential structure of lymphoid organs and sites of organized immune cells, including skin, the large intestine, and adipose tissue.

IMMUNE CELL DEVELOPMENT

Ontogeny of the Cells of the Immune System

In the first month of embryogenesis, stem cells capable of producing white blood cell progenitors are found in erythropoietic islands that are in the yolk sac.[1] The aorta–gonad–mesonephros (AGM), which is adjacent to the liver, produces progenitor cells that develop into hematopoietic stem cells (HSCs). In the sixth week of gestation, or just after the embryonic liver can be identified, progenitor stem cells in the liver begin blood cell production. By the eleventh week, the liver is the major source of hematopoiesis and remains so until the sixth month of gestation.[2]

HSC-derived progenitor cells can differentiate into granulocytes, erythrocytes, monocytes, megakaryocytes, and lymphocytes.[3] Subsequent to skeleton formation, which occurs between the second and fourth months of gestation, white blood cell development starts shifting to bone marrow. This transition is completed by 6 months' gestation. Cells that differentiate from early stem cells begin to populate the primary lymphoid organs, such as the thymus, by 7 to 8 weeks' gestation.[4] B-cell precursors initiate immunoglobulin (Ig) rearrangements by 7 to 8 weeks' gestation (Chapter 7), and T-cell precursors that have initiated T-cell receptor (TCR) rearrangement (Chapter 8) can be detected in thymus by 8 weeks' gestation. In bone marrow, B-cell progenitors congregate in areas adjacent to the endosteum and differentiate in the direction of the central sinus. HSC differentiation is a continuous process and is thus associated with many phenotypic stages. In the bone marrow of aged humans, there is evidence for a myeloid predominance, with a restricted diversity of HSCs.[5]

Stem cells with different characteristics and limited self-renewal can be induced into peripheral blood via injection of granulocyte–colony-stimulating factor (G-CSF).[4] Cord blood cells are being banked, and methods to increase levels of HSC self-renewal, including three-dimensional scaffolding, are under intensive study.[6]

Tools Essential to an Understanding of Immune Cell Biology

Understanding of the categorization and development of hematopoietic cells greatly depends on the use of monoclonal antibodies and flow cytometry to identify stage-specific leukocyte cell surface antigens. Leukocyte differentiation antigen workshops have grouped monoclonal antibodies that recognize the same single molecules on leukocytes by the cluster pattern of cells with which they are identified, hence the term *cluster of differentiation* (CD) antigen (Table 2.1, Appendix 1).

HEMATOPOIESIS AND LYMPHOPOIESIS

All mature cells of the hematopoietic and lymphoid lineages are derived from pluripotent stem cells that produce progenitors for lineage specific cells.[7] Hematopoietic progenitors mature into cells of the granulocytic, erythroid, monocytic–dendritic, and megakaryocytic lineages. Likewise, lymphoid progenitors mature into B, T, and innate lymphoid cells (ILCs), including natural killer (NK) cells (Fig. 2.1).

The site of development differs by cell type. After birth, granulocytes, monocytes, dendritic cells (DCs), erythrocytes, platelets, and B cells develop in bone marrow through the mature B-cell stage (Chapter 7) (Table 2.2). T-cell progenitors leave bone marrow and differentiate into αβ and γδ T cells in the thymus (Chapter 8). Some NK cells develop in the thymus.[7] Tissue-specific NK-cell development occurs outside the thymus, including bone marrow, lymph nodes, and the uterus.[8]

Characteristics of Hematopoietic Stem Cells

HSCs are rare in human bone marrow: 1 in 10 000 cells. They occupy distinct niches. One is close to bone and contains osteoblasts (endosteal niche), and the other is associated with the sinusoidal endothelium (vascular niche). Quiescent HSCs can be found near the arterioles in the endosteum. Actively dividing HSCs are located near the sinusoid regions close to the central veins. Different lineages of progenitors have preferred niches for development, and many HSCs are closely associated with perivascular mesenchymal stem cells.[3]

Long-term human HSCs divide once or twice per year. HSCs have characteristic flow cytometric light-scattering properties (low side scatter, medium forward scatter), no lineage-specific markers (Lin$^-$; *e.g.,* lacking CD2, CD3, CD5, CD7, CD14, CD15,

TABLE 2.1 Important Cell Surface Antigens on Hematopoietic Cells

Cell Type	Surface Antigens	Predominant Location
Hematopoietic Stem Cells		
Bone marrow hematopoietic stem cells (HSCs)	CD34$^+$, Lin$^-$, 90$^+$	Bone marrow
Peripheral blood HSCs	CD34$^+$, Lin$^-$, CD38$^+$, CD71$^+$	Blood
Myeloid Cells		
Monocytes	CD14, CD35 (CR1), CD64	Blood
Macrophages	(FcRγ1)CD68, CD13 CD64,CD35	Tissue
Langerhans cells	CD1a, CD207 (Langerin), CD35, CD64	Skin
Follicular dendritic cells	CD21, CD35 FcγRIIb	B-cell areas, lymph nodes
Interdigitating dendritic cells	CD80, CD56, Class II CD83 CD40	T-cell areas, lymph nodes
Myeloid dendritic cells	CD83, CD80, CD86, CD40 CD1a, CD11c	Mainly tissues
Plasmacytoid dendritic cells (interferon [IFN]-α producing)	CD4, CCR5, CXCR4, CD123	
Granulocytes		
Neutrophils	CD16 (FcγRIII), CD35 CD88 (C5aR)	Blood, tissues
Eosinophils	CD32 (FcγRII)	Blood, tissues
Basophils	CD23, (FcεRII), CD32	Tissues, blood
Mast cells	FcεRI	Tissues, blood
Lymphocytes		
T cells	CD7, CD3, CD4, CD8, CD28	Thymus, spleen, lymph nodes, mucosa-associated lymphoreticular tissue (MALT), blood
B cells	Surface Ig, class II, CD19, CD20, CD22, CD40	Bone marrow, spleen, lymph nodes, MALT, blood
Natural killer T cells (NKT cells)	CD16, CD56, CD94	Spleen, lymph nodes, mucosal tissues, blood
	CD3, CD56, Vα24 TCR	Blood, tissues
	CD4, CD25, Foxp3, GARP	
Regulatory T cells (Tregs)	CD4, CCR6, IL-17,	Thymus, blood, tissues
T-helper (Th)17 cells	RORγT	Intestine, blood, tissues
T Follicular helper (Tfh) cells	CD4, ICOS, PD-1, BCL-6	Germinal centers of lymph nodes

TABLE 2.2 Normal Distribution of Hematopoietic Cells in Bone Marrow

Cell Type	Approximate Proportion (%)
Stem cells	1
Megakaryocytes	1
Monocytes	2
Dendritic cells	2
Lymphocytes	15
Plasma cells	1
Myeloid precursors	4
Granulocytes	50–70
Red blood cell precursors	2
Plasma cells	1
Myeloid precursors	4
Granulocytes	50–70
Red blood cell precursors	2
Immature and mature red blood cells	10–20

or CD16), and expressing CD34, CD90, and CD49f.[6] As HSCs become active, they lose expression of CD90 and CD49f before diversion to lymphoid or myeloid precursors. Lymphoid precursors express CD10, CD45Ra, whereas myeloid precursors express CD135.

Transcription factors are unique for each population. For HSCs, these include SOX8, SOX18, and NFIB. Out of quiescence, HSCs express MYC and IKZF1. Key signaling pathways include Notch and Wnt/β-catenin.[3]

A long-lived stem cell has the capacity for self-renewal via asynchronous division.[9] HSCs circulate in peripheral blood with 10 to 100 times less frequency. Mobilization of “stem cells” to the periphery is induced by G-CSF. Of these induced cells, about 5–20% are true stem cells, although most are Lin$^-$.[9] Peripheral blood HSCs are more differentiated than bone marrow HSCs and have less self-renewal properties. Peripheral blood HSCs engraft 2 to 3 days faster than conventional bone marrow HSCs and are important in the reduction of bone marrow transplantation morbidities.

Regulation of Hematopoietic and Lymphopoietic Cell Growth and Differentiation

Regulation of stem cell differentiation occurs through interactions with a variety of microenvironmental factors. Cell surface receptors recognize either soluble ligands (*e.g.*, cytokines) released by other cells or surface ligands (*e.g.*, cell interaction molecules) expressed on adjacent cells. These receptors can facilitate differentiation. The differential expression of receptors on the stem cells allows control of proliferation and differentiation along one of the hematopoietic or lymphoid lineages.[6]

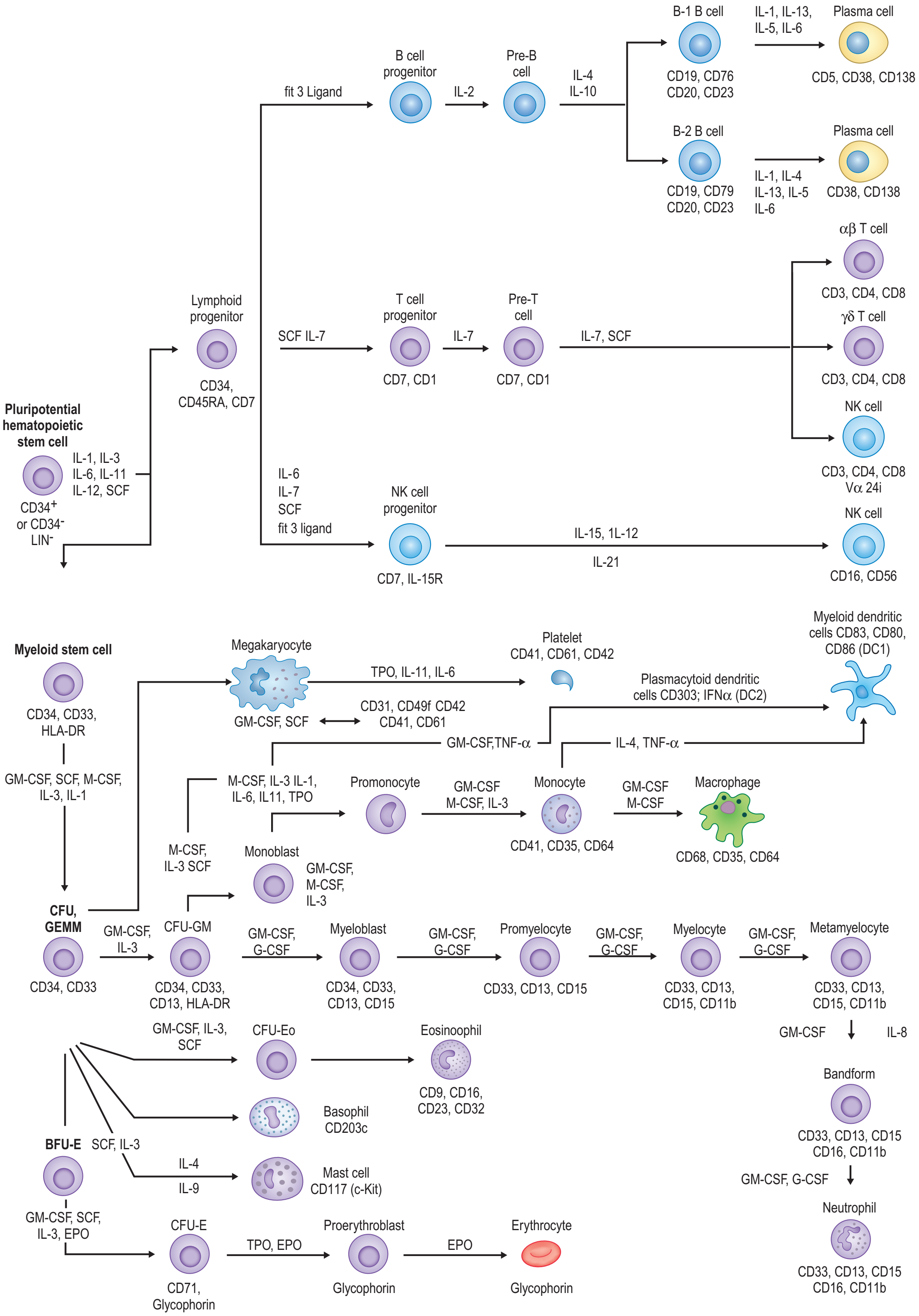

B-1 B cell
IL-1, IL-13,
IL-5, IL-6
Plasma cell
CD19, CD76
CD20, CD23
CD5, CD38, CD138
B cell
progenitor
Pre-B
cell
fit 3 Ligand
IL-2
IL-4
IL-10
B-2 B cell
Plasma cell
CD19, CD79
CD20, CD23
IL-1, IL-4
IL-13, IL-5
IL-6
CD38, CD138
αβ T cell
CD3, CD4, CD8
Lymphoid
progenitor
T cell
progenitor
Pre-T
cell
γδ T cell
SCF IL-7
IL-7
IL-7, SCF
CD34,
CD45RA, CD7
CD7, CD1
CD7, CD1
CD3, CD4, CD8
NK cell
Pluripotential
hematopoietic
stem cell
IL-1, IL-3
IL-6, IL-11
IL-12, SCF
CD3, CD4, CD8
Vα 24i
CD34+
or CD34-
LIN-
IL-6
IL-7
SCF
fit 3 ligand
NK cell
progenitor
NK cell
IL-15, 1L-12
IL-21
CD7, IL-15R
CD16, CD56
Myeloid dendritic
cells CD83, CD80,
CD86 (DC1)
Myeloid stem cell
Megakaryocyte
Platelet
CD41, CD61, CD42
TPO, IL-11, IL-6
Plasmacytoid dendritic
cells CD303; IFNα (DC2)
CD34, CD33,
HLA-DR
GM-CSF, SCF
CD31, CD49f CD42
CD41, CD61
GM-CSF,TNF-α
IL-4, TNF-α
GM-CSF, SCF, M-CSF,
IL-3, IL-1
M-CSF, IL-3 IL-1,
IL-6, IL11, TPO
Promonocyte
GM-CSF
M-CSF, IL-3
Monocyte
GM-CSF
M-CSF
Macrophage
CD41, CD35, CD64
CD68, CD35, CD64
M-CSF,
IL-3 SCF
Monoblast
GM-CSF,
M-CSF,
IL-3
CFU,
GEMM
GM-CSF,
IL-3
CFU-GM
GM-CSF,
G-CSF
Myeloblast
GM-CSF,
G-CSF
Promyelocyte
GM-CSF,
G-CSF
Myelocyte
GM-CSF,
G-CSF
Metamyelocyte
CD34, CD33
CD34, CD33,
CD13, HLA-DR
CD34, CD33,
CD13, CD15
CD33, CD13, CD15
CD33, CD13,
CD15, CD11b
CD33, CD13,
CD15, CD11b
GM-CSF, IL-3,
SCF
CFU-Eo
Eosinoophil
GM-CSF
IL-8
Bandform
CD9, CD16,
CD23, CD32
Basophil
CD203c
CD33, CD13, CD15
CD16, CD11b
BFU-E
SCF, IL-3
IL-4
IL-9
Mast cell
CD117 (c-Kit)
GM-CSF, G-CSF
Neutrophil
GM-CSF, SCF,
IL-3, EPO
CFU-E
TPO, EPO
Proerythroblast
EPO
Erythrocyte
CD71,
Glycophorin
Glycophorin
Glycophorin
CD33, CD13, CD15
CD16, CD11b

FIG 2.1 The differentiation of hematopoietic cells.

Cytokines (Chapter 9) affect the growth and maintenance of pluripotent stem cells, as well as the development and differentiation of specific cell lineages. The effect of the cytokine can depend on whether the cell has previously been or is concurrently being stimulated by other cytokines. The stage of differentiation, as well as the presence or absence of the cytokine's receptor on the cell surface, also affects the cellular response.

Stromal cells located within bone marrow and the thymus regulate hematopoietic and lymphoid cell growth and differentiation by releasing cytokines, such as interleukin 4 (IL-4), -6, -7, and -11; leukemia inhibitory factor (LIF); granulocyte macrophage–colony-stimulating factor (GM-CSF); G-CSF; and stem cell factor (SCF).[3] Stromal cells also participate in cell–cell interactions with progenitors that express fibroblast growth factor 1 (FGF-1) and FGF-2, which support HSC expansion. Stromal cells form the intercellular matrix (*e.g.,* fibronectin and collagen), which binds to the selectin and integrin receptors present on hematopoietic and lymphoid progenitors.[10]

Cytokines That Affect the Growth and Maintenance of Pluripotent and Multipotent Stem Cells

Pluripotent stem cells can reconstitute cells of the hematopoietic and lymphoid lineages. Maintenance of pluripotent capacity is mediated through factors that keep HSCs quiescent, including c-kit, N-cadherin, osteopontin, transforming growth factor-β (TGF-β), and Wnt. Factors that have a negative effect on quiescence include Hedgehog and the notch ligands Delta and Jagged.[3] Because the stem cell pool is depleted as the progeny differentiate, a low level of proliferation of stem cells is required to avoid exhaustion. The entry of stem cells into the cell cycle and the subsequent proliferation, as well as commitment to particular lineages, are controlled by cytokines and transcription factors. Data suggest that flt-3 ligand, c-kit ligand, and megakaryocyte growth and development factor all promote long-term stem cell expansion. The combination of c-kit ligand, IL-3, and IL-6 causes more rapid expansion but does not allow long-term extension of precursor cells.[11]

Several cytokines, alone or in combination, promote stem cell growth (Table 2.3).[4] Cytokine combinations are more effective at inducing stem cell growth compared with individual cytokines. IL-1 promotes stem cell growth by inducing bone marrow stromal cells to release additional cytokines and by synergistically stimulating these cells in the presence of other cytokines. IL-3 promotes the growth of hematopoietic progenitors. The effect is significantly enhanced by IL-6, IL-11, G-CSF, and SCF. Stromal cell–derived IL-11 enhances IL-3–induced colony formation in 5-fluorouracil–resistant stem cells. Cytokines secreted by stromal cells (*e.g.,* IL-6, G-CSF, and SCF) exert their effects by shortening the G0 period in stem cells. IL-3 acts on cells after they have left G0. IL-12 is unable to support the growth of primitive hematopoietic stem cells, either by itself or in conjunction with IL-11 or SCF. However, it acts in synergy with IL-3 and IL-11, or IL-3 and SCF, to enhance stem cell survival and growth.

In some situations, a cytokine can enhance the growth of hematopoietic and lymphoid cells, whereas in others the same cytokine can inhibit cell growth or enhance differentiation. LIF

TABLE 2.3 Cytokines Important for Hematopoietic Cell Growth and Differentiation

Cytokines	Stem Cells	Thymocytes	B Cells	Natural Killer (NK) Cells
Interleukin (IL)-1	Acts on stromal cells	Differentiation		
IL-2		Pleomorphic	Proliferation	Proliferation
IL-3	Proliferation			
IL-4		Pleomorphic	Promotes (low) Prevents (high)	Inhibits IL-2
IL-5			Proliferation/differentiation	
IL-6	Shortens G0	Enhances stimulation	Maintains potential	Enhances IL-2
IL-7		Survival/proliferation	Proliferation of pro- and pre-B cells	Activation
IL-10			Survival	
IL-11 Oncostatin M	Shortens G0		Maintains potential	
IL-12	Survival			Activation proliferation
IL-13			Activation/division of mature B cells	
IL-15			Proliferation	Development/survival
IL- 21			Proliferation	Expansion
Stem cell factor (SCF)/c-kit	Survival	Atrophy	Maintains potential	Expansion
G-CS	Shortens G0		Maintains potential	
FLt3 ligand	Growth factor		Increases proliferation	Expansion
Stromal cell–derived factor (SDF)1-α		Proliferation/regeneration	Chemoattractant	
Leukemia inhibitory factor (LIF)	Proliferation	Atrophy		
Thrombopoietin	Expansion/regulates self-renewal			
Tumor necrosis factor (TNF)-α	Proliferation: inhibits granulocytes			
Transforming growth factor (TGF)-β	Inhibits growth enhanced granulocytes			
Macrophage inflammatory protein (MIP)-1α	Inhibits			
Nerve growth factor (NGF)			Proliferation/differentiation	Expansion

can enhance the growth and development of bone marrow progenitor cells along multiple lineages in media containing IL-3, IL-6, and GM-CSF. However, in the absence of other cytokines or factors in serum, LIF has little effect on the growth and development of CD34$^+$ progenitors. TGF-β and IL-4 are potent inhibitors of hematopoietic progenitor cell growth; yet they enhance granulocyte development. Tumor necrosis factor-α (TNF-α) inhibits the development of granulocytes, but it can potentiate IL-3 effects on hematopoietic progenitor cell proliferation.

Other cytokines have effects on the proliferation and differentiation of multipotent progenitors of hematopoietic and lymphoid cells. GM-CSF and IL-3 promote development of granulocytes, macrophages, DCs, and erythrocytes. IL-6 participates in the development of neutrophils, macrophages, platelets, T cells, and B cells. Thrombopoietin signaling promotes stem cell self-renewal to increase transplantation success.[7]

Cytokines That Inhibit Hematopoietic Stem Cell Growth

Cytokines produced by mature cells can downregulate hematopoietic stem cell growth. Macrophage inflammatory protein-1α (MIP-1α) is an inhibitor of hematopoietic progenitor cell proliferation. Other factors regulate stem cell growth through a variety of mechanisms, including the promotion of terminal differentiation (*e.g.*, interferon-γ [IFN-γ] and TGF-β) or through the induction of apoptosis (*e.g.*, TNF-α). When pathologic conditions exist, these cytokines can have adverse effects on hematopoietic and lymphoid cell development.

Cytokines Affecting Development and Differentiation of Specific Cell Lineages

Differentiation begins with the commitment of pluripotent stem cells to a specific lineage. Cytokines can have lineage-specific effects that act specifically at late stages of differentiation. Erythropoietin regulates the later stages of erythrocyte differentiation, whereas G-CSF induces granulocyte differentiation and macrophage colony-stimulating factor (M-CSF) promotes macrophage maturation.[12] Cytokines that play an important role in the growth and development of specific cell lineages are described below under each cell type.

KEY CONCEPTS

Cells of the Immune System

- Pluripotent stem cells in bone marrow give rise to all lineages of the immune system, platelets, and red blood cells.
- Development and regulation of cells of the immune system is associated with programmed appearance of specific cell surface molecules called "cluster of differentiation" (CD) markers and with responsiveness to selective cytokines.
- Mature cells of the immune system include antigen-presenting cells (APCs); phagocytic cells, including neutrophils, eosinophils, and basophils; and lymphocytes, including T cells, B cells, and natural killer (NK) cells, as well as other innate lymphoid cells.
- APCs include monocytes, macrophages, dendritic cells (DCs), B cells, endothelial cells, epithelial cells, and adipocytes. They can direct the differentiation and function of both innate and acquired immune cells.
- Polymorphonuclear (PMN) granulocyte cells are important in the early response to stress, tissue damage, and pathogens. They include neutrophils, eosinophils, and basophils.
- Lymphocyte lineages have discrete subpopulations with specialized functions. These include CD4 and CD8 T cells, B-1 and conventional B-2 B cells, and NK and other innate lymphoid cells. CD4 T-helper (Th) subsets include regulatory T cells (Tregs), Th17, and Tfh cells.

Mature Cells of the Immune System

The mature cells of the immune system primarily arise from progenitor cells in bone marrow. They include both nonspecific and antigen-specific effector cells.

Antigen-Presenting Cells

The central player in both nonspecific and antigen-specific lines of defense is the antigen-presenting cell (Chapter 6). In addition to their nonspecific effector functions, these cells are crucial for the development of specific immune responses. With maturation, these cells enter the blood (Table 2.4) and circulate into the tissues and organs.

Antigen-presenting cells (APCs) are found in the solid lymphoid organs and skin (Chapter 19) at a frequency that varies from 0.1–1%. Specialized APCs in B-cell areas of lymph nodes and spleen are termed *follicular dendritic cells* (FDCs). They trap antigen–antibody complexes important in the generation and maintenance of memory B cells. FDCs do not express major histocompatibility complex (MHC) class II molecules as do other APCs. Instead, they have receptors for immunoglobulin G (IgG) (FcγRI [CD64]) and complement component C3b (CR1 [CD35]), respectively.

Monocytes–Macrophages

Monocyte–macrophage lineage cells exist in blood (~10% of leukocytes) primarily as monocytes, which are large 10- to 18-µm cells with peanut-shaped, pale purple nuclei as determined by Wright staining (see Table 2.4). The cytoplasm, which is 30–40% of the cell, is light blue and has azurophilic granules that resemble ground glass with intracytoplasmic lysosomes. The cells express MHC class II, CD14 (the receptor for lipopolysaccharide), and distinct Fc receptors (FcRs) for Ig. The latter include FcγRI (or CD64), which has a high affinity for IgG, and FcγRII (or CD32), which is of medium affinity and binds to aggregated IgG. FcγRIII (or CD16) has low affinity for IgG and is associated with

TABLE 2.4 Normal Distribution of White Blood Cells in the Peripheral Blood of Adults and Children

	APPROXIMATE PERCENTAGE		RANGE OF ABSOLUTE COUNTS (NO./µL)	
Cell Type	**Adults**	**Children (0–2 yr)**	**Adults**	**Children (0–2 yr)**
Monocytes	4–13		400–1000	ND
Dendritic cells	0.5–1	ND[a]	30–170	ND
Granulocytes	35–73		2500–7500	1000–8500
Lymphocytes	15–52	34–75	1450–3600	3400–9000
As % of Lymphocytes				
T cells	75–85	53–84	900–2500	2500–6200
CD4 cells	27–53	32–64	550–1500	1300–4300
CD8 cells	13–23	12–30	300–1000	500–2000
B cells	5–15	06–41	100–600	300–3000
Natural killer (NK) cells	5–15	03–18	200–700	170–1100

[a] Not determined.

Child data adapted from Shearer W, Rosenblatt H, Gelman R, et al. Lymphocyte subsets in healthy children from birth through 18 years of age: the pediatric AIDS Clinical Trials Group P1009 study. J Allergy Clin Immunol 2003;12:973–80.

antibody-dependent cellular cytotoxicity (ADCC). It is expressed on macrophages, but not on blood monocytes. Monocytes and macrophages also express CD89, the Fc receptor for IgA.[12]

Macrophages are more differentiated monocytes that reside in various tissues, including the lungs, liver, and brain.[13] Cells of the monocyte–macrophage lineage adhere strongly to glass or plastic surfaces. This process activates them, and this can confound functional studies when they are isolated in this way. Many cells of this lineage phagocytose organisms or tumor cells *in vitro*. Cell surface receptors, including CD14, Fcγ receptors, and CR1 (CD35), are important in opsonization and phagocytosis. This lineage expresses MHC class II, and some express the low-affinity receptor for IgE (CD23). Other cell surface molecules include myeloid antigens CD13 (aminopeptidase N) and CD15 (Gal (1–4) or [Fuc (1–3)] GlcNAc) and the adhesion molecules CD68 and CD29 or CD49d (VLA-4). Classic blood monocytes in humans (85%) express high levels of CD14, but no CD16. Nonclassic monocytes express less CD14, but more CD16. This later subset produces more IL-12, TNF-α, and IL-1β. These cells have receptors for various cytokines (*e.g.*, IL-4 and IFN-γ). Activated macrophages are a major source of cytokines, including IFN, IL-1, and TNF, as well as complement proteins and prostaglandins.

Macrophages, along with dentritic cells (DCs), are much more plastic in differentiation and function than previously realized. They can be alternatively activated and thereby become suppressive, developing antiinflammatory properties relevant in immune responses to cancer as well in maintaining adipose integrity. Alternative activation is induced by the T-helper cell 2 (Th2) (Chapter 16) cytokines IL-4 and IL-13.[14]

Monocytes and macrophages arise from colony-forming unit–granulocyte monocyte (CFU-GM) progenitors that differentiate into monoblasts, promonocytes, and then monocytes.[12] Mature monocytes leave bone marrow and circulate in the bloodstream until they enter tissues, where they develop into tissue macrophages (alveolar macrophages, Kupffer cells, intestinal gut macrophages, and microglial cells). Tissue macrophages appear to originate from fetal macrophages and seed tissues early in fetal development, where they are maintained by longevity and slow self-renewal.[13]

Several cytokines participate in the development of monocytes and granulocytes. SCF, IL-3, IL-6, IL-11, and GM-CSF promote development of myeloid lineage cells from CD34$^+$ stem cells, especially at early stages. M-CSF acts at later stages of development and is lineage specific, inducing macrophage maturation.[12]

Dendritic Cells

DCs express high levels of MHC class II molecules and are potent inducers of primary T-cell responses. Except for bone marrow, they are found in virtually all primary and secondary lymphoid tissues and in skin, mucosae, and blood. DCs are abundant in the thymus medulla for selection of thymocytes.

DCs are derived from CD34$^+$ MHC class II-negative precursors present in bone marrow, which also give rise to macrophages and granulocytes. GM-CSF and TNF-α are involved in DC development.[15] DCs residing in peripheral sites, such as skin, the intestinal lamina propria, lungs, the genitourinary tract, and so on, are typically immature. These cells are more phagocytic with less MHC class I, MHC class II, and costimulatory molecules. These immature DCs take up antigens in tissues for subsequent presentation to T cells, and as they migrate, they mature into efficient APCs.

The predominant APCs of the skin are the Langerhans cells[16] found in the epidermis and characterized by rocket-shaped granules called *Birbeck granules.* Immature tissue DCs in peripheral tissues engulf and process antigen and home to T-cell areas in the draining lymph nodes or spleen.[17] Mature DCs can directly present processed antigens to resting T cells to induce proliferation and differentiation, and this is a key functional difference between mature DCs and macrophages. The effector cells produced after this presentation then home to the site of the antigenic assault.

TNF-α maintains viability of Langerhans cells in skin and stimulates their migration. In Peyer patches (Chapter 20), immature DCs occur in the dome region underneath follicle-associated epithelium (FAE), where they actively endocytose antigens taken up by M cells. Mature interdigitating DCs are found in T-cell regions, where they can induce Th2 immune responses (Chapter 16).

Three types of DC are prominent—two types of “conventional” dendritic cells (cDCs) and plasmacytoid dendritic cells (pDCs). cDC1s derive from bone marrow are found in lymphoid tissues and express CD1a and CD11c. About 50% peripheral blood DCs are cDC1. cDC2s express CD141, are similar to murine CD8 α/α DCs in function, and are rare in peripheral blood but are common in lymph nodes. pDCs are high producers of IFN-α. They express CD123 and low levels of CD11c, along with BDCA 2 and 4. DCs can be derived from either myeloid or lymphoid lineages. DCs are largely influenced by stimulation with Toll-like ligands found on a variety of stimuli, which then direct the differentiation and function of innate and acquired immune cells.

Polymorphonuclear Granulocytes

Polymorphonuclear (PMN) granulocytes arise and mature in bone marrow. After release from bone marrow, their life span varies from a few to 5 to 6 days and is regulated by environmental conditions. They constitute 65–75% of the white blood cells in peripheral blood, are 10–20 μm in diameter, and have a multilobed pyknotic nucleus characteristic of cells undergoing apoptosis (see Table 2.4).[18] PMNs use diapedesis to gain access to tissues from blood.

Granulocytes are early responders to stress, tissue damage, or pathogen invasion. Because of their function in phagocytosis and killing, they possess granules with unique staining characteristics that are used to categorize the cells as neutrophils (Chapter 22), basophils (Chapter 23), or eosinophils (Chapter 24).

Neutrophils

Most circulating granulocytes are neutrophils (90%). Their granules are azurophilic and contain acid hydrolase, myeloperoxidase, and lysozymes. These granules fuse with ingested organisms to form phagolysosomes, which kill the invading organism. In some cases, there is extracellular release of granules after activation via the Fc receptors. Neutrophils express CD13, CD15, CD16 (FcγRIII), and CD89 (FcαR). In response to bacterial infection, the number of circulating granulocytes typically increases. This often includes the release of immature granulocytes, called *band* or *stab cells,* from bone marrow. In a mild infection, both the number and the function of neutrophils are increased. This is associated with a delay in apoptosis. With more severe infection, function may be impaired due to the immaturity of cells.

A newly described function of some neutrophils is to release neutrophil extracellular traps, which can capture microbes extracellularly and then use autophagy to digest them intracellularly.[19]

Neutrophils mature from CFU-GM progenitor cells and differentiate within a 10- to 14-day period. These progenitors give rise to myeloblasts, promyelocytes, myelocytes, and finally mature neutrophils. SCF, IL-3, IL-6, IL-11, and GM-CSF promote the growth and development of neutrophil precursors. Other cytokines are important for differentiation of CFU-GM progenitors into mature neutrophils.[20] G-CSF induces maturation of neutrophil precursors into mature neutrophils. IL-4 enhances neutrophil differentiation induced by G-CSF, while inhibiting the development of macrophages induced by IL-3 and M-CSF.

Eosinophils

Eosinophils typically comprise 2–5% of the white cells in blood. They exhibit a unique form of diurnal variation. Peak production occurs at night, perhaps because glucocorticoid levels are lower. Eosinophils are capable of phagocytosis followed by killing, although this is not their main function. The granules in eosinophils are much larger than in neutrophils and are actually membrane-bound organelles. The crystalloid core of the granules contains a large amount of the major basic protein (MBP), which can neutralize heparin and is toxic. During degranulation, the granules fuse to the plasma membrane, and their contents are released into the extracellular space. Organisms that are too large to be phagocytosed, such as parasites, can be exposed to cell toxins by this mechanism. The MBP can damage schistosomes *in vivo*, although damage is minimized because the MBP is confined to a small extracellular space. Eosinophils also release products that counteract the effects of mast cell mediators. Whether eosinophils are absolutely required for helminth control is controversial.

Eosinophils mature from a progenitor (CFU-Eo) into an eosinophilic myeloblast, then an eosinophilic promyelocyte, a myelocyte, and finally a mature eosinophil.[21] Three cytokines are important in the development of eosinophils: GM-CSF, IL-3, and IL-5. GM-CSF and IL-3 promote eosinophil growth and differentiation. SCF also has an effect on eosinophil function. Eotaxin (CCL11) promotes eosinophilia. IL-5 has more lineage-specific effects on eosinophil differentiation. Although it also affects some subsets of T and B cells, it is important for eosinophil survival and maturation. Eosinophils are involved in the pathophysiology of asthma, with contribution to airways dysfunction and tissue remodeling, and IL-5 is being targeted to correct eosinophilia.

Basophils and Mast Cells

Basophils represent less than 1% of the cells in the peripheral circulation. They are characterized by large, deep-purple granules. Mast cells are found in proximity to blood vessels and are much larger than peripheral blood basophils. Their granules are less abundant, and the nucleus is more prominent. There are two different types of mast cells—designated mucosal and connective tissue—depending on their location. Mucosal mast cells require T cells for their proliferation, whereas connective tissue mast cells do not. Both types have granules that contain effector molecules. After degranulation, which is effected by cross-linkage of cell surface IgE bound to cells via the high-affinity receptor for IgE, basophils and mast cells release heparin, histamine, and other effector substances to mediate an immediate allergic attack (Chapters 23 and 42).

Basophils and mast cells share a number of phenotypic and functional features that suggest derivation from a common precursor. Both basophils and mast cells contain basophilic-staining cytoplasmic granules; both express the high-affinity IgE receptor (FcεRI); and both release a number of similar chemical mediators that participate in immune and inflammatory responses, particularly anaphylaxis. They both have been implicated in allergic inflammation and in fibrosis. However, basophils and mast cells also have some distinct morphological and functional characteristics that suggest that they may be distinct lineages of cells, rather than cells at different stages within the same lineage. In human, transcription factor analysis places basophils closer to eosinophils than mast cells.[22]

Basophils mature from a progenitor (CFU-BM) into basophilic myeloblasts and then into basophilic promyelocytes, myelocytes, and finally mature basophils. Less is known about the stages of mast cell development, although they are probably derived from the same CFU-BM progenitor as basophils.

In humans, SCF induces the most consistent effects on the growth and differentiation of both basophils and mast cells. Both IL-3 and SCF are important for intestinal mast cell differentiation. Il-6 can also increase mast cell numbers. This probably explains why T cells are needed for their development.[28] Both IL-4 and IL-9 stimulate mast cell development in mice. However, in humans, only IL-9 acts in synergy with SCF to enhance mast cell growth. Additional cytokines that affect basophil growth include nerve growth factor and GM-CSF or TGF-β, and IL-5 for basophil differentiation.

Platelets and Erythrocytes

Hematopoietic stem cells give rise to platelets and erythrocytes. Platelets are necessary for blood clot formation and mediate a number of immune functions. Mature red blood cells are necessary for oxygen delivery to tissues.[23] Platelets derive from CFU-GEMM progenitors, which differentiate into burst-forming units for megakaryocytes (BFU-MEG). BFU-MEG then differentiate into CFU-MEG, promegakaryoblasts, megakaryoblasts, megakaryocytes, and finally platelets.[24] Several cytokines, particularly thrombospondin, IL-1, IL-3, GM-CSF, IL-6, IL-11, and LIF, affect the growth and differentiation of platelets.

Erythrocytes also derive from CFU-GEMM progenitors, but their progenitors are burst-forming units for erythrocytes (BFU-E), which, in turn, differentiate into CFU-E, pronormoblasts, basophilic normoblasts, polychromatophilic normoblasts, orthochromic normoblasts, reticulocytes, and finally erythrocytes. Again, several cytokines, notably GM-CSF, SCF, IL-9, thrombospondin, and erythropoietin, regulate erythrocyte development.

Lymphocytes

Lymphocytes, the central cell type of the specific immune system, represent about 25% of the white blood cells in blood (see Table 2.4). Small lymphocytes range from 7–10 μm in diameter and contain a nucleus that stains dark purple with Wright staining, and a small cytoplasm. Large granular lymphocytes range from 10–12 μm in diameter and contain more cytoplasm and scattered granules. The three types of lymphocytes that circulate in the peripheral blood—T cells, B cells, and ILCs, including NK cells—constitute approximately 80%, 10%, and 10% of the total blood lymphocyte population, respectively (Chapters 7, 8, and 17). In the thymus, most of the lymphocytes (90%) are T cells; however, in the spleen and lymph nodes, only about 30–40%

are T cells. The preponderant lymphocytes in these locations are B cells (60–70%).[25]

T Lymphocytes

T lymphocytes arise from lymphocyte progenitors in bone marrow committed to the T-cell lineage before moving to the thymus. In the early stages of embryogenesis, T-cell precursors migrate to the thymus in waves.[26] Associated with this migration is the developing ability of thymic education elements, epithelial cells, and DCs to select appropriate T cells.[27]

In the thymus, T cells rearrange their specific antigen receptors (TCRs) (Chapter 4) and then express CD3 along with the TCRs on their surface (Chapter 8). Resting T cells in blood typically range from 7–10 μm in diameter and are agranular, except for the presence of a structure termed *Gall body,* which is not found in B cells (see Table 2.4). The Gall body is a cluster of primary lysosomes associated with a lipid droplet. A minority of T cells in blood (about 20%) are of the large granular type, that is, they are 10–12 μm in diameter and contain primarily lysosomes that are dispersed in the cytoplasm. Golgi apparati also are found.

The preponderant form of the TCR, found on about 95% of circulating T cells, expresses TCRαβ.[28] Some $CD3^+$ cells do not express either CD4 or CD8 (double-negative, or DN) and express an alternative TCRγδ. Further differentiation in the thymus occurs from $CD3^+$ cells that express both CD4 and CD8 (double-positive, or DP) to cells expressing either CD4 or CD8 but not both (Chapter 8). These mature cells then circulate in peripheral blood at a ratio of about 2 : 1 (CD4:CD8) and populate lymph nodes, the spleen, and other secondary lymphoid tissues.

T-cell progenitors, which are $CD7^+$, arise in bone marrow from a multipotential lymphoid stem cell. After migration to the thymus, the $CD7^+$ progenitors give rise to a population of $CD34^+$, $CD3^-$, $CD4^-$, and $CD8^-$ T-cell precursors, which undergo further differentiation into mature T cells. Cytokines produced by thymic epithelial cells (*e.g.,* IL-1 and soluble CD23) promote differentiation into $CD2^+$, $CD3^+$ thymocytes (see Table 2.3). IL-7 induces the proliferation of $CD3^+$ DN ($CD4^-$ $CD8^-$) thymocytes, even in the absence of comitogenic stimulation. IL-7 is an absolute requirement for human T-cell development.[29]

IL-2 and IL-4 demonstrate complex effects on thymocyte development. Both can promote development of prothymocytes, as well antagonizing their development. IL-6 acts as a costimulator of IL-1– or IL-2–induced proliferation of DN thymocytes and can stimulate the proliferation of mature, cortisone-resistant thymocytes alone.

Subpopulations of T Cells

T cells can be divided into subsets based on surface expression of CD4 and CD8 as well as by function in immune response. CD4 and CD8 T cells were originally characterized by expression of the respective antigen and association with functional ability. For example, human T cells expressing CD4 provide help for antibody synthesis, whereas cells expressing CD8 develop into cytotoxic T cells. The distinction is better described on the basis of which antigen-presenting molecule is used for TCR interaction. Thus CD4 T cells recognize antigen in the context of MHC class II molecules, and CD8 T cells recognize antigen presented by class I molecules (Chapters 5 and 6). Memory T cells are divided on the basis of expression of CD45R0, CCR7, CD28, and CD95, which categorize the functions of cells as stem cell memory, central memory, transitional memory, effector memory, and terminal effector cells. Early memory T cells have high potential for lymph node homing and proliferation, but later stage cells home to the periphery, are effector cells, and do not proliferate.[30]

Th cells mature in response to foreign antigens. Their function is dependent on the production of cytokine patterns, which characterize them as Th type 1 (Th1), Th2, or Th17.[31] The precursor Th cell first differentiates into a Th0 cell, which produces IFN-γ and IL-4. The cytokine environment subsequently determines whether Th1 or Th2 cells predominate. Th1 cells produce primarily IFN-γ, IL-2, and TNF-α and are important in cell-mediated immunity to intracellular pathogens, such as the tubercle bacillus. Th1 cells primarily use T-bet transcription factor. Th2 cells produce predominantly IL-4, -5, -6, -10, and -13, as well as IL-2; they predominate in immediate or allergic type 1 hypersensitivity and primarily use GATA-3 transcription factor. Other populations of CD4 T cells can develop and rely on IL-23 or IL-12 action upon the cells. If T cells are exposed to IFN-γ, they upregulate both IL-12R and IL-23R, which then produce either conventional Th1 cells or another subset, Th17, which produces IL-17 and is important for controlling immune cell activation in the gastrointestinal (GI) tract. Overactive function of this subset has been associated with autoimmunity. The Th17 population preferentially uses RORγt transcription factor. T follicular helper cells are those classically determined to help B cell responses in germinal centers. They are $CD4^+$, $ICOS^+$, and $PD\text{-}1^+$; and they express transcription factor BCL-6. It is likely that there are other epigenetically altered T cells that allow diversity of function during an immune response.

A minor subpopulation (<5%) of $CD3^+$ cells in peripheral blood express γδ TCR molecules. Most of these cells do not express CD4 or CD8. However, some intraepithelial lymphocytes (IELs) that express γδ TCR also express CD8 αα homodimers in place of conventional CD8 αβ heterodimers. These cells, which are thymus independent, are involved in the initial response to bacterial antigens presented in mucosal epithelium.

Another minor subpopulation of T cells, natural killer T cells (NKT cells), can be $CD4^+$ or $CD8^+$ and express a single Vα chain, Vα24, which recognizes glycolipids in the context of CD1a, rather than a classic MHC molecule.[31,32] NKT cells express MIP-1 α and β and have a Th1 bias, but lack IL-10 production. The final subset comprises regulatory T cells (Tregs), which occur naturally and can be induced *in vitro*. They are $CD4^+$ and express high levels of CD25 and the transcription factor FOXP3 and perhaps GARP.[33] These cells are important in regulatory immune responses. Tregs are reduced in autoimmunity and in adipose tissue during obesity and increased in cancer.

B Cells and Plasma Cells

B cells represent 5–10% of lymphocytes in blood (see Table 2.4). They are typically 7–10 μm in diameter and lack Gall bodies and granules. The cytoplasm is characterized by scattered ribosomes and isolated rough endoplasmic reticulum (RER). Unless the cells are activated, the Golgi apparatus is not prominent. B cells express cell membrane immunoglobulin (mIg), the majority expressing both IgM and IgD.[34] A small minority of B cells express either surface IgG or IgA. Cell surface molecules found on B cells (Chapter 7) include CD19, CD20, CD23, CD40, CD72, CD79a and b, MHC class II, FcγRII (CD32), and complement receptors C3b (CR1a; CD35) and C3d (CR2a; CD21). B-cell mIg is a part of a B-cell receptor complex that consists of CD19, CD21, and CD81 (Chapter 4). B-cell proliferation and differentiation processes take place in the germinal centers of the lymph nodes.

Upon activation and cross-linking of surface Ig by specific antigen, B cells undergo proliferation and differentiation to produce plasma cells, which lose expression of mIg and MHC class II molecules. Plasma cells (10–15 μm) are not normally found in blood. They display an eccentric nucleus and a basophilic cytoplasm with a well-developed Golgi apparatus and parallel arrays of expanded Ig-containing RER. Plasma cells are nondividing, specialized cells terminally differentiated from B cells and whose primary function is to secrete Ig.

In vitro studies of cytokines involved in the development of early B-cell progenitors show that combinations of SCF (but not IL-3) with IL-6, IL-11, or G-CSF can maintain B-lymphoid potential.[34] Stromal cell–dependent differentiation of fetal pro-B cells occurs in conjunction with FLK-2/FLT-3 ligand and IL-7 and on several transcription factors, including PU.1, IKAROS, E2A, EBF, PAX5, and IRF8. Unlike in mice, IL-7 is not an absolute requirement for B-cell development in humans.[35]

IL-4 has a variety of important effects on B-cell growth and differentiation. Low doses of IL-4 induce pre-B cells to differentiate into B cells expressing surface membrane IgM, whereas higher doses of IL-4 inhibit differentiation of B cells. In mature B cells, IL-4 increases expression of MHC class II, CD23, and CD40; promotes activation and progression to the G1 stage of the cell cycle; enhances proliferation after stimulation through the Ig receptor; and induces immunoglobulin class switch in human to IgG4 and IgE (IgG1 and IgE in mouse). IL-13, which is closely related to IL-4, has many similar effects on B cells.

IL-2, -5, -6, and -11 and nerve growth factor act on mature B cells to either enhance their proliferation or promote their differentiation into immunoglobulin-secreting cells. IL-10 enhances the viability of B cells *in vitro*, increases MHC class II expression, and augments the proliferation and differentiation of B cells after stimulation through the Ig receptor or CD40. TGF-β1 is a major switch factor for IgA. This cytokine induces human B cells triggered by mitogen to switch to both IgA1 and IgA2. Stromal cell-derived factor 1 (SDF-1) attracts early-stage B-cell precursors and is a likely mechanism whereby B cells form islands in bone marrow.

There are at least two major populations of B cells: B-1 cells, which are found in the follicular mantle and peritoneal cavity, and conventional B-2 cells, which are found in lymphoid follicles. The B-1 lineage predominates early in gestation and produces natural antibodies of the IgM isotype.[36] There is good evidence for local expression of IgA plasma cell precursors in the ileum important for bacterial containment.[37]

Innate Lymphoid Cells

Natural Killer Cells

A new nomenclature has arisen to classify lymphocytes that have cytolytic or noncytolytic functions typical of T cells, but do not express a T cell receptor. The first described were cytolytic NK cells, which comprise about 10–15% of circulating lymphocytes (see Table 2.4). These cells are usually larger than typical lymphocytes (10–12 μm), but have less nuclear material and more cytoplasm. They possess electron-dense, peroxidase-negative granules and a developed Golgi apparatus. Functional NK cells can be found in the fetal liver as early as 6 weeks' gestation. These fetal NK cells express cytoplasmic CD3 proteins but no TCR rearrangements. Evidence suggests that an Fcγ receptor–positive cell that does not express lineage-specific markers (Lin$^-$) exists in the fetal mouse thymus, where it normally gives rise to T cells. However, if removed from the thymus, the cells develop into CD3$^-$ NK cells. Such CD3$^-$ cells with variable CD16 expression exist in the human thymus and can be induced to proliferate, express NK-associated antigens, and exhibit NK cell function *in vitro*. These cells also express substantial levels of CD3δ and CD3ε in the cytoplasm.[38]

Mature NK cells in blood do not express conventional antigen receptors, such as TCR or Ig, and the genes for these receptors remain unrearranged. Some express FcγRIII (CD16) and others express CD56, an adhesion molecule. More than 90% of these cells are CD11b$^+$ but CD27$^-$. In tissues, subsets of human NK cells express variable levels of both CD11b and CD27, and this defines their function (tolerant, cytotoxic, or regulatory). NK cells, like T cells, also express the CD2 molecule. NK cells express the β chain of the IL-2 receptor, CD122, which allows resting NK cells to respond directly to IL-2.

The function of some NK cells is to provide nonspecific cytotoxic activity toward virally infected cells and tumor cells (Chapter 17). When provided with an antibody, NK cells can kill specifically. This death delivery mechanism, known as ADCC, occurs via binding of the antibody to the Fcγ receptor CD16. After activation, NK cells produce cytokines, such as IFN-γ, and this affects proliferation and differentiation of other cell types, especially DCs. Some of the recognition molecules on human NK cells activate, some inhibit, and some act as receptors for MHC class I molecules.

NK cells express a number of membrane antigens in common with T cells, and they share functional properties with some T-cell subsets, suggesting a common origin. NK cells are found in fetuses before the development of T cells or the thymus, and they develop normally in nude, athymic mice. NK cells probably develop extrathymically, and data suggest that they can develop from stem cells in lymph nodes. NK cells arise from "triple-negative" (CD3$^-$CD4$^-$CD8$^-$) precursors that are CD56$^+$, but CD34$^-$ and CD5$^-$. T cells, in contrast, develop from triple-negative precursors that are CD34$^+$CD5$^+$CD56$^+$. It is likely that T cells and NK cells arise from a common triple-negative precursor with the phenotype CD7$^+$CD34$^+$CD5$^+$CD56$^+$.

The α chain of the IL-2 receptor, CD25, is the prime determinant of T cell versus NK lineage specificity. Once CD25 is upregulated, the cell is destined to become a T cell. IL-15 and IL-7 play major roles in the early development of NK cells. FLT ligand and c-KIT also facilitate NK cell expansion. In mature NK cells, IL-2 induces proliferation and activation. This probably occurs via the IL-2 receptor β chain (CD122), as NK cells do not express CD25. IL-2 also induces the growth of NK cells from precursors in bone marrow cultures. Both IL-7 and IL-12 activate NK cells. Although IL-4 inhibits the effects of IL-2 or IL-7 on NK cells, it acts synergistically with IL-12 to induce proliferation of CD56$^+$ cells. IL-6, despite having no effect by itself, enhances NK cell activity in thymocytes cultured with IL-2. Finally, IL-15 is involved in signaling NK cells for survival.[39] Subsets of human NK cells develop on the basis of responsiveness to TGF-β and IL-10 (tolerant); IL-12 and IGF-1 (cytotoxic); and TGF-β, IL-7, and IL-15.[40]

Noncytotoxic Innate Immune Cells

These cells, which are similar in function to T-helper subsets, are divided into three main groups, ILC1, ILC2, and ILC3, defined by the cytokines they produce.[41] ILC1 cells are noncytotoxic Lin$^-$ cells that produce INF-γ and TNF. ILC2 cells produce Th2 cytokines, such as IL-4, -5, -9, and -13; and some produce amphiregulin. ILC3 cells produce IL-17A, -17F, and -22; GM-CSF;

and TNFα and are the most heterogeneous. They also express CCR6 and CD117 and can be divided on the basis of expression of the NCR Nkp44. The role of these cells in normal host function and responses to chronic inflammatory stimuli and cancer is under intensive study.

KEY CONCEPTS

Tissues of the Immune System

- Stem cells proliferate and mature into effector cells in the primary lymphoid organs, which include bone marrow and the thymus.
- Mature immune cells reside in secondary lymphoid organs, where additional maturation occurs and immune responses are generated.
- The spleen and lymph nodes comprise the systemic immune system, which functions to protect the body from antigens in the lymphatic drainage and the bloodstream.
- The mucosal immune system (respiratory, gastrointestinal, and genital) and the skin and adipose tissue have distinguishing features that differentiate the immune system at these sites from those of the systemic immune system.
    - Mucosal sites include the mucosa-associated lymphoreticular tissue (MALT).
    - Commensal organisms at mucosal surfaces are an important component of the immune response at these sites.

MAJOR LYMPHOID ORGANS

The primary lymphoid organs are sites where lymphocytes differentiate from stem cells and then proliferate and mature into effector cells. In humans, from birth to old age, these functions are carried out only in bone marrow and the thymus.

Bone Marrow

Bone marrow provides the environment necessary for the development of most of the white blood cells of the body (Fig. 2.2). At birth, most bone cavities are filled with actively dividing blood-forming elements known as "red" marrow. By 3 to 4 years, however, the tibia and femur become filled with fat cells, limiting their role in hematopoietic development. The ribs, sternum, iliac crest, and vertebrae remain 30–50% cellular and produce hematopoietic cells throughout a person's life.[1] Main components of bone marrow include blood vessels, cells, and extracellular matrix. The production of cells from HSCs occurs in areas separated by vascular sinuses. The walls of the surrounding sinus contain a layer of endothelial cells with endocytic and adhesive properties. These specialized endothelial cells of the sinuses appear to produce type IV collagen and laminin for structural support via CXCL-12 (SDF-1) interactions. These cells also elaborate CSFs and IL-6. The outer wall of the sinus is irregularly covered with reticular cells that branch into areas where cells develop and provide anchors by producing reticular fibers. Megakaryocytes lie against this wall, touching the endothelial cells.

A functional unit of marrow, called a *spheroid,* contains adipocytes, stromal cell types, and macrophages. These reticular cell networks compartmentalize the developing progenitor cells into separate microenvironments called *hematons.* Osteoblasts and osteoclasts regulate the production of progenitor cell expansion.[3]

The distribution of stem and progenitor cells across the radial axis of bone suggests that HSCs are next to the bone surface, whereas the more mature progenitor cells are nearer to the central venous sinus to facilitate release of mature cells. The production of new progenitor cells from stem cells occurs as a result of interactions between stem cells and stromal cells. Given the right stimuli, most of the progeny proliferate and differentiate further, which may result in migration from the bone marrow. In migrating, the cells become detached from stromal elements and progress toward the central sinus.

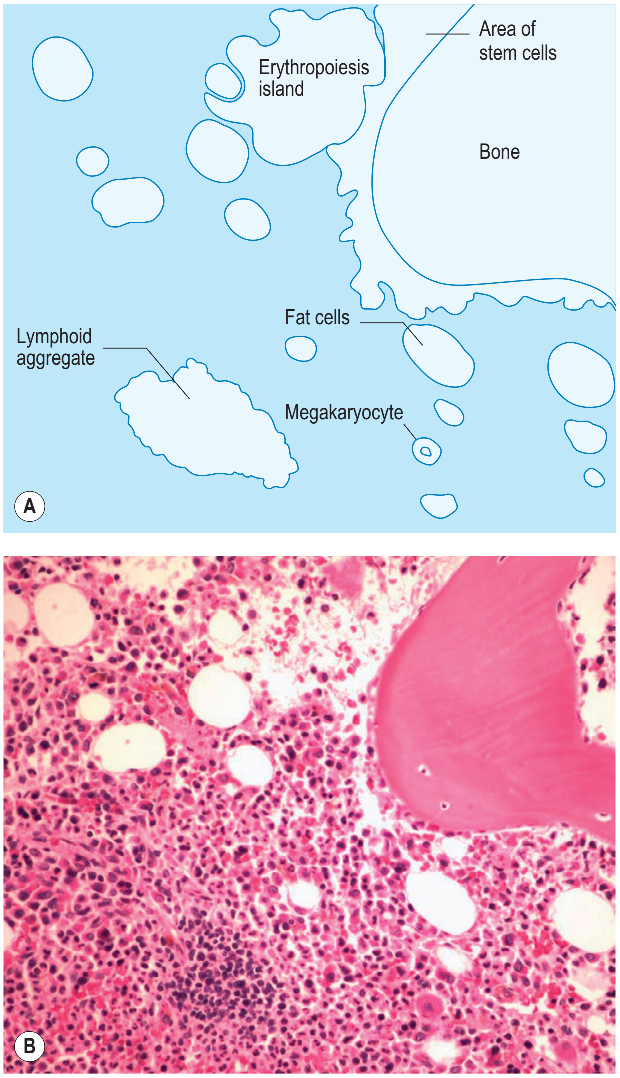


FIG 2.2 Structure of bone marrow, showing islands of erythropoiesis, granulopoiesis, and scattered lymphocytes.

Control of hematopoiesis is regulated by both positive and negative cytokines, and by upregulation and downregulation of various adhesion molecules in committed progenitor cells. The molecules involved include the fibronectin receptor, glycoproteins IIb and IIIa, ICAM-1 (CD54), LFA-1 (CD11, CD18), LFA-3 (CD58), CD2, and CD44. Adhesion molecules on stromal cell surfaces include fibronectin, laminin, ICAM-1 (CD54), types I, III and IV collagen, and N-CAM. The most clearly established role for adhesion molecules involves fibronectin, which allows erythroid precursors to bind to stromal cells and thus facilitates progression from erythroblast to reticulocyte. Molecular signals important for the HSC niche include N-cadherin, which regulates osteoblastic interactions with HSCs; WNT/B catenin signals, which are important for the self-renewal of HSCs; VEGF, which is important for coupling osteoblasts with vascular endothelial cells; and PDE2, which is an inflammatory mediator that can increase HSC numbers.

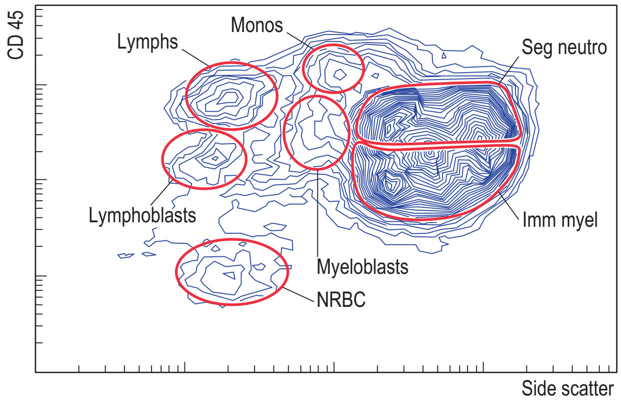


FIG 2.3 Flow cytogram of Normal Human Bone Marrow Based on CD45 Expression and Side Scatter. Most of the major hematopoietic populations can be delineated. In this example 1.5% are red blood cell precursors (NRBCs), 1.5% are lymphoblasts, 3% are mature lymphocytes (Lymphs), 3% are monocytes (Monos), 4% are myeloblasts, 45% are segmented neutrophils (Seg neutro), and 42% are immature myeloid cells (Imm myel).

Accessory cell populations in bone marrow regulate many aspects of hematopoiesis. The upregulation of growth of the earliest progenitor cells is mediated by cytokines. For example, macrophages produce IL-1, which then induces stromal cells to express growth factors, such as GM-CSF, IL-6, and IL-11. Downregulation can occur at any stage. For example, T cells regulate hematopoiesis by producing factors that act on the early erythroid progenitor cells, BFU-E. Later progenitors, CFU-E, are fully differentiated by erythropoietin. Activated T cells, however, produce factors that suppress BFU-E and CFU-E *in vitro*.

The cells in bone marrow were originally characterized by morphology. The predominant types are those of the myeloid lineage, which account for about 50–70% of the cells. Red blood cell precursors represent 15–40% of the total cells. Other lineages exist in lower proportions (<5%). With the advent of cell surface antigen markers and flow cytometry, a more precise delineation could be made (Fig. 2.3). Thus of the mature leukocytes in bone marrow, approximately 70% are $CD3^+$, $CD14^+$, $CD20^+$, or $CD11b^+$. Both memory T and B cells return to bone marrow after generation. These are designated as Lin^+. Of the Lin^- cells, about 6% are $CD33^+$ and primarily of myeloid lineage. A Lin^-CD71^+ population comprises about 18% of the total and is mostly of red blood cell lineage.

Thymus

The thymus is located in the mediastinum and below the sternum. This bilobed organ develops from the third and fourth pharyngeal pouches and is endodermal in origin. It is organized into a loose lobular structure, with areas in each lobe consisting of a cortex of rapidly dividing cells and a medulla that contains fewer, but more mature, T cells (Figs. 2.4 and 2.5). This arrangement has long suggested a scenario for differentiation, where cells progress from the cortex to the medulla. Nonlymphocyte cells play very important site-specific roles in the development of T cells. Epithelial cells are scattered throughout the thymus. Depending on their location, they are known as *nurse cells, cortical epithelial cells,* or *medullary epithelial cells.* Macrophage-type cells and interdigitating cells that are derived from bone marrow are located at the junction between the cortex and the medulla and are involved in T-cell selection.

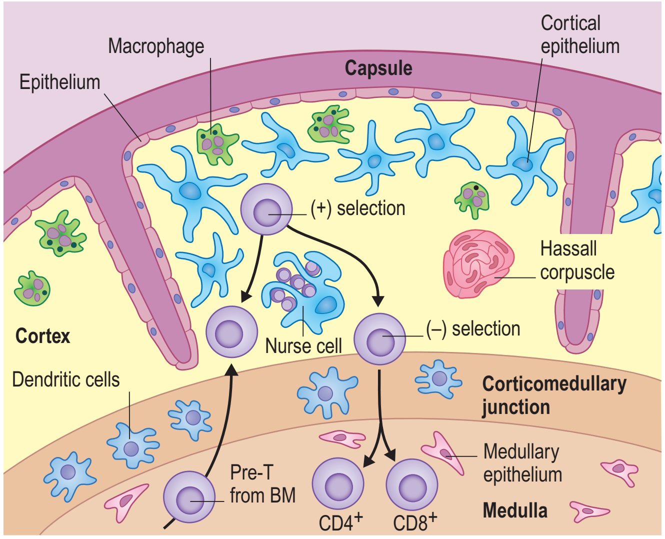


FIG 2.4 Structure of the Thymus, Showing pre–T Cells Entering from Bone Marrow (BM). Positive selection occurs on thymic epithelial cells; negative selection probably occurs during interactions with cortico-medullary dendritic cells. This may explain why single-positive CD4 or CD8 cells are found primarily in the medulla. Nurse cells appear to remove negatively selected cells. Hassall corpuscles are specialized cells producing thymic growth factors.

Enlarged, activated T-cell precursors from bone marrow begin by colonizing the subcapsular region of each lobe. These are actively proliferating and can self-renew. Selection begins when their progeny encounter MHC class II–bearing cortical epithelial cells. A further education process probably occurs by interaction with macrophage-like cells found at the cortico-medullary junction and in the medulla.

Thymus nurse cells, found in the cortex, were originally thought to contribute to the thymic education of T cells. Because large numbers of thymic cells (50–200) can be found inside each nurse cell, it was believed that these structures provided an environment where selection and expansion could occur. There is now evidence that secondary rearrangement of TCRα can occur in these structures.[42]

A structure known as *Hassall corpuscle,* which consists of concentric whorls of epithelial cells, is found in the medulla; but its function is unclear. Hassall medullary epithelial cells contain secretory granules, and this network of cells may be active in the production of thymic hormones, especially thymic stromal lymphopoietin, which has a role in production of DCs that select Tregs in the thymus.[43] In the fetus, these bundles of cells are widely scattered but become larger as the thymus matures. The center cells eventually become keratinized and die.

Thymic differentiation (Chapter 8) involves rearrangement of functional TCR (Chapter 4), surface expression of CD3, and both positive and negative selection (Chapter 8), which allow only a small percentage of T cells to survive. Pre–T cells in the thymus express CD2, CD5, and CD7, as well as activation antigens, such as CD38 and the transferrin receptor (CD71). Pre–T cells express intracytoplasmic CD3 and exhibit rearrangements in the TCRβ chain. Successful rearrangement of TCRα allows the cell to progress to the next stage of development, with functional

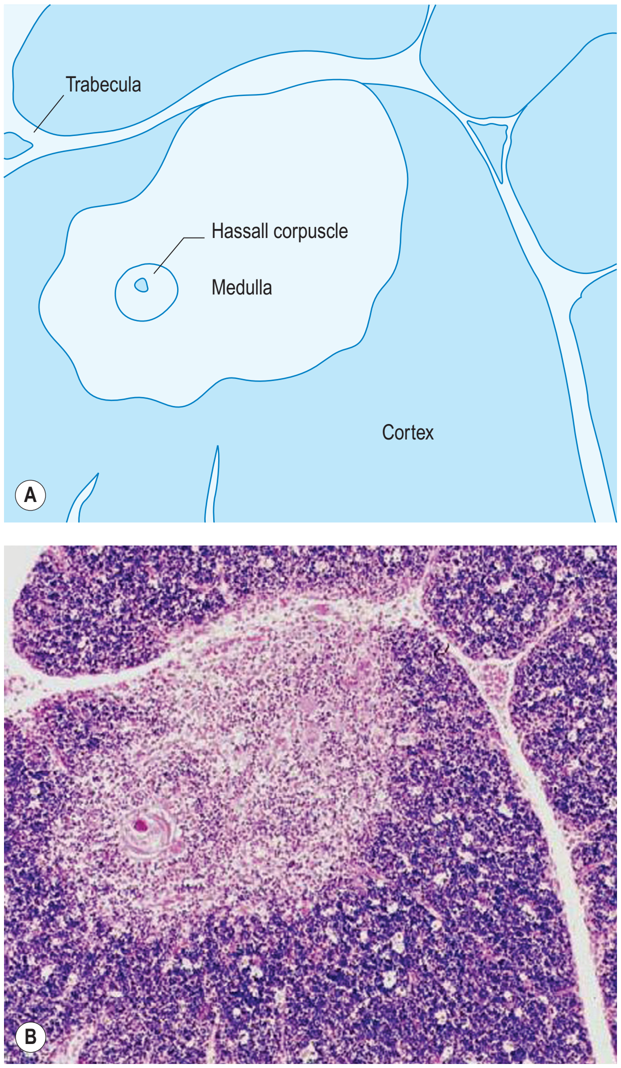


FIG 2.5 Human Thymus Showing Cortex and Medullary Areas. Cortical thymocytes are stained with an anti-CD1 antibody. Most medullary T cells do not express CD1.

TCR and CD3 on the cell surface.[27] Most of the cells in the thymus (85%) express both CD4 and CD8 on their surface and are termed DP. They also express CD1 and CD69, an activation marker. CD69 is expressed until the cell reaches the single-positive stage, where it expresses either CD4 or CD8, but not both.

T cells are CD45RO$^+$ at the DP stage into the single-positive stage. Before leaving the thymus, CD45RO is downregulated, and CD45RA appears. The most mature thymus cells lose CD1 expression and either CD4 or CD8 expression. Most of these mature cells are also negative for activation molecules (CD38 and CD71). However, they acquire an adhesion molecule called *CD44,* which is necessary for homing. Upon completion of this thymus selection and education process, mature CD4 or CD8 T cells leave the thymus and enter the peripheral circulation via the postcapillary venules at the cortico-medullary junction.

After birth and during childhood, the thymus continues to grow and select T cells. This process is probably necessary to develop a fully normal repertoire. Before puberty, however, the thymus begins to involute. The rapidly dividing cortex is the first to atrophy, leaving medullary areas intact. The sensitivity of cortical thymocytes to hormone-induced death probably accounts for the involution, although human thymocytes are less sensitive to glucocorticosteroids compared with murine thymocytes. However, an increase in steroids reduces immature thymocyte numbers and enhances thymus involution. Recent evidence suggests that active TCR rearrangements, and hence T-cell development, continue in the adult thymus, albeit at a lower level than during childhood. There is an age-associated decline in new T-cell production, such that by age 75 years, the ability to make new T cells in humans is severely reduced.

Development of Hematopoietic and Lymphoid Cells

Although most of the key steps during the growth and development of hematopoietic and lymphoid cells occur in bone marrow and the thymus, additional maturation steps occur after the cells leave those tissues. For example, monocytes and DC precursors migrate from blood vessels into tissues, where they mature into macrophages and DCs, respectively. There is recent evidence for a tissue-associated macrophage that is fetal in origin. Mast cells and eosinophils also undergo further differentiation in resident tissues. After leaving bone marrow and the thymus, B and T cells undergo further maturation and memory cell development in secondary lymphoid organs. There is strong evidence that some T cells, particularly the γδ T cells residing in mucosal epithelium, do not develop in the thymus.

SECONDARY LYMPHOID ORGANS

Secondary lymphoid organs are sites where mature lymphocytes reside and where immune responses are generated. Secondary lymphoid organs belong to either the systemic immune system or the mucosal immune system. The systemic immune system, which includes the spleen and lymph nodes, functions to protect the body from antigens in the lymphatic drainage and circulating in the bloodstream. The mucosal immune system responds to antigens that enter through mucosal epithelium and plays an important role in the inductive phase of the immune response. Unique features differentiate the mucosal immune system from the systemic immune system (Chapter 20). These include efferent, but not afferent, lymphatics, a specialized FAE involved in antigen sampling at the mucosal surface (Fig. 2.6), specialized DCs that rapidly process and present antigens to initiate antigen-specific immune responses, unique distribution and subsets, and an environment that promotes class switching to IgA.

Systemic Immune System

Spleen

The human spleen is surrounded by a capsule of fibrous tissue with many trabeculae traversing from the capsule into the tissue of the spleen. These trabeculae branch and anastomose, forming a complex framework of lobules. Splenic blood vessels enter and exit through the hilum of the spleen and branch into smaller vessels within the trabeculae. Splenic tissue is supported by a fine network of reticular cells and fibers, called the *reticulum,* which connects and supports the trabeculae, blood vessels, and the capsule. The lobules of the spleen can be functionally divided into two compartments, the red pulp and the white pulp. The largest compartment is the red pulp, which contains numerous venous sinuses situated between arteries and veins. Blood is filtered through these sinuses, which contain many macrophages that phagocytose senescent red and white blood cells, bacteria, and other particulate material. Other leukocytes, including

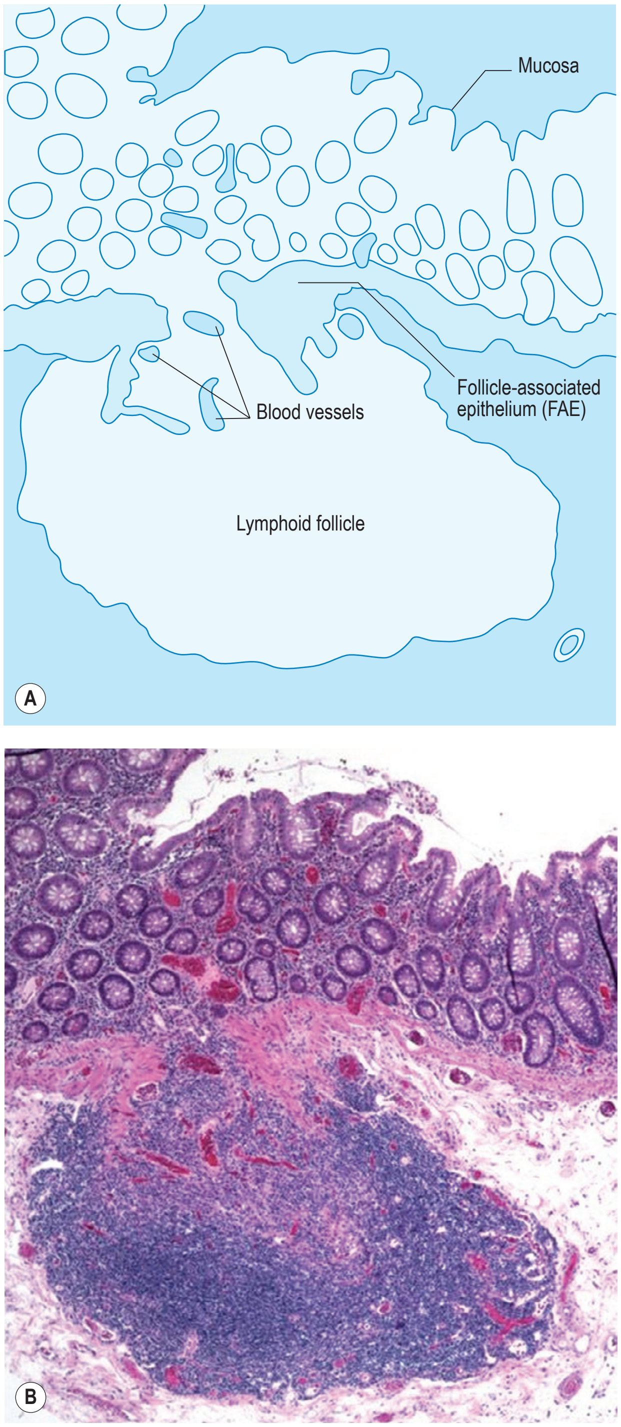

Mucosa
Follicle-associated epithelium (FAE)
Blood vessels
Lymphoid follicle
A
B

FIG 2.6 Lymphoid Follicles in the Human Large Intestine. FAE, follicle-associated epithelium.

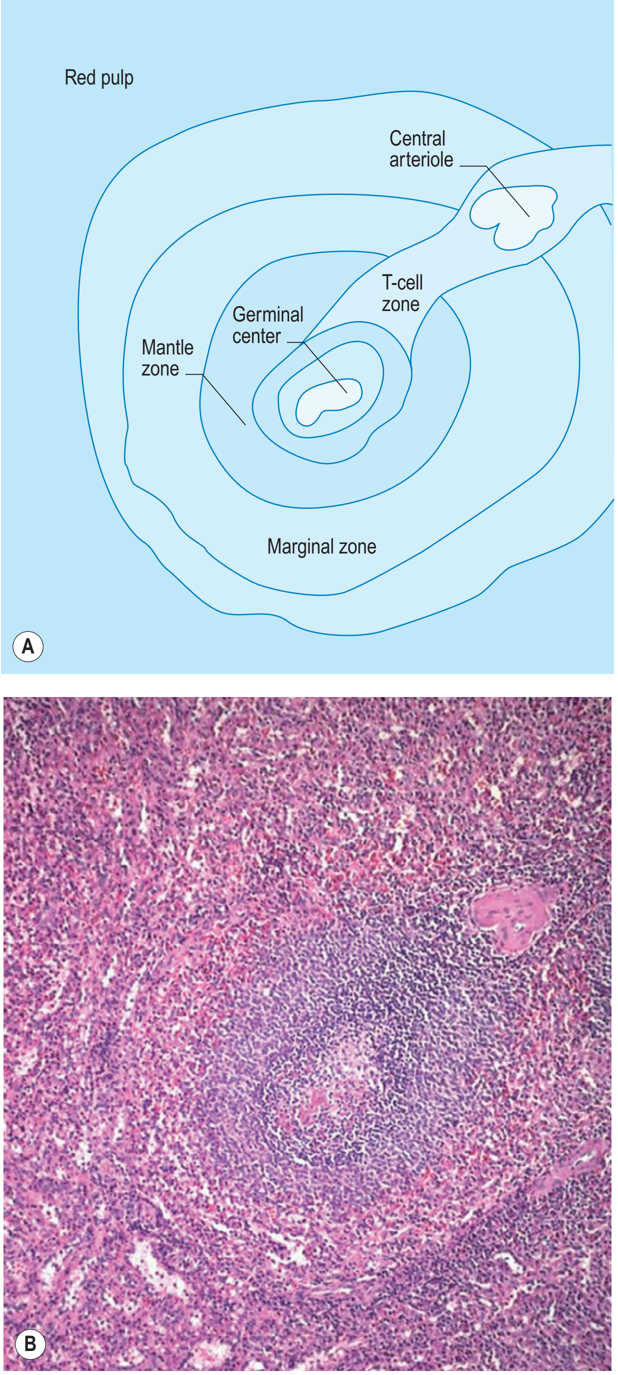

Red pulp
Central arteriole
T-cell zone
Germinal center
Mantle zone
Marginal zone
A
B

FIG 2.7 Human spleen showing a periarteriolar lymphoid sheath and germinal center.

neutrophils, eosinophils, and lymphocytes, particularly plasma cells, are found in the red pulp.[44]

The white pulp consists of lymphoid tissue surrounding central arterioles, which are branches of trabecular arteries. The human spleen is structurally different from rodent spleens because there is no central organization of follicles and the central artery.[44] Rather, a T cell–predominant area is found immediately surrounding a central arteriole, the so-called *periarteriolar lymphoid sheath*, which contains both CD4 and CD8 T cells. It is punctuated at intervals by B cell–predominant areas, follicles, or so-called *malpighian corpuscles.* These B cell–predominant areas contain primary and secondary follicles. Primary follicles consist of only a mantle zone, without germinal centers, whereas secondary follicles contain an inner germinal center in addition to the outer mantle zone (Fig. 2.7). Within the mantle zone are predominantly resting B cells, which express surface IgM/IgD and CD23 (FcεRII). It is within the germinal centers that immunoglobulin class switching, affinity maturation through somatic mutation, and the development of memory B cells occur. Germinal centers are more prevalent at younger ages and diminish with aging. CD4

T cells play a key role in B-cell responses through CD40L and other interactions. The signaling that occurs through this interaction is central to B-cell activation and class switching. In addition to activated B cells and CD4 T cells, the germinal center contains FDCs and macrophages.

At the interface between the white pulp and the red pulp is a region known as the *marginal zone,* which receives blood from branches of central arterioles opening into this region. It contains T cells, as well as subsets of macrophages and B cells. Marginal zone (MZ) B cells are distinct from follicular B cells. They express surface IgM and low levels of IgD and lack CD23. The initial encounter of T cells and B cells with antigen occurs in the marginal zone after blood enters through branches of the central arteriole. Antigen presentation is enhanced by MZ B cells, which are important in T cell–independent responses.

Lymph Nodes and Lymphatics

Lymph nodes occur as chains or groups located along lymphatic vessels. Lymph nodes exist in two major groups: those that drain the skin and superficial tissues (*e.g.,* cervical, axillary, or inguinal lymph nodes) and those that drain the mucosal and deep tissues of the body (*e.g.,* mesenteric, mediastinal, and periaortic lymph nodes). Lymph nodes are oval structures surrounded by adipose tissue with an indentation at the region of a hilus, where blood vessels enter and leave the node (Fig. 2.8). A lymph node is surrounded by a fibrous capsule contiguous with trabeculae traversing the node. Blood vessels and nerves, which enter through the hilus, branch through these trabeculae to various parts of the node. Immediately beneath the capsule is a subcapsular (marginal) sinus. Afferent lymph vessels enter this sinus opposite the hilus. DCs process antigen encountered in skin and migrate into lymph nodes from afferent lymphatics through the subcapsular sinus and into the lymph node. Lymph nodes vary in size, from being barely visible in an unstimulated state to several centimeters in size when undergoing an active immune response.

A lymph node is divided into two major regions, the cortex and the medulla. The cortex contains numerous primary and secondary lymphoid follicles, each approximately 0.5 mm in diameter, similar to those in the spleen. Surrounding the lymphoid follicles in the cortex is the paracortical region, which contains mostly T cells, along with some macrophages and DCs. Both CD4 and CD8 T cells are present, as are macrophages and B cells. The accessory cells, such as interdigitating DCs, present peptide antigens in association with MHC molecules to the TCR on T cells to activate the T cells. Additional accessory molecules (*e.g.,* B7 [CD80] or LFA-3 [CD58]) on the accessory cell and their ligands (CD28 or CD2, respectively) on the T cell provide important costimulatory signals required for activation of the T cell. Other surface antigens, particularly adhesion molecules, such as LFA-1 (CD18) and ICAM-1 (CD54), are involved in stabilizing cellular interactions, as well as providing additional signals between cells.

In the center of the lymph node, beneath the cortex, lies the medulla, which is divided into medullary cords that contain T cells, B cells, plasma cells, and macrophages. Surrounding the medullary cords are medullary sinuses that drain into the hilus. B and T cells migrate from the follicles and paracortical region to the medulla. The Ig produced by the plasma cells drains into medullary sinuses that empty into the hilus. Efferent lymphatic vessels leave the hilus carrying lipids and antibodies, together with mature B and T cells that migrate to other tissues and act as memory B and T cells. The lymphatic vessel system serves to carry lymphocytes derived from various tissue spaces through the network of lymph nodes and to the thoracic duct. Lymphatic capillaries are lined with lymphatic epithelial cells that serve as valves to move lymph fluid, cells, and nutrients around the body. These epithelial cells express high levels of Toll-like receptor-4 (TLR4), which allows them to be activated after lipopolysaccharide delivery to increase lymphaniogenesis.[45] Lymph from the nodes is drawn into the left subclavian vein and back into circulation. Cancer cells found in lymph nodes may take advantage of this system to seed the body. This system of transport develops early in gestation with both lymphatic muscle cells for propulsion and valves that regulate unidirectional lymph flow.

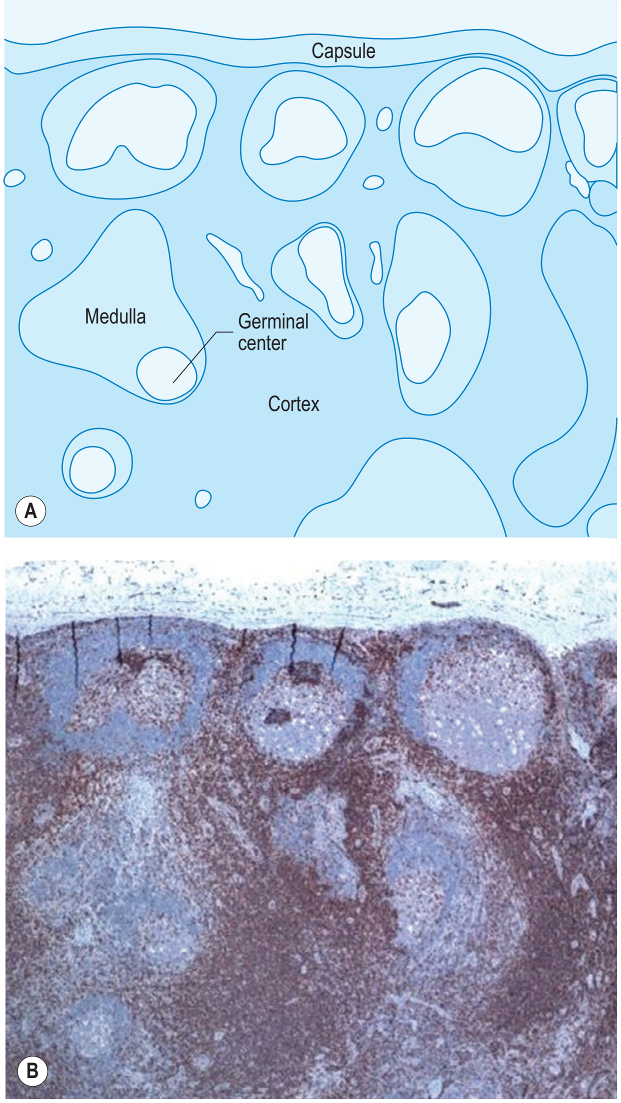


FIG 2.8 Human lymph node showing cortex, medullary areas, and germinal centers.

Adipose Tissue

Adipose tissue has been recently extensively studied in light of the recent obesity epidemic with the realization that immune cells play central roles in adipose homeostasis and in the chronic inflammation in obesity. Macrophages are a central component. They switch from the M2 to the M1 type in obesity. In lean adipose tissue, there are numerous Tregs and few CD8 T cells, which reverses during obesity when inflammation increases.[46]

Mucosal Immune System

The mucosal immune system is located along the surfaces of mucosal tissues (Chapter 20). Each mucosal immune system consists of organized secondary lymphoid structures, termed *mucosa-associated lymphoreticular tissue* (MALT), where inductive immune responses occur, and more diffuse tissues, such as the exocrine glands and lamina propria, where effector immune responses occur.[47] Some mucosal sites, such as the intestine and lungs, have well-developed MALT. Others, such as the vaginal–cervical mucosal surface, have minimally developed MALT.

MALT organization resembles lymph nodes with B cell follicles, intervening T-cell zones and numerous APCs, such as DCs and macrophages. Naïve T and B cells encounter antigen, become activated, exit the tissue via efferent lymphatics, and migrate to the local lymph nodes and then into the thoracic duct and the bloodstream. The cells home to effector sites, particularly the lamina propria of various mucosal tissues. IELs, contained within mucosal epithelium and the lamina propria located beneath the epithelium, are responsible for effector functions. They occur diffusely in mucosal tissues and lack the well-defined structure of the organized mucosal immune system.

The homing of activated lymphocytes from one inductive site to several mucosal surface effector sites has led to the concept of a common mucosal immune system, although significant compartmentalization remains in humans.[48] Trafficking from MALT to the lamina propria is well regulated. Expression of cell surface molecules, such as sphingosine 1 phosphate (S1P), MAdCAM-1, VLA-1, LFA-1, and VCAM-1; the cell surface integrin, α4β7; and chemokines, such as CCR9, CCL25, CCR10, and CCL28, are important in directing activated lymphocytes to the lamina propria surface.[49] The environment at the mucosal surface is favorable for induction of IgA. Low-affinity IgA inhibits commensal bacteria on mucosal surfaces. High-affinity IgA helps neutralize microbial pathogens.

Gastrointestinal Tract

The organized MALT of the GI system is termed *gut-associated lymphoreticular tissue* (GALT). It encompasses Peyer patches, cecal and rectal patches, and isolated lymphoid follicles. Isolated lymphoid follicles and cecal and rectal patches are found throughout the lamina propria and are similar to an individual follicle of a Peyer patch. Peyer patches consist of variably sized aggregates of closely associated lymphoid follicles located in the intestinal lamina propria, occurring predominantly in the ileum (Fig. 2.9).[50] Although these structures arise during fetal life, their full development, with follicles containing germinal centers, does not occur until several weeks after birth, presumably in response to antigenic stimulation. Their number and size increase until puberty and decline thereafter.

Peyer patches and lymphoid follicles have a structural organization that belies their function of presenting antigen from the intestinal lumen to T and B cells. The epithelium overlying lymphoid follicles and Peyer patches—that is, FAE (see Fig. 2.6)—lacks villi and contains very few goblet cells. Particulate antigen uptake via pinocytosis occurs in the FAE through specialized epithelial cells called *M cells* that do not express the polymeric immunoglobulin receptor (secretory component) required for secretion of IgA, expressed by crypt epithelial cells in villous epithelium.[51]

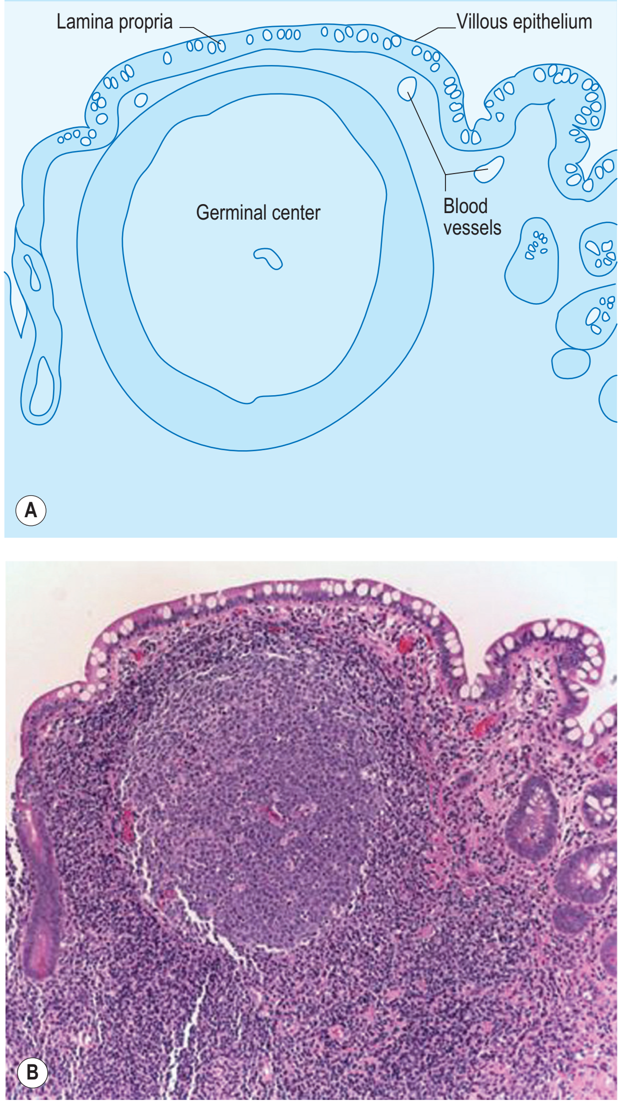


FIG 2.9 Germinal center in terminal ilium.

Beneath the epithelium that overlies individual follicles lies the dome. A substantial number of T cells, including CD4 T cells, are found in this subepithelial region. Macrophages, DCs, and naïve B cells are found here. Antigen, which is pinocytosed by M cells, is transported to the dome region, where antigen presentation to T cells occurs. Macrophages and DCs express high levels of MHC class II. Follicles lying beneath the dome contain mantle zones with predominantly resting B cells, most of which are surface IgM and IgD. Peyer patch follicles have germinal centers that contain activated B cells, FDCs, CD4 T cells, and tingible-body macrophages (so called because of their appearance after they have phagocytosed cellular debris). Many B cells within Peyer patch germinal centers express surface IgA, and it is believed that this is where IgA class switching occurs. Very few CD8 T cells are located within the follicles. An interfollicular region contains predominantly CD4 and CD8 T cells, as well as DCs, macrophages, and some B cells. CD4 T cells predominate over CD8 T cells in this region and the dome.

The diffuse tissue of the GI tract consists of two components: the lamina propria and IELs. The lamina propria is located immediately beneath the epithelium. It contains large numbers of B lymphocytes and plasma cells. A key effector function of the lamina propria is the secretion of antibodies, primarily IgA. IgM represents only 10–18% and IgG 3–5% of all Ig produced from the lamina propria. Two IgA subclasses, IgA1 and IgA2, occur. IgA1 represents greater than 90% of IgA in the respiratory tract and greater than 60% in the lamina propria of the small intestine.[51] IgA2 increases in the lower ileum and becomes predominant in the colon and the rectum. IgA is transported from the lamina propria into epithelial cells through polymeric immunoglobulin receptor-mediated uptake and subsequently secreted into the lumen. Normally, very few IgG B cells are present in the lamina propria. However, under certain inflammatory conditions, such as inflammatory bowel disease, the number of IgG-producing B cells and plasma cells increases dramatically.

The lamina propria also contains large numbers of both CD4 and CD8 T cells in a 2 : 1 ratio. Almost all lamina propria T cells (>95%) express αβ TCR.

Several specialized T cell subsets are present within the GI tract. IELs are found at the basal surface of the epithelium as well as interdigitated with epithelial cells. The vast majority of IEL (>90%) are T cells, which are either $CD8^+$ or $CD4^-CD8^-$. Although the majority of IEL T cells express TCRαβ, a substantial number express TCRγδ. The function of IELs remains incompletely understood, but they can be cytotoxic and also maintain oral tolerance. As part of their effector function, they produce several cytokines, including IL-5 and IFN-γ. Two other T-cell subsets play antagonizing roles in controlling inflammation within the intestinal lamina propria: Tregs and Th17 cells. Tregs are $FOXP3^+$ and function to repress inflammation, whereas Th17 cells mount inflammatory and autoimmune responses through production of IL-17.[52] Another T-cell subset, which comprises NKT cells, expresses the characteristics of both T cells and NK cells. NKT cells express perforin and granzymes but recognize antigens through non-MHC–mediated pathways.[53]

Other cells, including macrophages, DCs, eosinophils, mast cells, and a few neutrophils, are also found in the lamina propria and mediate effector functions. There is an elaborate network of APCs, including DCs and macrophages distributed within the lamina propria and GALT. Two major DC subsets, characterized by $CD103^+CD11b^-$ and $CD103^+CD11b^+$, develop into distinct lineages on the basis of secreted factor requirement. $CD103^+CD11b^+$ are predominately localized to the lamina propria and migrate to the mesenteric lymph nodes upon activation. In contrast, $CD103^+CD11b^-$ populations are localized to the Peyer patches. The GI tract contains the largest number of resident macrophages in the body. These macrophages are similar to macrophages in other tissues and express CD68, lysozyme, ferritin, MHC II, CD11b, CX3CR1, and CD74 but do not migrate to the mesenteric lymph nodes.

Respiratory Tract

Surrounding the entrance to the throat are three tonsillar groups: palatine tonsils, lingual tonsils, and pharyngeal tonsils or adenoids. Together, these are known as *Waldeyer ring.* Tonsils reach full development in childhood and involute by puberty. The palatine tonsils, one located on each side of the pharynx, each measures approximately 2.5 × 1.25 cm. Except at the pharyngeal surface, they are surrounded by a poorly organized capsule that is covered with stratified squamous epithelium. Trabeculae subdivide the tonsil into lobules. Blood vessels and nerves enter through the capsule and extend within trabeculae (Fig. 2.10). The surface of the tonsil is covered by pits, which are the openings of crypts. The crypts extend down into the tissue of the tonsil with branching, increasing surface area. Abundant lymphoid follicles in each lobule contain germinal centers that are predominantly B cells. The lymphoid tissue surrounding the follicles contains T cells, macrophages, DCs, and some B cells. The lingual tonsils consist of 35–100 separate crypts surrounded by lymphoid tissue and are located at the root of the tongue. The pharyngeal tonsils, or adenoids, are accumulations of lymphoid tissue, 2.5–4.0 cm long, located on the median dorsal wall of the nasopharynx. They contain a series of longitudinal folds, but not crypts. The lingual and pharyngeal tonsils also contain lymphoid nodules with germinal centers. The palatine tonsils and adenoids (nasopharyngeal tonsils) comprise the nasopharyngeal–associated lymphoreticular tissues.

Inductive immune responses to inhaled antigens within the respiratory tract occur mainly in the bronchus-associated lymphoid tissue (BALT). BALT consists of lymphoid aggregates located within the bronchial wall near bifurcations of the major bronchial branches (Fig. 2.11). These structures are analogous to the GALT present in the GI tract and function to provide T- and B-cell protection against inhaled microbes. BALT is present at birth and rapidly expands when exposed to antigenic stimulation. The specialized epithelium overlying the lymphoid aggregates consists of M cells heavily infiltrated with lymphocytes and with DCs below the epithelium. The main result of BALT immune induction is secretory IgA production.[54]

The diffuse mucosal tissue of the respiratory tract is minimal. Pools of lymphocytes are present within the lung interstitium, which is 10–20% T cells. Macrophages are present on both the air and the mucosal sides of the lungs and airways.[50] Minimal inflammation occurs in the bronchial mucosa as a result of Tregs that inhibit T-cell activation and expansion. Instead, antigen is carried by local macrophages to the regional lymph nodes, where most respiratory effector immune responses originate. Communication occurs between the GI and respiratory mucosae through cell trafficking. Antigen-reactive T and B cells from Peyer patches can populate the bronchial mucosa. This sharing feature has been exploited in the development of oral vaccines against respiratory microbes.[55]

Genital Tract

The male and female reproductive tracts are components of the common mucosal system. The genital tract immune system must maintain a delicate balance between tolerance of germinal center cells, spermatozoa, and the fetus and the recognition of microbes. The female reproductive tract has been studied the most. Its mucosal immune system is influenced by hormones that regulate all aspects of innate and adaptive immunity.[56] "Professional" APCs, including macrophages and DCs, are present in the stroma of both the uterus and the vaginal tract, where they have unique phenotypes. Reproductive tract NK cells play a role in host defense, implantation of the embryo, and pregnancy and also express a distinct phenotype. CD8 T cells predominate and, along with B cells and macrophages, form unique lymphoid aggregates. Formation of these nodules depends heavily on hormone regulation. Both secretory IgA and IgG are expressed in genital secretions, and levels vary with the stage of the menstrual cycle. The production and transport of antibody produced in the genital tract

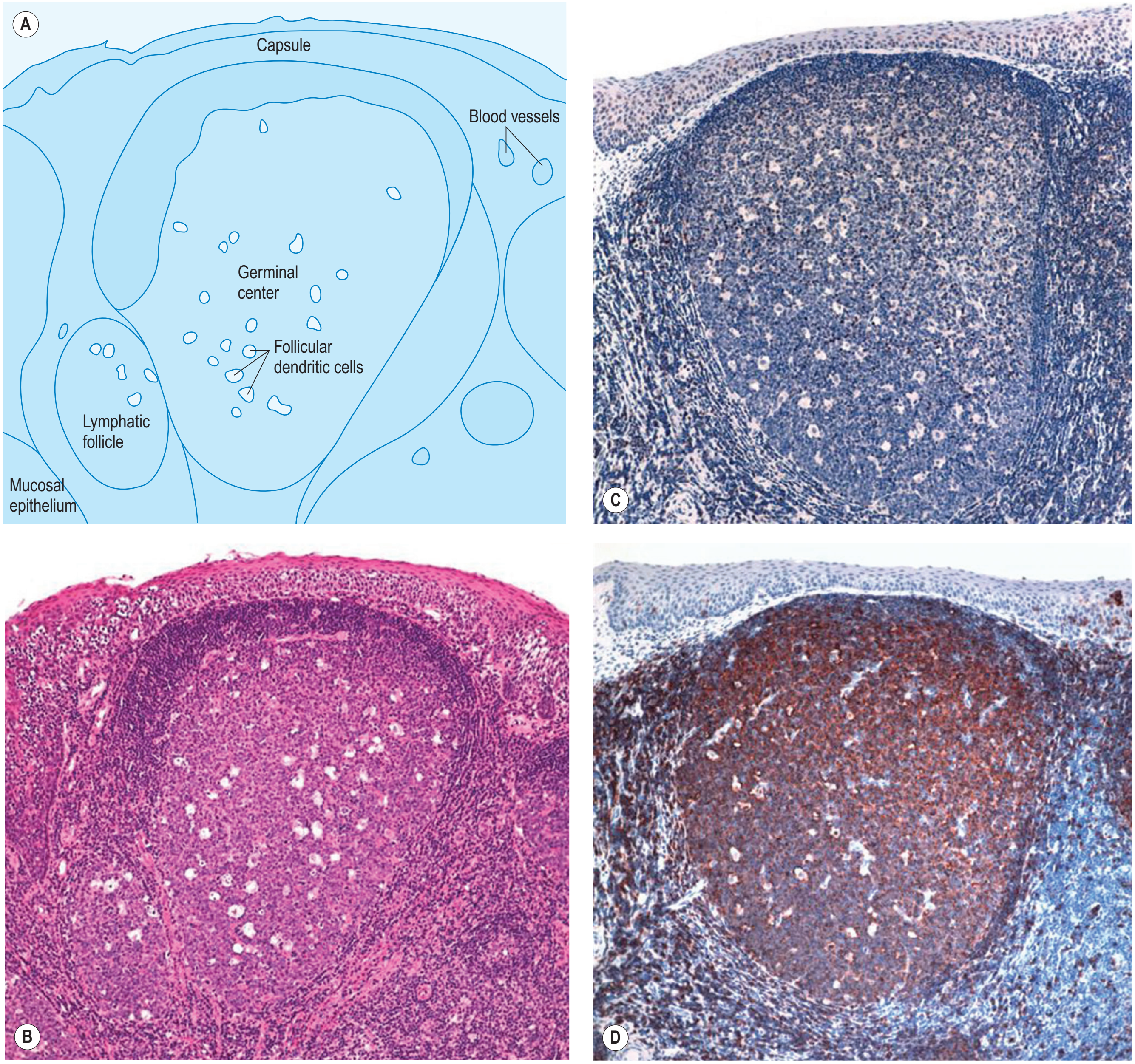


FIG 2.10 Human Tonsil. (A) Organization of lymphoid follicles and germinal centers. (B) Tonsillar tissue stained with hematoxylin and eosin. (C) Tonsillar tissue stained with anti-CD3 to demonstrate the distribution of T cells. (D) Tonsillar tissue stained with anti-CD19 to demonstrate the distribution of B cells.

depends on hormonal and local factors, including IL-1b, IL-6, and IL-10, all of which influence the maturation of B cells to plasma cells within the mucosa.[56]

SKIN

Skin also serves as a specialized secondary immune organ (Chapter 19). It contains two layers, the epidermis and the dermis. The epidermis is the outermost layer and contains three distinct cell types: keratinocytes, melanocytes, and Langerhans cells (Fig. 2.12). Keratinocytes are squamous epithelial cells and are the principal cell type. They secrete a number of cytokines and chemokines, including IL-1, -6, -10, -18, and -33; TGF-β; TNF-α; CXCL9; CXCL10; CXCL11; and CCL20. These can profoundly influence immune cell recruitment and responses. The second cell type is the pigment-producing melanocyte. Melanocytes derive from the neural crest and reside in the basal layer of the epidermis. The third cell type, and the one of particular importance for the immune system, is the Langerhans cell. Langerhans cells are scattered throughout the epidermis within the malpighian, or prickle cell, layer. They are important for both normal and pathologic cutaneous T-cell responses. After encountering antigen in the presence of keratinocyte cytokines, such as TNF-α and IL-6, Langerhans cells migrate from the epidermis to the dermis,

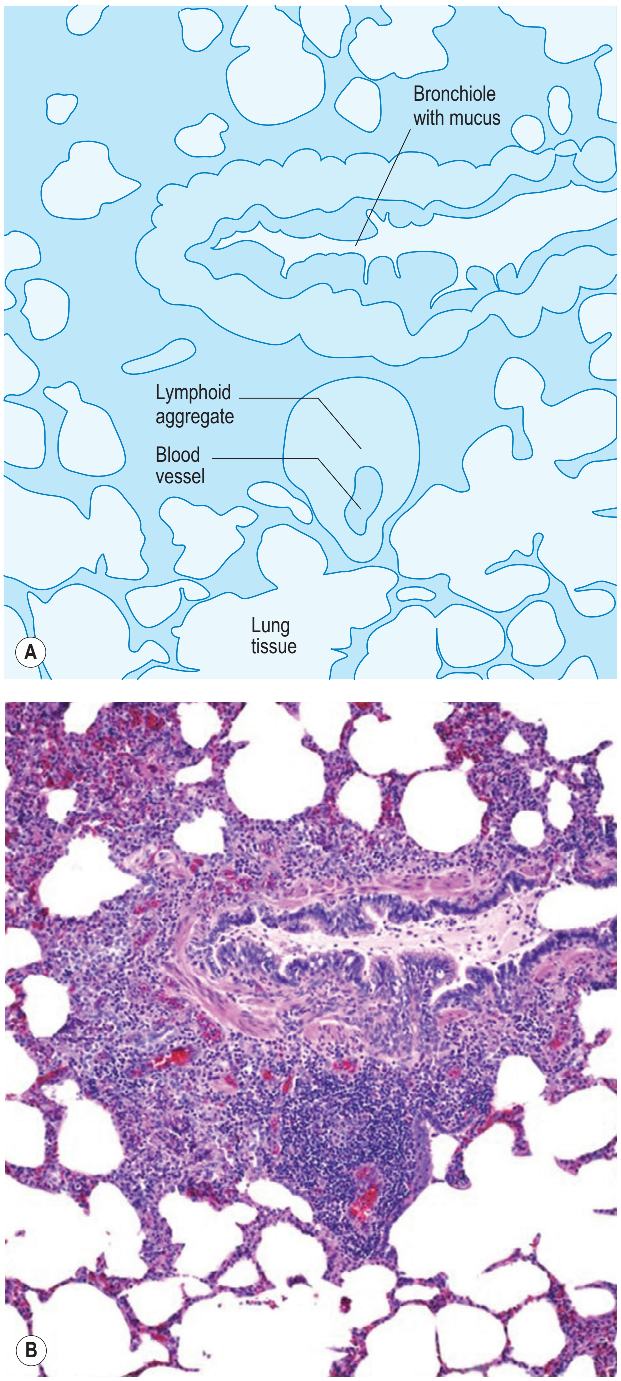


FIG 2.11 Lymphoid regions in the human lung.

enter the afferent lymphatics, and migrate to draining lymph nodes. There they present antigen to T cells to promote primary immune responses. Other types of dermal DCs, including those that are macrophage-like, expressing CD14, and those that are DC-like, expressing CD1c and CD14, have consequences for atopic dermatitis and psoriasis.[57]

The dermis lies under the epidermis. It contains abundant fibroblasts producing collagen, a principal component of skin. The dermis also contains blood vessels and various epidermal adnexal structures, such as hair follicles, sweat glands, and sebaceous glands. The dermal vasculature includes an extensive network of plexuses with arterioles, capillaries, and venules. Dermal lymphatics are associated with the vascular plexuses. In normal skin, a small number of lymphocytes can be found in perivascular areas. These lymphocytes are mostly T cells with distinctive features, including expression of a memory phenotype (CD45RO) and expression of a cutaneous lymphocyte–associated antigen that binds to the vascular addressing endothelial cell leukocyte adhesion molecule-1 (ELAM-1, or CD62E) present on the endothelium. This latter interaction plays an important role in homing of memory T cells to inflamed regions of skin. The dermis also contains mast cells important for immediate hypersensitivity reactions. Tregs are abundant in the skin. CD8 T cells are common in the epidermis. CD4 T cells are common in dermis infected with shingles.[57]

Commensal Organisms/Toll-Like Receptors

Humans live in symbiosis with over 1000 different species of viruses, bacteria, protozoa, and fungi that far outnumber human cells. Collectively termed *commensal microbiota,* these organisms are essential to the development, maturation, organization, and regulation of the mucosal immune system.[58] Many of their immune interactions involve TLR triggering. TLRs activate and prime both innate and specific immune responses. Production of IgA, induction of regulatory T cells, and stimulation of antiinflammatory cytokines are associated with commensal microbiota.[59] Thus the types and quantities of microorganisms present at a mucosal surface is an important component of the mucosal immune response.[60]

ON THE HORIZON

- Understanding how stem cells self-renew is a key to exploiting them for gene therapy.
- Exploiting innate and acquired immune cell function requires an understanding of the multiple subpopulations of cells and the manner in which they are induced.
- Increasing generation of new T cells and B cells later in life might enhance immune function, and thus prolong quality of life.
- The role of adipose tissue in hematopoietic stem cell (HSC) development in bone marrow and control of inflammation in obesity is fundamental to controlling the epidemic of obesity.
- Exploiting interactions between the mucosal immune system and commensal populations may improve health, prevent inflammation, and lead to less antibiotic use.
- Cellular migration via the lymphatics to the brain is unknown; inflammation and cancer spread might be reduced by controlling inflammation of the lymphatics.

ACKNOWLEDGMENTS

We thank Dr. Gregory R. Harriman at BioAdvance for his work on previous editions of this chapter; Dr. Edwina Popek, Pathology Department, Texas Children's Hospital, Houston, Texas, for providing histopathologic images of lymphoid tissues; Dr. Gregory Stelzer and Wendy Schober for the flow cytometric display; Eleanor Chapman, Anna Wirt, Terry Saulsberry, Yvette Wyckoff, and Pamela Love for help with the manuscript; and Dr. Jerry McGhee for critical review of the first edition.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

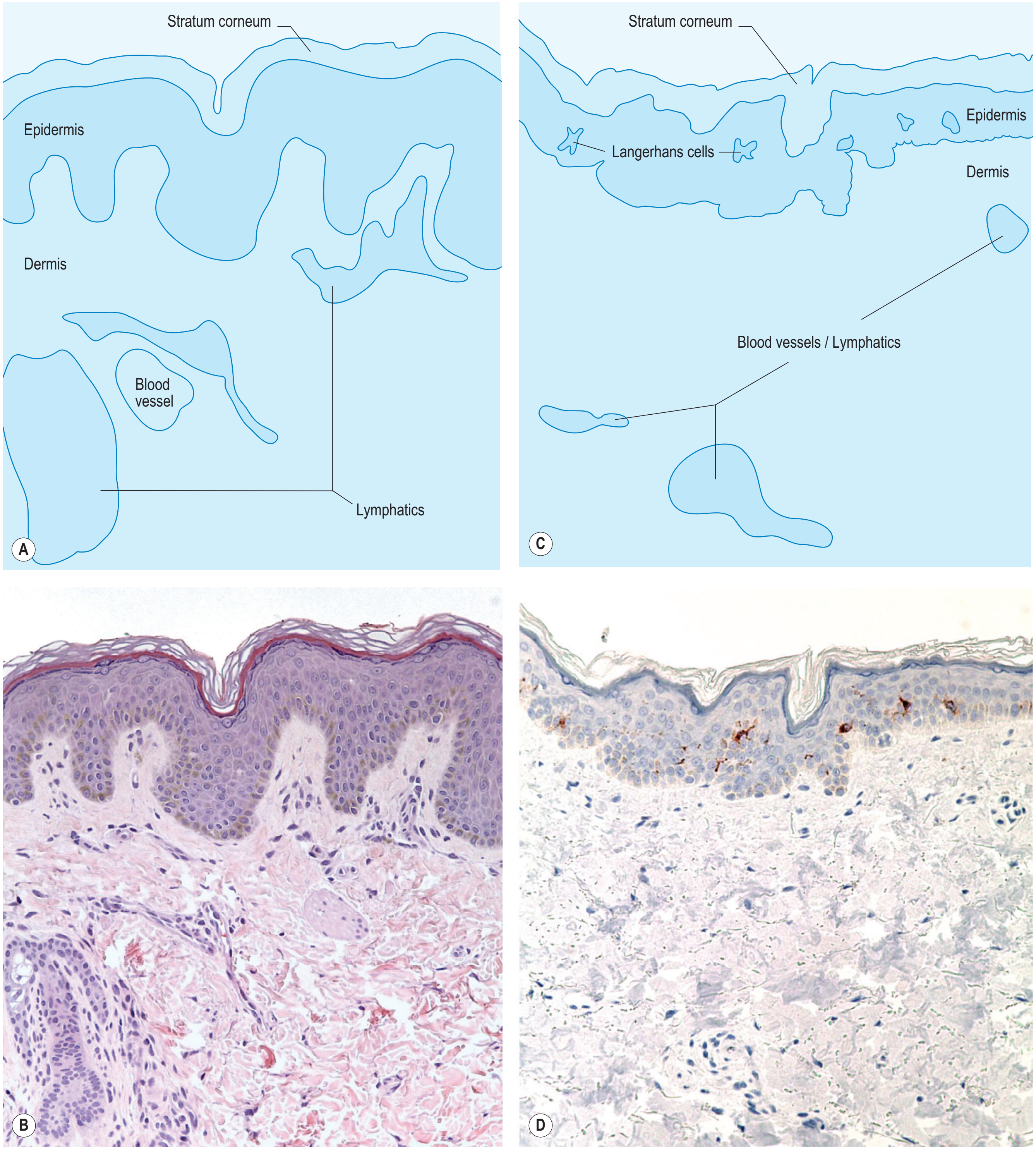

Stratum corneum
Epidermis
Dermis
Blood vessel
Lymphatics
A
Stratum corneum
Epidermis
Langerhans cells
Dermis
Blood vessels / Lymphatics
C
B
D

FIG 2.12 **Lymphoid Regions in Human Skin.** (A, C) Organization of the epithelial tissue. (B) Epithelial tissue stained with hematoxylin and eosin. (D) Epithelial tissue stained with anti-CD207 to demonstrate the distribution of Langerhans cells *(note brown cells).*

REFERENCES

1. Rampon C, Huber P. Multilineage hematopoietic progenitor activity generated autonomously in the mouse yolk sac: analysis using angiogenesis-defective embryos. Int J Dev Biol 2003;47:273–80.
2. Ciriza J, Thompson H, Petrosian R, et al. The migration of hematopoietic progenitors from the fetal liver to the fetal bone marrow: lessons learned and possible clinical applications. Exp Hematol 2014;41:411–23.
3. Smith JNP, Calvi LM. Current concepts in bone marrow microenvironmental regulation of hematopoietic stem and progenitor cells. Stem Cells 2013;6:1044–50.
4. Copley MR, Eaves CJ. Developmental changes in hematopoietic stem cell properties. Exp Mol Med 2013;45:e55.

5. Pang WW, Price EA, Sahoo D, et al. Human bone marrow hematopoietic stem cells are increased in frequency and myeloid-biased with age. Proc Natl Acad Sci 2011;108(50):20012–17.
6. Doulatov S, Notta F, Laurenti E, et al. Hematopoiesis: a human perspective. Cell Stem Cell 2012;10:120–36.
7. Metcalf D. Hematopoietic cytokines. Blood 2008;111:485–91.
8. Yu J, Freud AG, Caligiuri MA. Location and cellular stages of natural killer cell development. Trends Immunol 2013;34:573–82.
9. Wilson A, Laurenti E, Oser G, et al. Hematopoietic stem cells reversibly switch from dormancy to self-renewal during homeostasis and repair. Cell 2008;135:1118–29.
10. Zhang CC, Lodish HF. Cytokines regulating hematopoietic stem cell function. Curr Opin Hematol 2008;15(4):307–11.
11. Kondo M, Wagers MJ, Manz MG, et al. Biology of hematopoietic stem cells and progenitors: implications for clinical application. Annu Rev Immunol 2003;21:759–806.
12. Varol C, Mildner A, Jung S. Macrophages: development and tissue specialization. Annu Rev Immunol 2015;33:643–75.
13. Hoeffel G, Ginhoux F. Ontogeny of tissue-resident macrophages. Front Immunol 2015;6:486.
14. Bouhlel MA, Derudas B, Rigamonti E, et al. PPARγ activation primes human monocytes into alternative M2 macrophages with anti-inflammatory properties. Cell Metab 2007;6:137–43.
15. Boltjes A, van Wijk F. Human dendritic cell functional specialization in steady-state and inflammation. Front Immunol 2014;5:131.
16. Mende I, Karsunky H, Weissman IL, et al. Flk2þ myeloid progenitors are the main source of Langerhans' cells. Blood 2006;107:1383–90.
17. Randolph GJ, Angeli V, Swartz MA. Dendritic-cell trafficking to lymph nodes through lymphatic vessels. Nature Immunol Rev 2005;5:617–28.
18. Friedman AD. Transcriptional regulation of granulocyte and monocyte development. Oncogene 2002;21:3377–90.
19. Kruger P, Saffarzadeh M, Weber A, et al. Neutrophils: between host defence, immune modulation, and tissue injury. PLoS Pathog 2015;11(3):1–23.
20. Lieber JG, Webb S, Suratt BT, et al. The *in vitro* production and characterization of neutrophils from embryonic stem cells. Blood 2004;103:852–9.
21. Rosenberg HF, Dyer KD, Foster PS. Eosinophils: changing perspectives in health and disease. Nat Rev Immunol 2013;13:9–22.
22. Chirumbolo S. State-of-the-art review about basophil research in immunology and allergy: is the time right to treat these cells with the respect they deserve? Blood Transfus 2012;10:148–64.
23. Hattangadi SM, Wong P, Zhang L, et al. From stem cell to red cell: regulation of erythropoiesis at multiple levels by multiple proteins, RNAs, and chromatin modifications. Blood 2011;118:6258–68.
24. Woolthuis CM, Park CY. Hematopoietic stem/progenitor cell commitment to the megakaryocyte lineage. Blood 2016;127:1242–8.
25. Blom B, Spits H. Development of human lymphoid cells. Annu Rev Immunol 2006;24:287–320.
26. Haddad R, Guimiot F, Six E, et al. Dynamics of thymus colonizing cells during human development. Immunity 2006;24:217–30.
27. Spits H. Development of alpha beta T cells in the human thymus, thymic involution. Nature Immunol Rev 2002;2:760–72.
28. Gary DHD, Ueno T, Chidgey AP, et al. Controlling the thymic microenvironment. Curr Opin Immunol 2005;17:137–43.
29. Takhama Y. Journey through the thymus: stromal guides for T-cell development and selection. Nat Rev Immunol 2006;6:127–36.
30. Mahnke YD, Brodie TM, Sallusto F, et al. The who's who of T-cell differentiation: human memory T-cell subsets. Euro J Immunol 2013;43:2797–809.
31. Synder-Cappione JE, Tincati C, Eccles-James IG, et al. A comprehensive ex vivo functional analysis of human NKT cells reveals production of MIP1 alpha and MIP-1 beta, a lack of IL17 and a Th1bias in males. PLoS ONE 2010;5:15412–20.
32. Kronenberg M. Towards an understanding of NKT cell biology: progress and paradoxes. Annu Rev Immunol 2005;26:877–900.
33. Jonuleit H, Schmitt E. The regulatory T cell family: distinct subsets and their interrelations. J Immunol 2003;171:6323–7.
34. Nagasawa T. Microenvironmental niches in the bone marrow required for B-cell development. Nature Immunol Rev 2006;6:107–16.
35. Tobón GJ, Izquierdo JH, Cañas CA. B Lymphocytes: development, tolerance, and their role in autoimmunity—focus on systemic lupus erythematosus. Autoimmune Dis 2013;2013:827254.
36. Herzenberg LA, Tung JW. B cell lineages: documented at last. Nat Immunol 2006;7:225–6.
37. Barone F, Vossenkamper A, Boursier L, et al. IgA producing plasma cells originate from germinal centers that are induced by B cell receptor engagement in humans. Gastroenterology 2011;140:947–56.
38. Moretta L, Moretta A. Unravelling natural killer cell function: triggering and inhibitory human NK receptors. EMBO J 2004;23:255–9.
39. Farag SS, Caligiuri MA. Human natural killer cell development and biology. Blood 2006;20:123–37.
40. Fu B, Tian Z, Wei H. Subsets of human natural killer cells and their regulatory effects. Immunology 2013;141:483–9.
41. Artis D, Spits H. The biology of innate lymphoid cells. Nature 2015;517:293–301.
42. Guyden JC, Martinez M, Chilukuri RV, et al. Thymic nurse cells participate in heterotypic internalization and repertoire selection of immature thymocytes: their removal from the thymus of autoimmune animals may be important to disease etiology. Curr Mol Med 2015;15:828–35.
43. Hanabuchi S, Watanabe N, Liu Y-J. TSLP and immune homeostasis. Allergol Intl 2012;61:19–25.
44. Steiniger BS. Human spleen microanatomy: why mice do not suffice. Immunology 2015;145:334–46.
45. Choi I, Lee S, Hong Y-K. The new era of the lymphatic system: no longer secondary to the blood vascular system. Cold Spring Harb Perspect Med 2012;2:a006445.
46. Ferrante AW. The immune cells in adipose tissue. Diabetes Obes Metab 2013;15:34–8.
47. Brandtzaeg P. Mucosal immunity: induction, dissemination and effector functions. Scand J Immunol 2009;70:505–15.
48. Kunizawa J, Nochi T, Hiroshi K. Immunological commonalities and distinctions between airway and digestive immunity. Trends Immunol 2008;29:505–13.
49. Wright P. Inductive/effector mechanisms for humoral immunity at mucosal sites. Am J Reprod Immunol 2011;65:248–52.
50. Newberry RD, Lorenz RG. Organizing a mucosal defense. Immunol Rev 2005;206:6–21.
51. Bienenstock J, McDermott MRM. Bronchus- and nasal-associated lymphoid tissues. Immunol Rev 2005;206:22–31.
52. Grainger JR, Askenase MH, Guimont-Desrochers F, et al. Contextual functions of antigen-presenting cells in the gastrointestinal tract. Immunol Rev 2014;259:75–87.
53. Montalvillo E, Garrote JA, Bernardo D, et al. Innate lymphoid cells and natural killer T cells in the gastrointestinal tract immune system. Rev Esp Enferm Dig 2014;106:334–45.
54. Kyd JM, Foxwell AR, Cripps AW. Mucosal immunity in the lung and upper airway. Vaccine 2001;19:2527–33.
55. Boyaka PN, Tafaro A, Fischer R, et al. Therapeutic manipulation of the immune system: enhancement of innate and adaptive mucosal immunity. Curr Pharm Des 2003;9:1965–72.
56. Hickey DK, Patel MV, Fahey JV, et al. Innate and adaptive immunity at mucosal surfaces of the female reproductive tract: stratification and integration of immune protection against the transmission of sexually transmitted infections. J Reprod Immunol 2011;88:185–94.
57. Heath WR, Carbone FR. The skin-resident and migratory immune system in steady state and memory: innate lymphocytes, dendritic cells and T cells. Nature Immunol 2013;14:978–85.
58. Spasova DS, Surh CD. Blowing on embers: commensal microbiota and out immune system. Front Immunol 2014;5:318.
59. Hevia A, Delgado S, Sanchez B, et al. Molecular players involved in the interaction between beneficial bacterial and the immune system. Front Microbiol 2015;6:1285.
60. Kurashima Y, Goto Y, Kiyono H. Mucosal innate immune cells regulate both gut homeostasis and intestinal inflammation. Euro J Immunol 2013;43:3108–15.

MULTIPLE-CHOICE QUESTIONS

1. Characteristics of hematopoietic stem cells include:
   A. Deleted as progeny develop
   B. Derive all lineages of blood cells
   C. Form distinct niches in fetal liver
   D. Rely on stromal factors to regulate differentiation
2. Interactions between innate and acquired immune cells involve (more than one correct answer):
   A. Processing and presentation of antigens
   B. Stimulation of macrophages by TLR-ligands
   C. Only tissue based interactions
   D. Migration of immature dendritic cells to lymph nodes
3. Mucosal immune sites include (more than one correct answer):
   A. Vaginal and rectal tissues
   B. M cells which transport antigens across the lumen
   C. Inguinal and axial lymph nodes
   D. Mesenteric lymph nodes
4. Lymph nodes are (more than one correct answer):
   A. Located throughout the body
   B. Delivery sites for phagocytized organisms
   C. Location of T- and B-cell rearrangements
   D. Peyer patches in the gastrointestinal tract

3

Innate Immunity

Douglas R. McDonald, Ofer Levy

Innate immunity is the first line of host defense against infection. All living organisms are continually exposed to microbes. For example, the human gut is colonized by trillions of commensal bacteria. The innate immune system must accommodate commensal bacteria but be able to be activated by pathogens (*e.g., Salmonella* or *Shigella*). Potentially life-threatening infections can result from naturally occurring defects in the innate immune response (Chapter 36).

A defining characteristic of innate immunity is its existence before microbial exposure. Innate immune responses are initiated rapidly by microbes and precede the development of adaptive immune responses. The adaptive immune system is characterized by the tremendous diversity of its receptors and its antigen ligands. The innate immune system responds to a more constrained set of antigens that are typically essential and invariant structural components specific to microbes. These microbial components are known as *pathogen-associated molecular patterns* (PAMPs).[1] They include microbial cell wall components and nucleic acids. PAMPs are recognized by pattern recognition receptors (PRRs), and they are highly potent and effective in initiating inflammatory responses.

"Trained immunity" is used to describe the phenomenon of enhanced innate immune responses following microbial exposure.[2,3] This increase in host resistance to reinfection can provide "cross-protection" against other infectious agents. For example, macrophages and natural killer (NK) cells can expand and contract their cell populations, upregulate genes involved in pathogen recognition and presentation, and secrete cytokines that augment the antimicrobial activity of bystander cells. Thus there is a growing appreciation that the adaptive and innate immune systems have certain similar characteristics.

BARRIERS TO INFECTION

Skin and Mucosa

The epithelial layers of the skin and the linings of the gastrointestinal (GI), genitourinary (GU), and respiratory tracts provide a mechanical barrier to microbial entry and thus play an essential role in host defense. The stratum corneum of the skin is the first barrier encountered by microbes (Chapter 19). The skin is persistently colonized with numerous microbes. Thus an intact physical barrier is essential to prevent activation of the immune system under nonpathological conditions. Key cellular components of the skin immune barrier include keratinocytes, dendritic cells (DCs), macrophages, T lymphocytes, and mast cells. These cells express a wide variety of pathogen recognition receptors and secrete a broad range of cytokines, chemokines, and antimicrobial proteins and peptides (APPs) that mediate inflammatory responses to infection. Genetic disorders of the skin that compromise skin integrity, such as epidermolysis bullosa (Chapter 63), can result in life-threatening infections.

Skin disorders that impair barrier function, such as atopic dermatitis (AD) (Chapter 44) or eczema, are common.[4] Filaggrin (FLG) is a key a structural component of the outermost layer of the epidermis. Loss of function variants in filaggrin (R510X, 2282del4) is estimated to be present in up to 50% of patients with AD.[4] FLG mutations are a risk factor for the development of early-onset AD and thus for sensitization to food allergens (Chapter 45), allergic rhinitis, and asthma (Chapter 72) (the atopic march). Eczematous skin can lead to reduced expression of APPs and increased susceptibility to cutaneous bacterial (*e.g., Staphylococcus*, *Streptococcus*) and viral (*e.g.,* herpes) infections.[5]

The luminal surfaces of the intestines are sites of continual exposure to massive numbers of microbes. Intestinal epithelial cells (IECs) (Chapter 20) protect against infection by forming a physical barrier through tight junctions and by producing mucus (goblet cells) and APPs. IECs express apical junction complexes, including E-cadherin, ZO-1, claudin, and occludin, which function to form a tight monolayer that prevents penetration by bacteria.[6] A breakdown in epithelial gut homeostasis can lead to inflammatory bowel diseases (*e.g.,* Crohn disease, ulcerative colitis) (Chapter 75) and increased susceptibility to bacterial infection.[7]

KEY CONCEPTS

The Innate Immune System

- The innate immune system provides the initial immune response to pathogens.
- Although less specific than the adaptive immune system, the innate immune system must differentiate commensal from pathological microbes.
- The innate immune system comprises barriers to the environment (*e.g.,* skin, mucosa), antimicrobial peptides and proteins, cells (*e.g.,* neutrophils, macrophages, monocytes), and soluble factors (*i.e.,* cytokines, chemokines, complement).
- Pathogen detection is mediated by a variety of germline-encoded pathogen recognition receptors (PRRs) that can recognize invariant microbial structures known as pathogen-associated molecular patterns (PAMPs).
- Activation of the innate immune system leads to subsequent activation of the adaptive immune system.
- The innate immune system has a form of memory or "trained immunity" such that innate immune activation can modulate innate immune responses to subsequent unrelated stimulus/infection.

Influenza viruses and respiratory syncytial virus replicate in airway epithelial cells, leading to cell death and inflammation. The impaired barrier function of the airways can lead to increased susceptibility to secondary invasive bacterial infections by *Streptococcus pneumoniae* and other pyogenic bacteria. Inflammatory bowel diseases also result in impaired barrier functions of the small and large intestines, which can be associated with increased translocation of bacteria across gut mucosa, potentially leading to serious infection.

CLINICAL PEARLS

Innate Immunity Barriers

- Innate immune barriers consist of epithelial layers, including those of skin and the gastrointestinal, respiratory, and genitourinary tracts.
- Barrier function is an underappreciated component of the innate immune system.
- Defects of barrier function, such as epidermolysis bullosa and atopic dermatitis, increase the risk of infection.
- Production of antimicrobial peptides and proteins at barrier sites plays a vital role in preventing invasion by microbes.

Antimicrobial Proteins and Peptides

Among the APPs produced by the skin, GI, GU, and respiratory tract epithelia are bactericidal/permeability-increasing protein (BPI), defensins (β-strand peptides connected by disulfide bonds), and cathelicidins (linear α-helical peptides)[8] (Table 3.1). Most APPs have a net positive charge, which enhances their affinity for negatively charged microbial cell membranes. Binding of APPs to microbes can increase membrane permeability and target cell death.

BPI is a ~55-kilodalton (kDa) cationic and hydrophobic protein with high affinity for the lipid A region of lipopolysaccharide (endotoxin). It is found in neutrophil primary (azurophilic) granules and is also inducible in epithelial cells. BPI inhibits gram-negative bacteria by its endotoxin neutralizing and microbicidal and opsonic properties.[9] Neutralization of endotoxin may also help limit inflammatory responses to gram-negative bacteria.

Some APPs, such as lysozyme (Lz), have enzymatic activities, which cleaves peptidoglycans found in bacterial cell walls. Other APPs bind to and compete for nutrients, a form of so-called nutritional immunity. Lactoferrin (Lf), for example, binds iron, a nutrient essential to bacterial survival.[9]

Defensins are classified by the linking pattern of cysteines and their sizes. α-defensins are expressed in neutrophils and Paneth cells of the small intestine, whereas β-defensins are expressed by mucosal surface epithelia, including those of skin, eyes, and the oral, urogenital, and respiratory tracts.[10] Defensins have a broad specificity of antimicrobial activities against bacteria, mycobacteria, fungi, parasites, and viruses (Table 3.2). They have also been shown to enhance antigen uptake and processing, and to stimulate the chemotaxis of monocytes, macrophages, and mast cells.[10,11] The expression of several of the defensins is constitutive. For others, inflammatory stimuli (bacterial products, proinflammatory cytokines) will increase defensin expression (human neutrophil proteins 1–3 and human β-defensin-2). Given the increasing incidence of antibiotic resistant bacteria, there is great interest in the potential uses of APPs as treatment for bacterial infections and infections with multidrug-resistant organisms.[12,13]

HUMORAL INNATE IMMUNITY

The Acute Phase Response

A variety of soluble proteins found in plasma help recognize PAMPs and function as mediators of innate immunity. Tumor

TABLE 3.1 Epithelial Antimicrobial Proteins and Peptides (APPs)

Antimicrobial Peptide	Source	Target Organism
Dermicidin	Eccrine sweat glands	Broad spectrum
Psoriasin	Keratinocytes, sebocytes	G^-
RNase 7	Keratinocytes	Broad spectrum
RNase 5/angiogenin	Keratinocytes	*C albicans*
Cathelicidin (LL-37)	Keratinocytes, sebocytes	G^+, G^-
BPI	Epithelia-oral, GI, urogenital tract	G^-, (G^+, fungi)
hBD-1	Keratinocytes, sebocytes	G^-
hBD-2	Keratinocytes, sebocytes	G^-
hBD-3	Keratinocytes	Broad spectrum
hBD-4	Keratinocytes	G^+, G^-
SLPI	Keratinocytes	Broad spectrum
Elafin	Keratinocytes	Broad spectrum
Adrenomedullin	Keratinocytes, hair follicles, eccrine/apocrine sweat glands, sebocytes	G^+, G^-
MIP-3α/CCL20	Keratinocytes	Broad spectrum
Lysozyme	Keratinocytes, sebocytes, hair bulb cells	G^+, G^-
Lactoferrin	Milk, saliva, tears, nasal secretions, neutrophils	Broad spectrum

RNase, ribonuclease; BPI, bactericidal/permeability-increasing protein; CCL, chemokine ligand; G^+, gram-positive; G^-, gram-negative; GI, gastrointestinal; hBD, human β-defensin; MIP, macrophage inflammatory protein; SLPI, secretory leukocyte peptidase inhibitor.

TABLE 3.2 Neutrophil-Derived Antimicrobial Proteins and Peptides (APPs)

Neutrophil APP	Granule Type	Target Organism
Lysozyme	Azurophil, specific	G^+, G^-
Azurocidin	Azurophil, secretory	G^+, G^-*Candida albicans*
Elastase	Azurophil	G^+, G^-
Cathepsin G	Azurophil	G^+, G^-
Proteinase 3	Azurophil	G^+, G^-
BPI	Azurophil	G^-, (G^+, fungi)
α-defensins (HNP-1 to -4)	Azurophil	G^+, G^-, fungi, viruses
Cathelicidin (hCAP-18)	Specific	G^+, G^-, mycobacteria
Lactoferrin	Specific	G^+, G^-, fungi, viruses
SLPI	Specific	G^+, G^-, *Aspergillus fumigatus, C. albicans*
NGAL	Specific	G^+, G^-, fungi
Lysozyme	Azurophil, specific	G^+, G^-
Azurocidin	Azurophil, secretory	G^+, G^-, *C. albicans*
Elastase	Azurophil	G^+, G^-
Cathepsin G	Azurophil	G^+, G^-

BPI, bactericidal/permeability-increasing protein; G^+, gram-positive; G^-, gram-negative; hCAP, human cathelicidin antimicrobial protein; HNP, human neutrophil peptide; NGAL, neutrophil gelatinase-associated lipocalin; SLPI, secretory leukocyte peptidase inhibitor.

KEY CONCEPTS

Humoral Innate Immunity

- Cytokines and chemokines are essential mediators of the innate immune response.
- Cytokines are redundant and pleiotropic. To avoid host tissue response, their synthesis is tightly controlled.
- Acute phase reactants (*i.e.,* cross-reactive protein [CRP]) are induced by cytokines (interleukin [IL]-6) and play roles in opsonization of microbes. Plasma CRP is used to monitor infections and inflammation.
- Defects in the complement system result in invasive bacterial infections, primarily from encapsulated bacteria.

necrosis factor-α (TNF-α) and interleukin-1β (IL-1β) induce production of acute phase reactants in hepatocytes, including members of the pentraxin family (*e.g.,* serum amyloid A [SAA], serum amyloid P [SAP], and cross-reactive protein [CRP]). These pentraxins bind to components of the bacterial cell wall.[14] TNFα and IL-1β also induce production of IL-6 from mononuclear phagocytes, endothelial cells, and fibroblasts. IL-6 is another potent inducer of acute phase reactants, including CRP and fibrinogen. CRP, SAA, and SAP function as opsonins and can bind phosphorylcholine and phosphatidylethanolamine expressed on bacteria and apoptotic cells, enhancing phagocytosis of bacteria and apoptotic cells by macrophages.

Lipopolysaccharide-binding protein (LBP) is an acute phase reactant synthesized by the liver in response to gram-negative bacterial infections. LBP binds to LPS and subsequently forms a complex with CD14, TLR4, and MD-2, which functions as a high-affinity receptor for LPS.[15]

Mannose-binding lectin (MBL) is a member of the calcium-dependent (C-type) lectins (collectins) produced by the liver in response to infection. MBL binds to carbohydrates with terminal mannose and fucose residues that are expressed on microbial cell surfaces.[16] MBL can bind to the C1q receptor on macrophages to enhance phagocytosis and can activate the complement system via the lectin pathway (discussed below).

Surfactant protein-A and surfactant protein-D are collectins expressed in the lung and can bind a variety of microbes and inhibit their growth.[13,17] They also function as opsonins that promote ingestion by alveolar macrophages.

Finally, ficolins are plasma proteins capable of binding to several types of bacteria and can activate complement.

CLINICAL PEARLS

Complement

- Deficiency of early components of the complement cascade create susceptibility to invasive infections with encapsulated bacteria and development of a lupus-like syndrome.
- Deficiency of late components of the complement cascade (C5–9) result in susceptibility to meningitis caused by *Neisseria meningitis.*
- Deficiency of C1-inhibitor protein (or function) results in hereditary angioedema.
- Deficiency of factor H is associated with development of membranoproliferative glomerulonephritis, hemolytic–uremic syndrome, and age-related macular degeneration.
- Deficiency of mannose-binding lectin can result in susceptibility to bacterial infection in individuals with comorbid conditions (*e.g.,* cystic fibrosis).

The Complement System

The complement system comprises a collection of plasma proteins activated by microbes (Chapter 21). It helps mediate microbial destruction and inflammation.[18] Complement activation can occur via three pathways: the classical pathway (CP), the alternative pathway (AP), and the lectin pathway (LP).

In the CP, complement C1 detects immunoglobulin M (IgM), IgG1, or IgG3 bound to the surface of a microbe. C1 is composed of the C1q, C1r, and C1s subunits. These form multimeric complexes that recognize IgM or IgG bound to microbial surfaces. C1r and C1s are serine proteases. Activated C1s generates a C3 convertase composed of C4b and C2b (C4b2b) bound to the microbial surface. C3 convertase cleaves C3, generating C3b. C3b binds covalently to C4b2b, generating C5 convertase. C5 convertase then activates the late steps of complement activation leading to assembly of the membrane attack complex (MAC) and subsequent cytolysis.

The AP is initiated by small amounts of C3b, which are spontaneously generated in plasma. C3b that remains unbound to a cell surface is rapidly hydrolyzed and inactivated. C3b bound to a microbe becomes a binding site for factor B. Bound factor B is cleaved by factor D, generating factor Bb that binds covalently to C3b, forming the AP C3 convertase, which activates the late steps of complement activation, as in the CP.

The lectin pathway is activated by MBL or ficolins binding to microbial surfaces. MBL then binds to MBL-associated serine proteases (MASPs)-1, -2, and -3. MASP-2 cleaves C4 and C2 to activate the complement cascade, as in the CP (Fig. 3.1).

Complement components also function as opsonins. Complement-coated microbes can be phagocytosed via complement receptors on phagocytes. Complement receptor type 1 (CR1) is a high-affinity receptor for the C3b and C4b fragments of complement and mediates the internalization of C3b- and C4b-coated particles. On erythrocytes, it mediates clearance of immune complexes from the circulation. Complement type 2 receptor (CR2, also known as CD21) is expressed on B cells and follicular dendritic cells (FDCs). It binds C3 proteolytic fragments, including C3d, C3dg, and iC3b. CR2 augments humoral immune responses by enhancing B-cell activation by antigen and by promoting trapping of antigen–antibody complexes in germinal centers.[19] CR2 is also the receptor for Epstein-Barr virus (EBV), allowing EBV to enter B cells. Complement receptor 3 (CR3) is composed of a heterodimer of CD18 and CD11b and is expressed in polymorphonuclear neutrophils (PMNs) and monocytes or macrophages. CR3 binds to iC3b bound to the surface of microbes, leading to phagocytosis and destruction of the pathogen. Activation of complement via the AP can greatly enhance monocyte-generated TNF-α elicited by gram-positive bacteria, such as group B *Streptococcus.*[20]

There are multiple regulatory proteins within the complement pathways. C1-inhibitor (C1-INH) binds to, and inhibits, the enzymatic functions of C1r and C1s within the CP (21). Properdin stabilizes C3bBb complex, increasing the life span of the AP C3 convertase. Conversely, factor H inhibits the formation of, and degrades, C3bBb complex. Factor I inactivates C3b. CD55 (decay accelerating factor) and CD59 are cell surface, GPI-linked proteins that block complement-mediated cytolysis by inhibiting formation of C3bBb complex and binding of C9 to C5b678 complex, respectively. Paroxysmal nocturnal hemoglobinuria, an acquired defect in the *PIGA* gene that causes a deficiency of GPI-linked proteins, is the result of absent cell surface expression of CD55

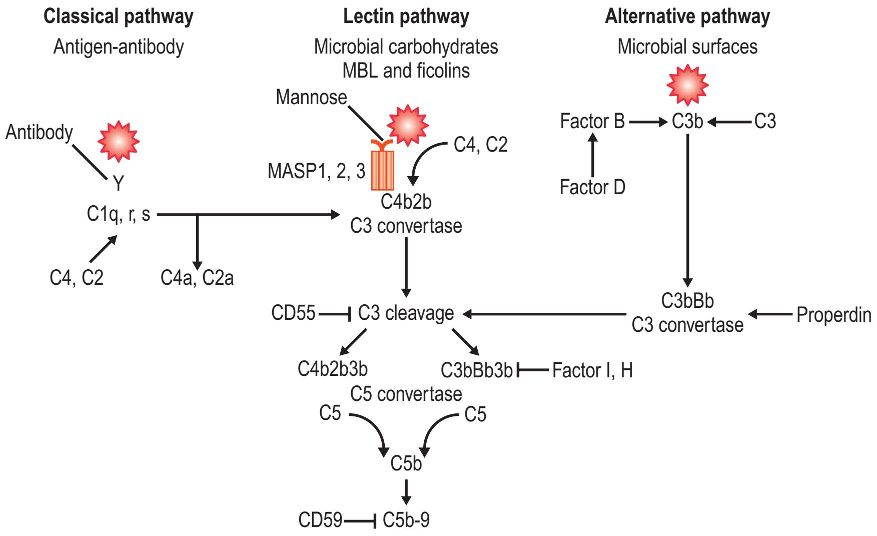


FIG 3.1 **Complement Activation Pathways.** The classical complement cascade is activated by antibody bound to microbial surfaces, which is a binding site for the C1 complex. The alternative pathway is activated by the binding of spontaneously generated C3b to microbial surfaces. Microbe-bound C3b binds factor B, which is converted to factor Bb, forming a C3 convertase. The lectin pathway is activated by the binding of mannose-binding lectin (MBL) to mannose residues on microbial surfaces. MBL binds MBL-associated serine proteases, which bind and cleave C4 and C2, forming a C3 convertase.

and CD59 that leads to hemolytic anemia caused by complement-mediated lysis of red blood cells (RBCs).

Complement Deficiency Diseases

Deficiencies of early components of the complement pathway are associated with invasive bacterial infections caused by encapsulated organisms (Chapter 21). Lack of early components of the complement pathway are also associated with rheumatic disorders, including a lupus-like syndrome that may be caused by impaired immune complex clearance, impaired clearance of apoptotic cells, and loss of complement-dependent B cell tolerance (Chapter 50). Deficiency of factor I is also associated with increased incidence of invasive infection with encapsulated bacteria, as well as glomerulonephritis and autoimmune disease. Deficiency of C1-INH protein and function, either hereditary or acquired, leads to hereditary angioedema (HAE) or acquired angioedema (AAE) (Chapter 42). C1-INH inhibits C1, factors XIa and XIIa, and kallikrein. Dysregulation of these cascades leads to generation of vasoactive products that result in angioedema.[21,22] Deficiencies of late components of complement, including C5 through C9, as well as factors B, D, and properdin create susceptibility to meningococcal infections.[23] Deficiency of factor H function is associated with membranoproliferative glomerulonephritis (Chapter 68), hemolytic–uremic syndrome, and age-related macular degeneration (AMD) (Chapter 74).[24] Deficiency of MBL is associated with increased susceptibility to bacterial infections in infancy and in individuals with other comorbid conditions, such as cystic fibrosis.[25]

CELLULAR INNATE IMMUNITY

Polymorphonuclear Leukocytes

PMNs are the most abundant leukocyte (Chapter 22). They have a short life span of ~6 hours in circulation, and in the healthy adult, $\sim 10^9$ PMNs are produced per hour. PMNs are readily identified by light microscopy by segmented nuclei divided into 3–5 lobules. Their cytoplasm contains four types of granules: azurophilic (or primary), specific (or secondary), gelatinase, and secretory. PMN granules contain a wide variety of APPs with a broad spectrum of antimicrobial activities (see Table 3.2). Azurophilic granules contain enzymes, such as proteinase 3, cathepsin G, and elastase as well as α-defensins and BPI. Specific granules contain lactoferrin and the proforms of cathelicidin peptides. Gelatinase granules are rich in gelatinase and are a marker of terminal neutrophil differentiation. Secretory granules contain a variety of receptors that are inserted into the cell membrane upon activation. Exocytosis of these receptors convert PMNs into cells responsive to inflammatory stimuli. PMNs are

KEY CONCEPTS

Cellular Innate Immunity

- Polymorphonuclear leukocytes (neutrophils) are the most abundant cells of the innate immune system, are short-lived, and are the earliest responders to infection.
- Monocytes and macrophages are the predominant immune cells several days after an infection.
- Activated neutrophils, monocytes, and macrophages kill phagocytosed bacteria through production of reactive oxygen intermediates and antimicrobial proteins and peptides (APPs).
- Dendritic cells (DCs) are efficient in uptake and presentation of foreign antigen and provide a critical link between innate and adaptive immunity.
- Natural killer (NK) cells can kill infected or malignant cells without prior activation.
- Mast cells are present are found at the interface between host and environment and are first responders to microbes and recruit other inflammatory cells.

the earliest responders to infection. Those not recruited to sites of infection undergo apoptosis and are cleared by the reticuloendothelial system. Individuals with severely low numbers of neutrophils (<500 cells/μL [microliter]) are susceptible to overwhelming bacterial infections.

CLINICAL PEARLS

Neutrophils

- Myeloperoxidase deficiency is asymptomatic in the majority of individuals, although *Candida* infections (mucocutaneous and invasive) have been reported.
- Impaired production of reactive oxygen intermediates causes chronic granulomatous disease, which manifests as susceptibility to invasive bacterial and fungal infections and impaired wound healing.
- Severe primary, or secondary, neutropenia (<500 cells/μL [microliter]) creates susceptibility to overwhelming bacterial infections.

Mononuclear phagocytes include monocytes and macrophages. Monocytes originate in bone marrow and migrate into the peripheral circulation. $CD14^+$ monocytes are effective at phagocytosis and production of reactive oxygen intermediates (ROIs) and proinflammatory cytokines in response to a wide variety of microbial stimuli. A subset of monocytes with low CD14 expression ($CD14^{dim}$), but expressing CD16, is associated with vascular endothelia and appears to be specialized for response to viruses and nucleic acid–containing immune complexes and may also be involved in the pathogenesis of autoimmune disorders.[26] $CD14^+$ monocytes enter tissues where they mature into macrophages. Distinct macrophages populations in different tissues are given specific names, including *Kupffer cells* in the liver, *alveolar macrophages* in the lung, *osteoclasts* in bone, and *microglia* within the brain. Macrophages differ from PMNs in that they are not terminally differentiated and can proliferate at sites of infection. They are longer lived than PMNs and are the predominant innate immune cell several days after an infection. Macrophages display plasticity in their functions, depending on the cytokine milieu. Classically activated macrophages (M1) are induced by interferon-γ (IFN-γ) and bacterial products. They have microbicidal activity and proinflammatory functions. In contrast, alternatively activated macrophages (M2) are induced by IL-4 and IL-13 and have anti-inflammatory functions. M2 macrophages have been shown to inhibit T-cell activation through production of IL-10 and transforming growth factor-β (TGF-β).[27]

Activated PMNs and macrophages kill phagocytosed bacteria by releasing microbicidal molecules both extracellularly and within phagolysosomes. Microbes are detected by pattern recognition receptors, as well as by Fc and complement C3 receptors.[28] Bacteria are internalized into phagosomes. Phagosomes fuse with lysosomes containing proteolytic enzymes (elastase, cathepsin G) to form phagolysosomes.

Activated PMNs and macrophages produce ROIs, which are toxic to microbes. ROIs are produced by phagocyte-derived nicotinamide adenine dinucleotide phosphate (NADPH) oxidase, a multisubunit enzyme that consists of five subunits, $p22^{phox}$, $p40^{phox}$, $p47^{phox}$, $p67^{phox}$, and $gp91^{phox}$. The phagocyte oxidase is activated following engulfment of opsonized bacteria (oxidative burst). Genetic defects in components of the NADPH oxidase complex create susceptibility to invasive infections with bacteria and fungi (chronic granulomatous disease), as well as impaired wound healing. A variety of stimuli activate the phagocyte oxidase complex, including the complement fragment C5a; formylated peptides, such as FMLP (N-formyl-methionine-leucine-phenylalanine); LTB4 (leukotriene B4); PAF (platelet-activating factor); and pattern recognition receptors, such as TLR4. Upon cellular activation, $p40^{phox}$, $p47^{phox}$, and $p67^{phox}$ are phosphorylated and recruited to cellular membranes, where they associate with membrane-bound $gp91^{phox}$ and $p22^{phox}$ (flavocytochrome b_{558}) and GTP-bound Rac1 (monocytes) or Rac2 (PMNs).[29] The activated enzyme generates superoxide radicals, which are then converted to hydrogen peroxide by superoxide dismutase. Hydrogen peroxide is combined with halide ions by myeloperoxidase to generate hypohalous acids, which are toxic to bacteria.

The phagocyte oxidase complex also generates an environment within the phagolysosome conducive to proteolytic enzyme activation. The oxidase functions as an electron pump that generates an electrochemical gradient across the phagolysosomal membranes, which is compensated by the movement of ions into the vacuole. This results in the increase in vacuolar pH and osmolarity required for activation of the anti-microbial proteases elastase and cathepsin G.[30]

Macrophages (Chapter 6) produce reactive nitrogen intermediates in response to microbes. Nitric oxide (NO) is produced by inducible nitric oxide synthetase (iNOS). Expression of iNOS is induced by activation of Toll-like receptors (TLRs), and expression is augmented further by IFN-γ.[31] iNOS catalyzes the conversion of arginine to citrulline, releasing diffusible nitric oxide gas. Within phagolysosomes, NO combines with hydrogen peroxide or superoxide to produce peroxynitrite radicals, which contribute to microbial killing. Although ROIs and NO are effective antimicrobial agents, they are nonspecific and are also capable of inducing damage to host tissues.

DCs (Chapter 6) have long membranous extensions for surveying the local environment and are highly phagocytic. They link innate to adaptive immune responses after activation by microbes. DCs express a variety of PRRs, which allow them to respond to microbes by antigen uptake and cytokine secretion. Activated DCs rapidly uptake antigen and then home to draining lymph nodes where they localize to T-cell zones. During their migration to lymph nodes, DCs mature and become efficient antigen presenting cells (APCs). Once in the lymph node, DCs express high levels of costimulatory molecules, such as B7 and IL-12p70, and present antigen to naïve T cells, inducing their differentiation into effector T cells (Th1 T cells). Plasmacytoid dendritic cells (pDCs) are specialized for response to viral infection and secrete large amounts of type 1 IFNs.

One subset of DC characterized by $CD11c^{high}CD103^+$ expression in the lamina propria of the small intestine facilitates the differentiation of regulatory T cells in a retinoic acid- and TGF-β–dependent manner. Such DC subsets may play a role in the development of tolerance to commensal bacteria (Chapter 14).

NK cells are derived from common lymphoid progenitor cells and constitute 5% to 20% of mononuclear cells in the periphery (Chapter 17). They do not express somatically rearranged antigen receptors. Target cells are identified using germline DNA-encoded receptors. NK cells are divided into two subsets, $CD56^{bright}CD16^-$ and $CD56^{dim}CD16^+$, which have different functions. $CD56^{dim}$ NK cells account for roughly 90% of NK cells in the periphery and express the low affinity Fcγ receptor (CD16), which mediates antibody-dependent, cell-mediated cytotoxicity. $CD56^{bright}$ NK cells are poorly cytotoxic but produce large amounts of cytokine and represent the majority of NK cells in peripheral lymphoid organs. NK cells are a major source of IFN-γ, which augments the microbicidal functions of macrophages. Conversely, NK cells are primed by IL-15 derived from DCs and IL-12 or IL-18 derived from macrophages, demonstrating the regulatory interactions that occur between NK cells and other cells of the immune system.

NK-cell function is regulated by a delicate balance between signals generated by inhibitory and activating receptors.[32] NK cells possess the ability to recognize and kill infected or malignantly transformed cells, while leaving healthy host cells unharmed. Inhibitory receptors on NK cells recognize class I major histocompatibility complex (MHC) molecules expressed on most healthy cells in the body.

NK inhibitory receptors include three families of receptors: heterodimers composed of CD94 and NKG2A, the Ig-like transcripts (*i.e.,* ILT-2), and the killer cell Ig-like receptor (KIR) family. The immunoreceptor tyrosine-based inhibition motifs (ITIMs) in their cytoplasmic tails recruit phosphatases (Src homology region 2 [SH2] domain-containing phosphatase-1 [SHP-1], SHP-2, and SHIP [SH2-containing inositol polyphosphate 5-phosphatase]), which oppose the effects of kinases activated by activating receptors. When NK cells encounter host cells expressing MHC class I molecules, protein tyrosine phosphatases are activated, reducing signaling downstream of activating receptors and opposing NK-cell activation (Fig. 3.2).

NK cells also possess activating receptors. CD16 mediates antibody-dependent, cell-mediated cytotoxicity and natural cytotoxicity receptors (*i.e.,* NKp46, NKp30, NKp44, NKG2D, CD94/NKG2C, 2B4). Activating receptors are linked to molecules (*i.e.,* CD3-ζ, FcR-γ, or DAP12) that contain immunoreceptor tyrosine-based activation motifs (ITAMs). Upon ligand binding, tyrosine residues within ITAMs are phosphorylated by Src family kinases. The tyrosine phosphorylated ITAMs serve as binding sites for the activation of other protein tyrosine kinases, such as Syk and ZAP-70, which activate downstream effector molecules in a signaling cascade. Infection of host cells with some viruses can lead to reduced MHC class I expression, thereby reducing viral antigen presentation to T cells. Concomitantly, ligands for activating receptors are expressed by the infected cell, leading to NK-cell activation and killing of the infected cell.

NK cells play an important role in immunosurveillance against tumors.[33] In humans, NK cell receptors that mediate tumor recognition include NKp46, NKp30, NKp44, DNAM-1 (DNAX accessory molecule-1), and NKG2D. Ligands expressed on target cells include MHC I-related chain (MIC)-A, MICB, unique long 16-binding proteins (ULBP), poliovirus receptor (PVR), and Nectin-2. DNAM-1 specific ligands include PVR and Nectin-2, which are expressed in cell lines that include carcinomas, melanomas, and neuroblastomas. Nectin expression is not specific to tumors, since nectins are expressed on normal cells. DNAM-1-nectin interactions on normal cells do not result in NK cell lysis because normal cells are protected by MHC class I expression. Conditions favoring NK cell–mediated lysis include tumors in which nectins are overexpressed and/or MHC class I expression is reduced, favoring NK-cell activation (see Fig. 3.2).

Natural killer T (NKT) cells are a small, but highly variable, population of thymus-derived T cells that express NK cell markers and a restricted repertoire of T-cell receptors (TCRs), which recognize lipids bound to the MHC-like molecule CD1d.[34] Type 1 NKT cells (also referred to as invariant NKT [iNKT] cells) express the invariant Vα24 and Jα28 TCR α chain, whereas type 2 NKT cells have a more diverse TCR repertoire. Mature human NKT cells can be further divided into three groups, including $CD4^{+}CD8^{-}$, $CD4^{-}CD8^{-}$, and $CD4^{-}CD8^{+}$ subsets. The most completely characterized NKT antigen is the lipid α-galactosylceremide (α-GalCer), which is often used to activate NKT cells experimentally. Identification of natural NKT ligands

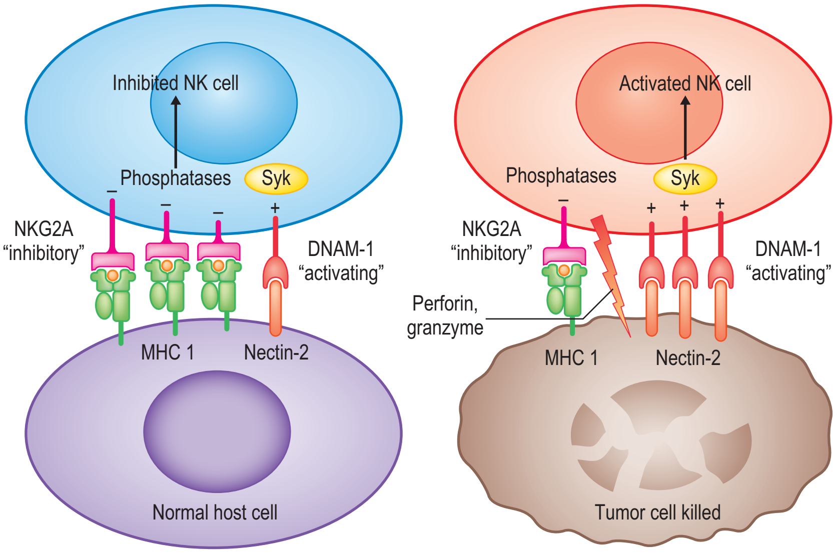


FIG 3.2 Regulation of Natural Killer (NK) Cell Function. Upon encountering normal host cells, inhibitory receptors on NK cells that contain immunoreceptor tyrosine-based inhibitory motifs (ITIMs) preferentially activate phosphatases (e.g., SHP-1/2, SHIP) that send inhibitory signals, inhibiting NK-cell function. NK cells that encounter virally infected cells or tumor cells receive signals through activating receptors that contain immunoreceptor tyrosine-based activation motifs (ITAMs) that activate tyrosine kinases, including Syk, leading to NK-cell activation, release of perforin and granzyme, and target cell death.

has proven difficult. NKT cells express perforin and granulysin and are capable of cytotoxic activity. NKT cells are also able to influence innate and adaptive immune responses through release of large amounts of cytokines, including IFN-γ, TNF-α, IL-4, IL-13, IL-10, and granulocyte macrophage–colony-stimulating factor (GM-CSF). In general, NKTs found in blood can produce large amounts of cytokines, whereas NKTs in the thymus are poor cytokine producers.

Decreased NKT cell frequency and/or function may increase susceptibility to some autoimmune diseases, including type 1 diabetes (Chapter 71) and multiple sclerosis (Chapter 66). Mice with NKT defects are susceptible to tumors and adoptive transfer of normal NKTs can provide protection against tumors (Chapter 77). NKT cells may also contribute to the pathogenesis of the airway hyperresponsiveness (AHR) in asthma (Chapter 72), which is dependent on IL-4 and IL-13 production in the airways. iNKT cells are necessary for AHR in several murine models of asthma, since NKT-deficient mice fail to develop AHR following allergen challenge, ozone challenge, or viral infection.[35] iNKT cell deficiency was associated with severe varicella infection, demonstrating a role for iNKT cells in innate antiviral immunity.[36]

Intraepithelial Lymphocytes, Innate Lymphoid Cells, B1 and MZ B Cells, and Mast Cells

Barrier epithelia of skin and the GI tract contain specific types of lymphocytes, including intraepithelial T lymphocytes (IELs) and B1 B cells (Chapter 7), which respond to commonly encountered microbes. Because of their more limited diversity of receptors, IELs can be considered part of the innate immune system. The main immune cell populations within the epidermal layer include keratinocytes, melanocytes, a type of DC known as the *Langerhans cell*, and IELs (Chapter 19).

Keratinocytes and melanocytes express a variety of PRRs, enabling detection of microbes, resulting in secretion of cytokines that can contribute to innate immune responses through recruitment and activation of phagocytes.[37] Langerhans cells form an elaborate network of dendritic processes that allow them to capture antigens that gain access to skin. Following activation by microbes, Langerhans cells migrate to draining lymph nodes and express chemokine receptor-7 (CCR7), which allows them to migrate to the T-cell zones within lymph nodes in response to the chemokine ligands (CCL)-19 and CCL-21 and present antigen to T cells.[38]

Intraepidermal T lymphocytes constitute roughly 2% of lymphocytes within the skin. This lymphocyte subset expresses a more restricted set of antigen receptors, which include both αβ and γδ TCRs, similar to IELs found in the intestines. These specialized T cells appear to be committed to recognizing microbial peptide antigens commonly found at epithelial surfaces and thus function as components of the innate immune system.[39]

IELs are a significant component of the GI immune system and reside at the basolateral side of the intestinal epithelial cell layer (Chapter 2). IEL are among the first immune cells to encounter pathogens that have breached intestinal epithelia. IEL consist of CD8 T cells, as well as memory-effector T cells bearing αβ or γδ TCRs. IELs contain a greater proportion of TCRγδ$^+$ cells than is found in the peripheral circulation.[40]

CD4$^+$TCRαβ$^+$ and CD8$^+$αβTCRαβ$^+$ IELs are class II and class I restricted, respectively. These IELs have likely undergone thymic selection and subsequently homed to the gut after antigenic stimulation. As such, these IELs are likely specific for foreign antigens.[40] They have a memory phenotype and an oligoclonal TCR repertoire. IELs in the small intestine frequently express of CD8$^+$αα (CD4$^+$CD8$^+$αα or CD8$^+$αβCD8$^+$αα), a characteristic of activated mucosal T cells within the gut microenvironment. Upon antigenic stimulation, CD8$^+$αβTCRαβ$^+$ IELs are cytolytic and kill via granzymes and perforin or through engagement of Fas.[40]

TCRγδ$^+$ IELs emigrate from the thymus and subsequently take up residence in the intestinal epithelium. They constitute roughly 10% of intestinal IELs in humans, and the majority expresses CD8αα. TCRγδ$^+$ IELs recognize nonclassic MHC molecules, such as thymus leukemia antigen or MHC class I-like molecules, MICA (MHC I-related chain [MIC]-A) and MICB, and may help modulate inflammatory immune responses. These IELs can be cytolytic and express FasL. TCRγδ+ IELs can produce keratinocyte growth factor and promote intestinal epithelial integrity.[40]

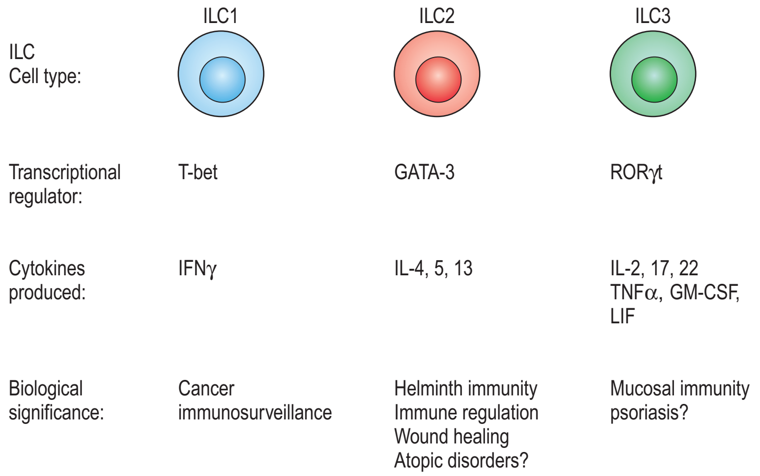
The B1 and marginal zone (MZ) subsets of B lymphocytes have been characterized as innate-like B cells. They express antigen receptors enriched for germline sequence. These cell types have been mostly studied in mice. Their identity in humans is less clear. B1 cells and MZ B cells can function as APCs, but unlike conventional B cells, B1 cells and MZ B cells do not develop into memory B cells. B1 cells and MZ B cells share characteristics: *(i)* they are the main source of natural antibodies; (ii) they express high surface levels of IgM and low surface levels of IgD; and *(iii)* they are rapidly activated by microbes through pattern recognition receptors to produce large amounts of natural antibodies.[41] B1 cells and MZ B cells produce IL-10 upon activation, which may downregulate immune responses. The natural antibodies produced by B1 cells and MZ B cells function as the first line of defense against invading microbes.

Innate lymphoid cells (ILCs) are a heterogeneous population of cells that have been recently recognized. ILCs do not express rearranged antigen-specific receptors. This lymphoid subset includes killer ILCs (*e.g.,* NK cells) and helper ILCs. Helper ILCs are further classified as ILC1, ILC2, and ILC3.

ILC1 cells express T-bet, similar to NK cells, and produce IFN-γ but lack cytolytic activity. They are found mainly within tissues and are barely detectable in peripheral blood. Oncogene-induced murine cancer model studies suggest that ILC1 cells may play a role in cancer immunosurveillance.[42]

ILC2 development is dependent on expression of the transcription factor GATA-3 and produces the cytokines IL-5 and IL-13. ILC2 cells were first identified in mice as a source of T helper (Th)2 cytokines (IL-4, -5, -13). ILC2 cells may play a role in antihelmintic immunity, immune surveillance, immune regulation, and wound healing. ILC2 cells have been shown to accumulate in the skin of patients with AD and within nasal polyps of patients with chronic rhinosinusitis. ILC2 cells produce IL-4, -5, and -13 in response to epithelial-derived IL-33, IL-25, and thymic stromal lymphopoietin (TSLP). The production of Th2 cytokines by ILC2 cells may represent an early step in the development of atopic disorders.[43]

ILC3 cells express the retinoic acid receptor related orphan receptor γt (RORγt) and produce IL-17 and IL-22 (Fig. 3.3).[44] ILC3 cells include fetal lymphoid tissue inducer (LTi) cells, which drive secondary lymphoid organ development during embryogenesis. LTi cells can induce upregulation of adhesion molecules on stromal cells and the release of chemokines involved in the recruitment of T and B lymphocytes and DCs to lymph nodes, leading to subsequent differentiation of naïve T cells into effector T cells and B-cell activation and production of antibody-secreting



	ILC1	ILC2	ILC3
ILC Cell type:			
Transcriptional regulator:	T-bet	GATA-3	RORγt
Cytokines produced:	IFNγ	IL-4, 5, 13	IL-2, 17, 22 TNFα, GM-CSF, LIF
Biological significance:	Cancer immunosurveillance	Helminth immunity Immune regulation Wound healing Atopic disorders?	Mucosal immunity psoriasis?

FIG 3.3 Developmental Regulation and Functions of Innate Lymphoid Cells. The development of innate lymphoid cells (ILCs) is regulated by master transcriptional regulators, including T-bet, GATA-3, and RORγt. Three subsets are recognized: ILC1, ILC2, and ILC3. ILC play roles in tumor immunosurveillance, immune regulation, wound healing, mucosal immunity, atopic disorders, psoriasis, and mucosal immunity.

cells. Postnatal ILC3 cells influence tissue homeostasis and host defense against extracellular organisms. ILC3-produced IL-22 induces expression of APPs by intestinal epithelial cells. Production of IL-17 by ILC3 cells likely contributes to host defense against *Candida*. ILC3 cells can also produce IL-2, GM-CSF, TNF-α, and leukemia inhibitory factor (LIF). Production of GM-CSF by splenic ILC3 cells is believed to promote survival and activation of splenic neutrophils. A significant population of ILC3 cells has been found in lesional skin of patients with psoriasis.[44]

Mast cells (Chapter 23) are components of innate immunity that are also commonly found at the interface between host and environment. They are derived from progenitors in bone marrow and circulate as immature precursors to the periphery. Mast cells take up residence and mature in skin, airways, and the GI tract. They are positioned to be first responders to environmental stimuli, including infectious agents.[45] Stem cell factor (SCF, also known as *c-kit ligand*) is their main survival and developmental factor. Mast cells express TLR-1 through TLR-9 and therefore are capable of responding to a wide variety of pathogens. TLR-induced mast-cell activation leads to production of proinflammatory cytokines and chemokines. Murine models of peritonitis, such as cecal ligation and puncture, demonstrate that mast cells enhance resistance to bacterial infection.[46] Mast cells are also well known for mediating allergic reactions through IgE-bound allergens that are anchored by FcεRI on the mast-cell surface. Ligation of FcεRI leads to release of tryptase, histamine, leukotrienes, prostaglandins, and cytokines, which cause type 1 hypersensitivity reactions (Chapter 42).

ACTIVATING INNATE IMMUNITY

The innate immune response is initiated when cells of the innate immune system encounter pathogens and recognize them by means of PRRs binding to microbial molecules (*e.g.,* lipopolysaccharide, DNA, RNA). These interactions activate signaling pathways that lead to the production of secreted factors involved in the inflammatory response, including cytokines (Chapter 9) and chemokines (Chapter 10). Characteristics of cytokines include pleiotropism (*e.g.,* the ability to act on multiple cell types) and redundancy. Cytokines can function locally and distantly and can affect the production of other cytokines. Exposure to cytokines can induce changes in gene expression that affect cell function (*e.g.,* enhanced microbicidal activity or proliferation). Secretion of cytokines (IL-1β, IL-6, TNF-α) is a transient event, thereby limiting potential destruction of host tissue. However, severe infections (*e.g.,* bacteremia and sepsis) can lead to overproduction of TNF-α, IL-1β, and IFN-γ, which leads to vascular collapse, disseminated intravascular coagulation, and metabolic disturbances (septic shock), which are often fatal.

Cytokine synthesis is a transient process because the messenger RNA of most cytokines is unstable, thus limiting cytokine production. Production of certain cytokines is also regulated by a posttranslational process. For example, TNF-α is a membrane-bound protein that is proteolytically cleaved by a membrane-associated metalloproteinase. IL-1β is a 33-kDa protein that is proteolytically processed by the IL-1β-converting enzyme caspase-1 to generate the biologically active, 17-kDa, mature IL-1β (described below).

TNF-α and IL-1β can recruit PMNs and monocytes to sites of infection and enhance their ability to eliminate microbes. TNF-α and IL-1β induce expression of adhesion molecules (Chapter 11), such as selectins (P-selectin, E-selectin) and the integrin ligands ICAMs (intercellular adhesion molecules) and VCAMs (vascular cell adhesion molecules) on vascular endothelial cells near the sites of infection. Expression of selectins on vascular endothelium induces leukocyte rolling on endothelium. Chemokines, such as CXCL8, activate PMNs and monocyte integrins and increase their affinity for ligands (ICAMs, VCAMs) on vascular endothelium, allowing migration of PMNs and monocytes through endothelium to sites of infection. TNF-α and IL-1β both induce prostaglandin synthesis in the hypothalamus, which induces fever.

PATTERN RECOGNITION RECEPTORS

Our understanding of the mechanisms by which pathogens are detected has increased greatly over the past decade. Pathogen

KEY CONCEPTS

Pattern Recognition Receptors (PRRs)

- Toll-like receptors consist of an extracellular domain containing leucine-rich repeats (LRRs) for ligand binding and a cytoplasmic Toll/IL-1 receptor domain that links to adapter proteins and complex signaling pathways.
- Nucleotide oligomerization domain (NOD)-like receptors are a family of 22 proteins that contain LRRs for potential ligand binding, a NOD, and a caspase activation and recruitment domain (CARD), Pyrin domain, or a baculovirus inhibitor of apoptosis repeat (BIR) domain for initiation of signaling.
- Retinoic acid-inducible gene (RIG)-like receptors consist of two N-terminal CARD for signaling and an RNA helicase domain.
- C-type lectin receptors (CLRs) contain a C (Ca^{2+})-type recognition domain and mediate diverse functions, depending on the signaling pathways they activate.
- Scavenger receptors are a diverse group of receptors that recognize a variety of ligands and mediate the uptake of oxidized lipoproteins and may be involved in atherosclerotic plaque formation.

recognition by the innate immune system leads to engulfment and destruction of invading pathogens, but clearance is often incomplete. The subsequent adaptive immune response is required to complete clearance.

The innate immune system expresses a wide variety of PRRs that mediate pathogen recognition. These include TLRs, nucleotide oligomerization domain (NOD)-like receptors (NLRs), and the retinoic acid-inducible gene-I (RIG-I)-like receptors (RLRs).[47] These receptors play an essential role in initiating the innate immune response. Unlike T-cell and B-cell antigen receptors, PRRs are germline encoded, do not undergo somatic recombination, and are expressed constitutively by immune and nonimmune cells. PRRs recognize PAMPs, components of pathogens that are invariant and required for pathogen survival (Table 3.3). Although PRRs detect the PAMPs expressed by microbes, they may also recognize self-molecules (*i.e.,* host nucleic acids), which may underlie some autoimmune diseases, such as lupus and rheumatoid arthritis.

TABLE 3.3 Classes of Pattern Recognition Receptors (PRRs)

Pattern Recognition Receptor		Ligand
Toll-like receptors (TLRs)	TLR1/2	Triacyl lipopeptides
	TLR2	Zymosan
	TLR3	dsRNA
	TLR4	LPS, RSV glycoprotein, HSPs, pneumolysin
	TLR2/6	Diacyl lipopeptide
	TLR7	ssRNA
	TLR8	ssRNA
	TLR9	dsDNA, hemozoin
	TLR10	?
	TLR11	Profilin-like protein
NOD-like receptors (NLRs)	NOD1	DAP, MDP
	NOD2	MDP
	CIITA	?
	NAIP	*Legionella pneumophilia,* flagellin?
	IPAF	PAMPs
	NLRP1	PAMPs, MDP, microbial toxins
	NLRP2	?
	NLRP3	PAMPs, toxins, DAMPs
	NLRP4–14	?
RIG-like receptors (RLRs)	RIG-I	dsRNA, ssRNA
	MDA5	dsRNA, ssRNA
C-type lectin receptors (CLRs)	Mannose receptor	Bacterial carbohydrates
	Dectin-1	Fungal wall glucans
Scavenger receptors	SRA, SRB	Oxidized lipoproteins, apoptotic cells
	CD36, CD68,	β-amyloid

dsRNA, double-stranded RNA; CIITA, class II major histocompatibility complex transactivator; DAMP, danger-associated molecular pattern; DAP, meso-diaminopimelic acid; MDP, muramyl dipeptide; IPAF, IL-1beta converting enzyme protease activating factor; LPS, lipopolysaccharide; MDA-5, melanoma differentiation-associated gene-5; PAMP, pathogen-associated molecular pattern; NAIP, neuronal apoptosis inhibitory protein; NLRP, NOD-like receptor related protein; RSV, respiratory syncytial virus; SR, scavenger receptor; ssRNA, single stranded RNA.

CLINICAL PEARLS

Toll-Like Receptors (TLRs), Caspase Activation and Recruitment Domain (CARD), NLR, and Inflammasomes

- TLRs play a nonredundant role in host defense. Impaired TLR function (*i.e.*, IL-1 receptor associated kinase-4 [IRAK-4], myeloid differentiation factor 88 [MyD88] deficiency) results in susceptibility to invasive, pyogenic infections.
- Increased infection in newborns and infants with TLR pathway defects (MyD88, IRAK-4, etc.) suggest this system is particularly important in early life.
- Missense mutations in nucleotide oligomerization domain 2 (NOD2) have been associated with Crohn disease and Blau syndrome.
- Mutations in NLRP3 are associated with Muckle Wells syndrome, familial cold autoinflammatory syndrome, and neonatal-onset multisystem inflammatory disease (NOMID).
- Mutation in CARD9 leads to susceptibility to chronic mucocutaneous candidiasis.

Toll-Like Receptors

Toll was initially identified in *Drosophila melanogaster* as a receptor required for dorsal–ventral patterning. Subsequently, the Toll pathway was found to be an essential component of host defense against fungal infection, which led to cloning of mammalian homologues, the TLRs. Mammalian TLRs consist of 11 members that can recognize a wide variety of PAMPs. TLRs are type 1 integral membrane glycoproteins characterized by an extracellular domain with varying numbers of leucine-rich repeats (LRRs) and a cytoplasmic signaling domain homologous to the IL-1 receptor (IL-1R), referred to as the Toll/IL-1R (TIR) homology domain. The TIR domain links the receptor to adaptor proteins (*e.g.,* myeloid differentiation factor 88 [MyD88]) and downstream signaling molecules. This leads to transcription of genes that regulate inflammation (Fig. 3.4).

TLRs are widely expressed on or within cells of the immune system and the epithelia. TLRs detect a wide variety of pathogens (see Table 3.3). They are classified into subfamilies based on their genetic tree. The TLR1, TLR2, and TLR6 subfamily recognizes bacterial lipoproteins, whereas the TLR3, TLR7, TLR8, and TLR9 subfamily recognizes nucleic acids. TLR4, in conjunction with MD-2, recognizes lipopolysaccharide (LPS), TLR5 binds bacterial flagellin, and TLR11, which is functional in mice but probably not in humans, recognizes a profilin-like molecule of *Toxoplasma gondii.* However, ligand binding by TLRs can be

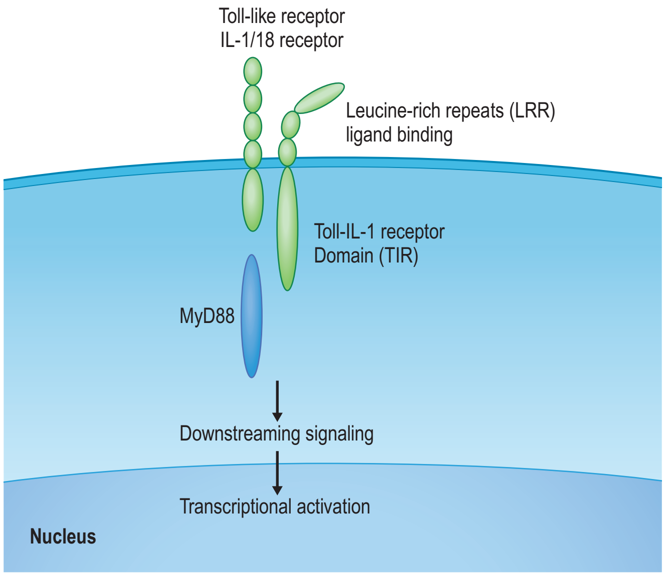


FIG 3.4 Toll-Like Receptors (TLRs) and Interleukin (IL)-1/-18 Receptors Share a Common Signaling Pathway. Upon ligand binding, signals are transduced intracellularly by the interaction of the adaptor protein, MyD88, with the TIR domain of receptor. MyD88 interacts with IL-1 receptor associated kinase (IRAK)-4 through death domain interactions, activating a signaling cascade, which results in transcriptional activation of genes involved in inflammation.

promiscuous. For example, TLR4 can also bind respiratory syncytial virus F protein and pneumolysin of *S. pneumoniae*[48,49]; and TLR9 binds malarial hemozoin and hypomethylated CpG-rich DNA. TLRs can also recognize endogenous danger signals, for example, damage-associated molecular patterns (DAMPs) that include heat shock proteins.[50]

The cellular localization of TLRs varies. TLR1, -2, -4, -5, -6, -10, and -11 are found on cell surfaces, whereas TLR3, -7, -8, and -9 are located within endosomes. The cell surface expression of TLRs, such as TLR4, which recognizes LPS, allows recognition of extracellular molecules released from pathogens. Endosomal expression of TLR3, -7, -8, and -9 allows recognition of microbial nucleic acids following their uptake and degradation in phagolysosomes. Endosomal expression of TLR3, -7, -8, and -9 may prevent activation by host nucleic acids and the development of autoimmunity. The broad cellular expression of TLRs and their diverse and promiscuous agonist recognition allows detection of a wide variety of pathogens despite the existence of a limited number of TLRs.

TLR-mediated cellular responses are essential to host defense. TLR activation stimulates a brief burst of macropinocytosis, which results in antigen uptake at sites of infection and allows antigen presentation to T cells. TLR activation leads to production of proinflammatory cytokines (*e.g.,* TNF-α and IL-6) and chemokines (*e.g.,* CXCL8). TLR pathway engagement induces transcription and translation of messenger RNA (mRNA) encoding pro-IL-1β, but production of mature IL-1β requires activation of the inflammasome (described below). Production of proinflammatory cytokines recruits phagocytes to sites of infection and augments their antimicrobial functions. Production of IL-12p70 leads to activation of naïve T cells and their subsequent differentiation into effector Th1 cells. Presentation of foreign peptides and increased expression of MHC molecules along with expression of costimulatory molecules (*e.g.,* B7–1, B7–2, and IL-12p70) results in subsequent development of adaptive immune responses. IL-12p70 stimulates IFN-γ production by T cells, which further augments the microbicidal activities of phagocytes. Stimulation of TLR3, -7, -8, and -9 elicits the production of proinflammatory cytokines, as well as type 1 IFNs, which play a crucial role in innate antiviral immunity and also influence adaptive immune responses.

Engagement of TLRs activates complex signaling pathways that have been characterized through biochemical analyses and in gene-targeted mice.[51-53] TLR, IL-1R, and IL-18R share similar signaling pathways (Fig. 3.5). Upon ligand binding, the cytoplasmic adaptor protein MyD88 is recruited to the TIR domain of the receptor for all TLRs, except TLR3. Recruitment of MyD88 leads to recruitment of IL-1 receptor associated kinase-4 (IRAK-4), likely through death domain interactions. IRAK-4 activation leads to recruitment and activation of IRAK-1 and IRAK-2. In monocytes, IRAK-M is also recruited to this complex and functions as a negative regulator of signaling. Both IRAK-1 and IRAK-2 activation are required for full activation of NFκB and MAP kinases. IRAK activation leads to interaction with TNF receptor–associated factor 6 (TRAF6), which is an E3 ubiquitin ligase. Along with the E2 conjugating complex of Ubc13 and Uev1a, TRAF6 is K-63 ubiquitinated, recruiting TGF-β–activated protein kinase-1 (TAK-1). TAK-1 then activates the inhibitor of NFκB (IκB) kinase complex (IKK), which consists of NFκB essential modifier (NEMO), IKKα, and IKKβ, leading to phosphorylation of IκB (inhibitor of NFκB) proteins and their subsequent K-48 linked ubiquitination and degradation. NFκB is subsequently released from inhibition, allowing translocation to the nucleus, where it mediates transcriptional activation of numerous genes involved in inflammation. The transcription factor interferon regulatory factor-5 (IRF-5) is also activated downstream of TRAF6 and is required for production of proinflammatory cytokines. TAK-1 activation also leads to activation of p38 MAP kinase and c-Jun N terminal kinase (JNK), which then activates the AP1 transcriptional complex (see Fig. 3.5).

TLR signaling can proceed via multiple pathways, impacting both the kinetics and nature of the subsequent innate response. For example, TLR4 also interacts with the adaptors MAL (MyD88-like adaptor protein), TRAM (translocating chain-associating membrane protein), and TRIF (TIR domain–containing adaptor-inducing IFN-β) (see Fig. 3.5). TLR4 initially recruits MAL and MyD88 to trigger "early phase" NFκB and MAP kinase activation. TLR4 is subsequently endocytosed and trafficked to the endosome, where it forms a signaling complex with TRAM and TRIF, which leads to activation of TANK-binding kinase-1 (TBK-1), IKKε, and IRF-3, and "late phase" activation of NFκB and MAP kinases. Activation of IRF3 induces IFN-β production.

Antiviral TLRs are located in endosomes and interact with an endoplasmic reticulum membrane protein called UNC93B (see Fig. 3.5).[54] Upon activation, TLR3 does not recruit MyD88, but rather TRIF, leading to recruitment of TRAF3 and activation of TBK1, IKKι, IRF-3, and IFN-β production. TRIF also recruits RIP1 and TRAF6, which leads to activation of NFκB. While the other antiviral TLRs, TLR7, -8, and TLR9, are MyD88 dependent; they activate a pathway utilizing IRAK-1, IKKα, TRAF3, and intracellular osteopontin (iOPN), which activates IRF7, leading to production of IFN-α.[55] TLR7, -8, and -9 also utilize a TRAF6-dependent pathway that leads to NFκB/MAP kinase activation.

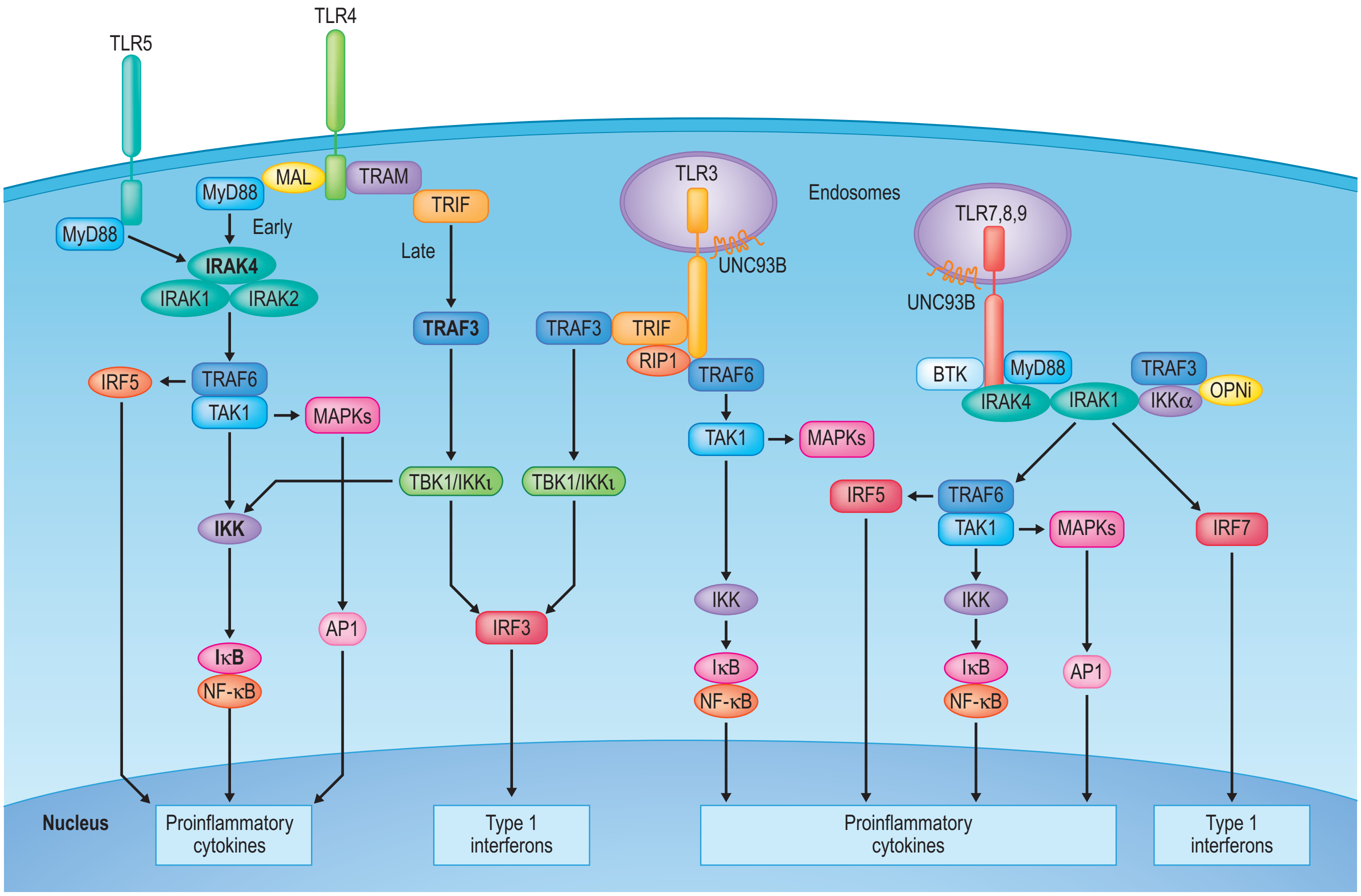

TLR5
TLR4
MAL
TRAM
MyD88
TRIF
MyD88
Early
Late
IRAK4
IRAK1
IRAK2
TRAF3
TRAF3
TRIF
RIP1
IRF5
TRAF6
TAK1
MAPKs
TBK1/IKKι
TBK1/IKKι
IKK
AP1
IRF3
IκB
NF-κB
TLR3
UNC93B
Endosomes
TLR7,8,9
UNC93B
TRAF6
TAK1
MAPKs
IKK
IκB
NF-κB
BTK
MyD88
TRAF3
IRAK4
IRAK1
IKKα
OPNi
IRF5
TRAF6
TAK1
MAPKs
IRF7
IKK
IκB
NF-κB
AP1
Nucleus
Proinflammatory cytokines
Type 1 interferons
Proinflammatory cytokines
Type 1 interferons

FIG 3.5 **MyD88-Dependent and -Independent Toll-Like Receptor (TLR) Signaling Pathways.** Engagement of MyD88-dependent TLR (TLR5) results in activation of IL-1 receptor associated kinase (IRAK)-4 and IRAK-1, IRAK-2, leading to activation of TRAF6 and TAK-1. Subsequently, activation of the IκB kinase (IKK) complex and MAP kinases activates NF-κB and AP1 transcription factors, respectively. The transcription factor IRF5 is also activated downstream of TNF receptor–associated factor 6 (TRAF6). The TLR4 signaling pathway utilizes four adaptor proteins. The adaptors MAL and MyD88 are activated upon ligand interaction at the cell surface, leading activation of an "early" signaling cascade through the IRAKs. Subsequently, TLR4 is internalized and a "late" signaling cascade, which is dependent on the adaptors TRAM and TRIF, is activated. TLR3 activates a TRIF-dependent pathway that activates RIP1 and TBK1/IKKι, leading to production of proinflammatory cytokines and IFN-β. TLR7, -8, and -9 are MyD88 dependent and activate the transcription factors NF-κB, IRF5, AP1, and IRF7, resulting in production of proinflammatory cytokines and type 1 interferons.

Bruton tyrosine kinase (BTK) (Chapter 34) also plays a critical role in TLR8- and TLR9-induced production of TNF-α and IL-6.[56]

The innate immune system *in natura* encounters intact pathogens that express multiple PAMPs, including bacterial cell wall components as well as microbial DNA and RNA. Thus DCs and other phagocytes are activated through multiple PRRs. Activation of DCs through combinations of TLRs, such as TLR4 and TLR8, can induce synergistic production of Th1 cell–inducing cytokines, as well as the Th1-inducing ligand, Delta-4, leading to stronger Th1 differentiation of T cells than occurs following activation of single TLR.[57] Interestingly, the use of combinations of TLR agonists on virus-sized nanoparticles containing antigen induce enhanced and better-sustained antibody responses in mice and nonhuman primates.[58] Thus the use of combinations of TLR agonists as adjuvants in vaccines may result in enhanced efficacy of future vaccines (Chapter 90), as well as in immunotherapy against tumors (Chapter 77).

During an infection, multiple factors can mitigate TLR-induced inflammation, and these include adenosine, an endogenous purine metabolite whose levels rise during stress or hypoxia. Adenosine binds receptors expressed on leukocytes, leading to increased intracellular concentrations of cyclic adenosine monophosphate (cAMP), dampening TLR-mediated production of Th1-polarizing cytokines while preserving production of Th2 and antiinflammatory cytokines. Antiinflammatory/proresolving lipid metabolites, such as resolvins and lipoxins, can differentially regulate TLR4-mediated responses of macrophages, inhibiting TNF response to pure LPS but enhancing uptake, killing, and TNF-α production to whole gram-negative bacteria.[59]

Within the GI tract, the factors that maintain tolerance to commensal host flora while detecting/containing pathogenic

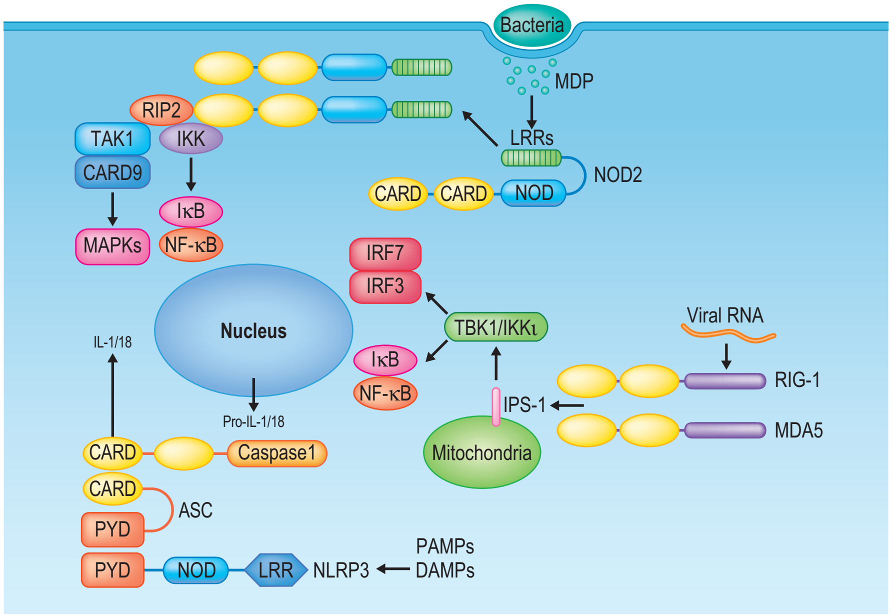


FIG 3.6 NOD-Like Receptor (NLR), RIG-I–Like Receptor (RLR), and Inflammasome Signaling. NOD2 is activated following exposure to bacterial muramyl dipeptide (MDP) and leads to dimerization and activation of RIP2, TAK1, CARD9, and the IκB kinase (IKK) complex. This results in inflammatory gene transcription. RLRs, like RIG-I and MDA-5, are activated by double-stranded RNA generated by intracellular, replicating viruses, inducing activation of transcription factors, including NF-κB, IRF3, and IRF7, leading to production of proinflammatory cytokines and type 1 interferons. Inflammasomes can be activated by microbial products (PAMPs), as well as endogenous products released by damaged host cells (DAMPs), leading to processing of pro-IL-1β to active IL-1β.

bacteria with appropriate inflammatory responses are incompletely understood. The detection of common PAMP in pathogenic and nonpathogenic bacteria would be anticipated to activate the same inflammatory response. Nevertheless, the detection of commensal bacteria within the intestines can induce tolerance. TLR signaling can contribute to intestinal homeostasis by regulating intestinal epithelial cell proliferation and epithelial integrity. Expression and localization of TLRs in the intestinal epithelium may directly relate to their role in maintaining homeostasis versus inducing inflammation. For example, within the intestinal epithelium, TLR9 activation through the apical membrane induces tolerance, whereas TLR9 activation via the basolateral membrane induces an inflammatory response through the canonical NF-κB pathway.[60] Differential spatial expression of PRR in epithelia may constitute a critical mechanism of distinguishing nonpathological from pathological bacteria.

NOD-Like Receptors

NLRs are cytoplasmic PRRs that are unrelated to the transmembrane PRRs. NLRs mediate detection of intracytoplasmic bacterial products. Among the NLRs are five members of the NOD family, 14 members of the NALP family, CIITA, IPAF, and NAIP (see Table 3.3). NLR family proteins possess LRRs for ligand detection; a NOD domain (also referred to as a NACHT domain); a domain for initiation of signaling, such as caspase activation and recruitment domain (CARD); pyrin domains; or baculovirus inhibitor of apoptosis repeat (BIR) domains. NOD1 and NOD2 were the first NLRs identified. They detect components of bacterial peptidoglycan: NOD1 detects mesodiaminopimelic acid, and NOD2 detects muramyl dipeptide. Direct ligand binding by NLR has not yet been demonstrated, leaving open the possibility that detection of pathogens and other signals by NLRs may be indirect. Following activation, NODs oligomerize and recruit the protein kinase RIP2 and CARD9 via CARD domains, leading to activation of NF-κB and MAP kinases, respectively (Fig. 3.6). NOD2 may also play a role in activation of some inflammasomes (described below). In humans, missense mutations in NOD2 that impair function have been associated with susceptibility to Crohn disease. Conversely, missense mutations in NOD2 that lead to constitutive activation of NF-κB lead to Blau syndrome, an autosomal dominant disorder characterized by granulomatous arthritis, iritis, and skin granulomas.[61]

RIG-I–Like Receptors

RLRs detect the presence of viral nucleic acids generated by intracellular, replicating viruses. The RLRs consist of two receptors: RIG-I and melanoma differentiation-associated gene-5 (MDA-5). Both have two N-terminal CARDs and an RNA helicase domain and mediate virus-induced type 1 IFN expression in fibroblasts and conventional DCs. A third RLR, laboratory of genetics and physiology 2 (LGP2), lacks the N-terminal CARD domains and plays a role in repression of signaling. RIG-I and MDA-5 are activated by double-stranded RNA (dsRNA) generated during viral replication with distinct specificities for viral recognition. RIG-I detects Orthomyxoviridae, Rhabdoviridae, Paramyxoviridae, and Flaviviridae, whereas MDA-5 detects Picornaviridae, Caliciviridae, and Coronaviridae. Poly inosine:cytosine (poly I:C) is a nonspecific dsRNA analogue used experimentally to activate TLR3 and RIG-I/MDA-5. Relatively short poly I:C (<1 kb) is recognized preferentially by RIG-I, whereas long poly I:C (>1 kb) is preferentially recognized by MDA-5.[48]

dsRNA-induced activation of RIG-I and MDA-5 induces their association with a mitochondria-associated adaptor known as interferon-β promoter stimulator-1 (IPS-1) or mitochondrial antiviral signaling protein (MAVS) through CARD domain interactions. Downstream effectors include TBK-1 and IKKι, which activate IRF3 and IRF7, leading to production of type-1 interferons (see Fig. 3.6). Of note, live bacteria are more effective inducers of STAT-1, type I IFN, and inflammasome pathways compared with killed organisms, a property that may reflect the importance of bacterial RNA to innate recognition of live infection.[62]

C-Type Lectin Receptors

C-type lectin receptors (CLRs) are a diverse group of receptors originally identified as Ca^{2+}-dependent carbohydrate-binding proteins.[63] CLRs are defined as any protein that contains a C-type carbohydrate recognition domain (CRD), regardless of calcium- or carbohydrate-binding ability. CLRs include numerous members with diverse functions, including cell adhesion, regulation of NK-cell function, phagocytosis, endocytosis, platelet activation, complement activation, tissue remodeling, and innate immunity. In myeloid cells, CLRs can mediate internalization of microbes to allow for antigen processing and presentation. Some CLRs function analogously to TLR, resulting in direct cellular activation and generation of inflammatory responses. Other CLRs are capable of binding PAMPs, but function to modulate cell activation. The functions of myeloid CLRs are dictated by the signaling pathways they activate.

Dectin-1 is a CLR expressed on DCs and other myeloid cells that recognizes β-1, 3-linked glucans present in the cell wall of fungi, mycobacteria, and plants. Dectin-2 recognizes high mannose structures and α-mannans found in fungi, mycobacteria, and dust mites. Following ligand binding, Dectin-1 and Dectin-2 activate signaling pathways by utilizing the tyrosine kinase Syk, CARD9, and Raf-1, leading to activation of the transcription factors NF-κB, NFAT, and AP-1 and production of proinflammatory cytokines (Fig. 3.7). The production of cytokines (IL-1, IL-6, TGF-β, IL-23) downstream of Dectin-1 and Dectin-2 induces the subsequent differentiation of naïve T cells into Th17 T cells, which play a critical role in antifungal immunity. Activation of Syk induces generation of ROIs, which can activate the NLRP3 inflammasome, leading to processing of pro-IL-1β to mature IL-1β. The importance of CLR function in antifungal immunity is demonstrated by inactivating mutations in Dectin-1 and CARD9 that lead to chronic mucocutaneous candidiasis, as well as invasive fungal infections in the case of CARD9-deficiency.[64]

Mincle (macrophage inducible C-type lectin) recognizes α-mannans and glycolipids and associates with the FcRγ chain. Upon ligand binding, Syk is recruited to the ITAM of FcRγ, leading to cellular activation. Mincle also binds the endogenous nucleoprotein SAP130, which is exposed by dead cells. The Mincle-mediated response to dead cells leads to infiltration of PMNs into damaged tissue and may contribute to tissue repair.

Other CLRs, such as DCIR (DC-inhibitory receptor), possess an inhibitory ITIM motif. DCIR is expressed on myeloid cells, DCs, and B cells. DCIR inhibits TLR8-induced IL-12p70 and TNF-α production by myeloid DCs and TLR9-induced IFN-α/β production by pDCs. Inhibition of TLR responses may reflect inhibition of tyrosine kinases and/or PI3 kinase pathways.

Pathogens express multiple PAMPs that activate a variety of PRRs. Indeed, CLRs and TLRs cooperate in antimicrobial responses. Coordinate activation of Dectin-1 and TLR2, for example via stimulation with zymosan, enhances production of proinflammatory cytokines, via activation of both Dectin-1-Syk and TLR2-MyD88 signaling pathways (see Fig. 3.7). DC-SIGN, which recognizes mycobacteria and viruses, can enhance TLR-induced NF-κB activation through a Raf-1–dependent signaling pathway.[65]

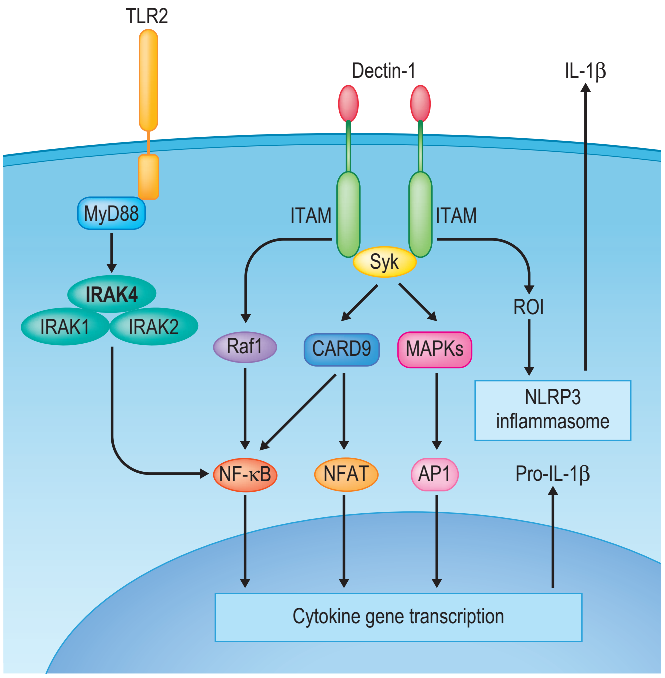


FIG 3.7 **C-Type Lectin Signaling.** C-type lectin receptors (CLRs), like Dectin-1, contain immunoreceptor tyrosine-based activation motifs (ITAMs) that interact with the cytosolic tyrosine kinase, Syk, leading to activation of transcription factors including NF-κB, NFAT, and AP1. Dectin-1 activates the serine kinase, Raf1, which contributes to NF-κB activation. Toll-like receptors (TLRs), such as TLR2, can cooperate with Dectin-1 signaling to activate NF-κB, leading to an enhanced inflammatory response.

Scavenger Receptors

Scavenger receptors are a diverse group of receptors that include CD36, CD68, SR class A, and SR class B.[66] The receptors mediate the uptake of oxidized lipoproteins into cells. Scavenger receptors also mediate the uptake of microbes and contribute to the response of macrophages to mycobacteria. SR class A can also mediate an inflammatory response to β-amyloid fibrils that may contribute to Alzheimer disease (see Table 3.3). Scavenger receptors play a pathological role in the generation of cholesterol-laden foam cells that comprise atherosclerotic plaques in blood vessels.

Inflammasomes

A variety of stimuli, including PAMPs, bacterial toxins, the common vaccine adjuvant aluminum hydroxide (alum), and ultraviolet (UV) light, as well as endogenous "danger signals" released by stressed or damaged host cells, referred to as DAMPs (*e.g.*, adenosine triphosphate [ATP], uric acid, hyaluronan), induce the processing of pro-IL-1β to mature IL-1β. The cytosolic cellular machinery responsible for IL-1β processing is termed the *inflammasome*. Prototypical inflammasomes include the NOD-like receptor-related protein-1 (NLRP1) inflammasome, the NLRP3

ON THE HORIZON

- Better understanding of populations of innate immune cells found within tissues affected by a variety of diseases (*e.g.,* psoriasis, atopic dermatitis, nasal polyps, tumors) will likely lead to targeted therapies (*e.g.,* anti–interleukin [IL]-17 for treatment of psoriasis).
- Knowledge of how individual and combined pattern recognition receptor (PRR) activation directs both innate immune activation (including trained immunity) as well as differentiation of naïve T cells may inform development of more effective vaccines.
- A better understanding of age-specific innate immune function should lead to development of new vaccine strategies to protect vulnerable populations, including newborns, young infants, and older adults.

inflammasome, and the IL-1β-converting enzyme protease-activating factor (IPAF) inflammasome. NLRP recruit ASC (apoptosis-associated speck-like protein containing a CARD) via homotypic PYRIN domain interactions. ASC is an adaptor protein that contains a PYRIN domain and a CARD domain. Caspase-1 is subsequently recruited via CARD domains, leading to proteolytic processing of IL-1β and IL-18 (see Fig. 3.6). Although NLRP1, NLRP3, and IPAF form prototypical inflammasomes, other NLR, including NOD2 and neuronal apoptosis inhibitory protein (NAIP), can also form inflammasomes or modulate their activities.[67]

Mutations in the *NLRP3* gene are associated with a spectrum of autosomal dominant inherited autoinflammatory diseases, including Muckle Wells syndrome (sensorineural deafness, urticaria, fevers, chills, arthritis), familial cold autoinflammatory syndrome (rash, conjunctivitis, fever, chills, arthralgias elicited by cold exposure), and neonatal-onset multisystem inflammatory disease (NOMID) (rashes, arthritis, chronic meningitis). These disorders are the result of constitutive production of IL-1β. Treatment of these disorders involves antiinflammatory therapy, including IL-1 antagonists.[68]

INNATE IMMUNITY IN CLINICAL PRACTICE

Our growing understanding of the molecular and cellular basis for innate immunity is paralleled by increasing appreciation of its importance to clinical medicine. Examples include *(i)* recognition of novel primary immunodeficiencies in the innate immune pathways presenting with heightened susceptibility to bacterial (*e.g.,* MyDD88-deficiency and IRAK-4-deficiency) and fungal (*e.g.,* CARD9-deficiency) infection; *(ii)* appreciation of immune ontogeny (*i.e.,* the change in immune function in a given individual across their life span, which parallels age-specific susceptibility to infection); *(iii)* clinical studies of the potential use of APP congeners to prevent or treat infection; *(iv)* leveraging PRR agonists, such as TLR agonists, as vaccine adjuvants (e.g., the TLR4 agonists monophosphoryl lipid A [MPLA] a detoxified form of lipopolysaccharide or endotoxin used in a licensed human papillomavirus [HPV] vaccine [Cervarix]); and *(v)* employing certain live attenuated vaccines, such as Bacille Calmette-Guérin (BCG) or measles, mumps, rubella (MMR) vaccine that activate multiple PRRs to reduce infections to antigenically distinct pathogens—so called beneficial heterologous or “nonspecific” effects—that may be mediated via epigenetic reprograming of monocytes and enhancement of innate immune responses.[2,12,69,70]

Our understanding of the complexity of innate immunity has advanced considerably over the past decade. The importance of the innate immune system to human health is underscored by single gene mutations, such as IRAK4 deficiency, that result in immune deficiencies and infection, particularly in early life.[71] Continued elucidation of the cell types comprising the innate immune system has expanded our understanding of the delicate balance between tolerance to nonpathogenic commensal bacteria and inflammatory responses to pathogenic, invasive microbes. As innate immunity is expressed in an age-specific manner, a better understanding of the ontogeny of the innate immune system, as well as the mechanisms by which innate and adaptive immunity interact, will guide development of adjuvants, resulting in more effective vaccines and tumor immunotherapy.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Janeway CA Jr, Bottomly K. Signals and signs for lymphocyte responses. Cell 1994;76(2):275–85.
2. Netea MG, Joosten LA, Latz E, et al. Trained immunity: a program of innate immune memory in health and disease. Science 2016;352:aaf1098.
3. Netea MG, Quintin J, van der Meer JW. Trained immunity: a memory for innate host defense. Cell Host Microbe 2011;9:355–61.
4. Bantz SK, Zhu Z, Zheng T. The atopic march: progression from atopic dermatitis to allergic rhinitis and asthma. J Clin Cell Immunol 2014;5:pii, 202.
5. Ong PY, Ohtake T, Brandt C, et al. Endogenous antimicrobial peptides and skin infections in atopic dermatitis. N Engl J Med 2002;347:1151–60.
6. Brandner JM, Zorn-Kruppa M, Yoshida T, et al. Epidermal tight junctions in health and disease. Tissue Barriers 2015;3:e974451.
7. Nuding S, Fellermann K, Wehkamp J, et al. Reduced mucosal antimicrobial activity in Crohn’s disease of the colon. Gut 2007;56:1240–7.
8. Hazlett L, Wu M. Defensins in innate immunity. Cell Tissue Res 2011;343:175–88.
9. Martin L, van Meegern A, Doemming S, et al. Antimicrobial peptides in human sepsis. Front Immunol 2015;6:404.
10. McDermott AM. Defensins and other antimicrobial peptides at the ocular surface. Ocul Surf 2004;2:229–47.
11. Schutte BC, McCray PB Jr. [beta]-defensins in lung host defense. Annu Rev Physiol 2002;64:709–48.
12. Kosikowska P, Lesner A. Antimicrobial peptides (AMPs) as drug candidates: a patent review (2003–2015). Expert Opin Ther Pat 2016;1–14.
13. Hiemstra PS, Amatngalim GD, van der Does AM, et al. Antimicrobial peptides and innate lung defenses: role in infectious and noninfectious lung diseases and therapeutic applications. Chest 2016;149:545–51.
14. Deban L, Jaillon S, Garlanda C, et al. Pentraxins in innate immunity: lessons from PTX3. Cell Tissue Res 2011;343:237–49.
15. Miyake K. Endotoxin recognition molecules, Toll-like receptor 4-MD-2. Semin Immunol 2004;16:11–16.
16. Sharon N. Carbohydrates as recognition determinants in phagocytosis and in lectin-mediated killing of target cells. Biol Cell 1984;51:239–45.
17. Ledford JG, Pastva AM, Wright JR. Review: collectins link innate and adaptive immunity in allergic airway disease. Innate Immun 2010;16:183–90.
18. Ricklin D, Hajishengallis G, Yang K, et al. Complement: a key system for immune surveillance and homeostasis. Nat Immunol 2010;11:785–97.
19. Rickert RC. Regulation of B lymphocyte activation by complement C3 and the B cell coreceptor complex. Curr Opin Immunol 2005;17:237–43.
20. Levy O, Jean-Jacques RM, Cywes C, et al. Critical role of the complement system in group B streptococcus-induced tumor necrosis factor alpha release. Infect Immun 2003;71:6344–53.

21. Kaplan AP. Enzymatic pathways in the pathogenesis of hereditary angioedema: the role of C1 inhibitor therapy. J Allergy Clin Immunol 2010;126:918–25.
22. Frank MM. Complement disorders and hereditary angioedema. J Allergy Clin Immunol 2010;125:S262–71.
23. Daha MR. Role of complement in innate immunity and infections. Crit Rev Immunol 2010;30(1):47–52.
24. Herbert AP, Deakin JA, Schmidt CQ, et al. Structure shows that a glycosaminoglycan and protein recognition site in factor H is perturbed by age-related macular degeneration-linked single nucleotide polymorphism. J Biol Chem 2007;282:18960–8.
25. Eisen DP. Mannose-binding lectin deficiency and respiratory tract infection. J Innate Immun 2010;2:114–22.
26. Cros J, Cagnard N, Woollard K, et al. Human CD14dim monocytes patrol and sense nucleic acids and viruses via TLR7 and TLR8 receptors. Immunity 2010;33:375–86.
27. Italiani P, Boraschi D. From monocytes to M1/M2 macrophages: phenotypical vs. functional differentiation. Front Immunol 2014;5:514.
28. Dale DC, Boxer L, Liles WC. The phagocytes: neutrophils and monocytes. Blood 2008;112:935–45.
29. Bokoch GM, Diebold B, Kim JS, et al. Emerging evidence for the importance of phosphorylation in the regulation of NADPH oxidases. Antioxid Redox Signal 2009;11:2429–41.
30. Reeves EP, Lu H, Jacobs HL, et al. Killing activity of neutrophils is mediated through activation of proteases by K^+ flux. Nature 2002;416:291–7.
31. Kobayashi Y. The regulatory role of nitric oxide in proinflammatory cytokine expression during the induction and resolution of inflammation. J Leukoc Biol 2010;88:1157–62.
32. Bryceson YT, Chiang SC, Darmanin S, et al. Molecular mechanisms of natural killer cell activation. J Innate Immun 2011;3:216–26.
33. Stojanovic A, Cerwenka A. Natural killer cells and solid tumors. J Innate Immun 2011;3:355–64.
34. Berzins SP, Smyth MJ, Baxter AG. Presumed guilty: natural killer T cell defects and human disease. Nat Rev Immunol 2011;11:131–42.
35. Umetsu DT, Dekruyff RH. Natural killer T cells are important in the pathogenesis of asthma: the many pathways to asthma. J Allergy Clin Immunol 2010;125:975–9.
36. Levy O, Orange JS, Hibberd P, et al. Disseminated varicella infection due to the vaccine strain of varicella-zoster virus, in a patient with a novel deficiency in natural killer T cells. J Infect Dis 2003;188: 948–53.
37. Miller LS. Toll-like receptors in skin. Adv Dermatol 2008;24:71–87.
38. Villablanca EJ, Mora JR. A two-step model for Langerhans cell migration to skin-draining LN. Eur J Immunol 2008;38:2975–80.
39. Hayday A, Theodoridis E, Ramsburg E, et al. Intraepithelial lymphocytes: exploring the Third Way in immunology. Nat Immunol 2001;2:997–1003.
40. Qiu Y, Wang W, Xiao W, et al. Role of the intestinal cytokine microenvironment in shaping the intraepithelial lymphocyte repertoire. J Leukoc Biol 2015;pii: jlb.3RU1014-465R.
41. Zhang X. Regulatory functions of innate-like B cells. Cell Mol Immunol 2013;10:113–21.
42. Dadi S, Chhangawala S, Whitlock BM, et al. Cancer immunosurveillance by tissue-resident innate lymphoid cells and innate-like T cells. Cell 2016;164:365–77.
43. Bonefeld CM, Geisler C. The role of innate lymphoid cells in healthy and inflamed skin. Immunol Lett 2016;179:25–8.
44. Montaldo E, Vacca P, Vitale C, et al. Human innate lymphoid cells. Immunol Lett 2016;179:2–8.
45. Galli SJ, Tsai M. Mast cells in allergy and infection: versatile effector and regulatory cells in innate and adaptive immunity. Eur J Immunol 2010;40:1843–51.
46. Echtenacher B, Mannel DN, Hultner L. Critical protective role of mast cells in a model of acute septic peritonitis. Nature 1996;381:75–7.
47. Kawai T, Akira S. Toll-like receptor and RIG-I-like receptor signaling. Ann N Y Acad Sci 2008;1143:1–20.
48. Kawai T, Akira S. The role of pattern-recognition receptors in innate immunity: update on Toll-like receptors. Nat Immunol 2010;11:373–84.
49. Malley R, Henneke P, Morse SC, et al. Recognition of pneumolysin by Toll-like receptor 4 confers resistance to pneumococcal infection. Proc Natl Acad Sci USA 2003;100:1966–71.
50. Vabulas RM, Wagner H, Schild H. Heat shock proteins as ligands of toll-like receptors. Curr Top Microbiol Immunol 2002;270:169–84.
51. Beutler B, Jiang Z, Georgel P, et al. Genetic analysis of host resistance: Toll-like receptor signaling and immunity at large. Annu Rev Immunol 2006;24:353–89.
52. Kawai T, Akira S. Toll-like receptors and their crosstalk with other innate receptors in infection and immunity. Immunity 2011;34:637–50.
53. O'Neill LA. The interleukin-1 receptor/Toll-like receptor superfamily: 10 years of progress. Immunol Rev 2008;226:10–18.
54. Tabeta K, Hoebe K, Janssen EM, et al. The Unc93b1 mutation 3d disrupts exogenous antigen presentation and signaling via Toll-like receptors 3, 7 and 9. Nat Immunol 2006;7:156–64.
55. Shinohara ML, Lu L, Bu J, et al. Osteopontin expression is essential for interferon-alpha production by plasmacytoid dendritic cells. Nat Immunol 2006;7:498–506.
56. Doyle SL, Jefferies CA, Feighery C, et al. Signaling by Toll-like receptors 8 and 9 requires Bruton's tyrosine kinase. J Biol Chem 2007;282:36953–60.
57. Napolitani G, Rinaldi A, Bertoni F, et al. Selected Toll-like receptor agonist combinations synergistically trigger a T helper type 1-polarizing program in dendritic cells. Nat Immunol 2005;6:769–76.
58. Kasturi SP, Skountzou I, Albrecht RA, et al. Programming the magnitude and persistence of antibody responses with innate immunity. Nature 2011;470:543–7.
59. Palmer CD, Mancuso CJ, Weiss JP, et al. 17(R)-Resolvin D1 differentially regulates TLR4-mediated responses of primary human macrophages to purified LPS and live E. coli. J Leukoc Biol 2011;90:459–70.
60. Ignacio A, Morales CI, Camara NO, et al. Innate sensing of the gut microbiota: modulation of inflammatory and autoimmune diseases. Front Immunol 2016;7:54.
61. Philpott DJ, Sorbara MT, Robertson SJ, et al. NOD proteins: regulators of inflammation in health and disease. Nat Rev Immunol 2014;14:9–23.
62. Sander LE, Davis MJ, Boekschoten MV, et al. Detection of prokaryotic mRNA signifies microbial viability and promotes immunity. Nature 2011;474:385–9.
63. Osorio F, Reis e Sousa C. Myeloid C-type lectin receptors in pathogen recognition and host defense. Immunity 2011;34:651–64.
64. Gross O, Gewies A, Finger K, et al. Card9 controls a non-TLR signalling pathway for innate anti-fungal immunity. Nature 2006;442:651–6.
65. Gringhuis SI, den Dunnen J, Litjens M, et al. C-type lectin DC-SIGN modulates Toll-like receptor signaling via Raf-1 kinase-dependent acetylation of transcription factor NF-kappaB. Immunity 2007;26:605–16.
66. Greaves DR, Gordon S. The macrophage scavenger receptor at 30 years of age: current knowledge and future challenges. J Lipid Res 2009;50:S282–6.
67. Schroder K, Tschopp J. The inflammasomes. Cell 2010;140:821–32.
68. Geyer M, Muller-Ladner U. Actual status of antiinterleukin-1 therapies in rheumatic diseases. Curr Opin Rheumatol 2010;22:246–51.
69. Alsina L, Israelsson E, Altman MC, et al. A narrow repertoire of transcriptional modules responsive to pyogenic bacteria is impaired in patients carrying loss-of-function mutations in MYD88 or IRAK4. Nat Immunol 2014;15:1134–42.
70. Levy O, Netea MG. Innate immune memory: implications for development of pediatric immunomodulatory agents and adjuvanted vaccines. Pediatr Res 2014;75:184–8.
71. Picard C, von Bernuth H, Ghandil P, et al. Clinical features and outcome of patients with IRAK-4 and MyD88 deficiency. Medicine (Baltimore) 2010;89:403–25.

MULTIPLE-CHOICE QUESTIONS

1. Which of the following statements regarding the complement system is true?
   - **A.** Defects in early components of the complement cascade (*i.e.,* C1, C2, or C3) are associated with increased susceptibility to meningococcal infections.
   - **B.** Paroxysmal nocturnal hemoglobinuria results from mutations within the *CD55* gene.
   - **C.** The classical complement pathway is initiated by the spontaneous generation of C3b in the plasma.
   - **D.** Hereditary angioedema results from mutations in the C1 esterase inhibitor gene.

2. Which of the following are characteristics of natural killer (NK) cells?
   - **A.** NK cells require two signals to kill targets.
   - **B.** NK-cell function is regulated by a balance between inhibiting and activating signals.
   - **C.** CD56$^{\text{bright}}$ NK cells have potent cytotoxicity activity.
   - **D.** NK cells express somatically rearranged receptors.

3. Which of the following cell types is responsible for mediating type 1 hypersensitivity reactions:
   - **A.** Dendritic cells
   - **B.** Neutrophils
   - **C.** Natural killer T (NKT) cells
   - **D.** Mast cells

4. Inflammasome activation leads to release of mature interleukin-1 (IL-1) and IL-18. Activation of which of the following receptors leads to activation of the inflammasome?
   - **A.** Nucleotide-binding oligomerization domain (NOD)-like receptor–related proteins
   - **B.** Toll-like receptors (TLRs)
   - **C.** Complement receptors
   - **D.** RIG-I–like receptors

4

Antigen Receptor Genes, Gene Products, and Coreceptors

Harry W. Schroeder, Jr., John B. Imboden, Raul M. Torres

In 1890, von Behring and Kitasato reported the existence of an agent in blood that could neutralize the diphtheria toxin. The following year, glancing references were made to "*Antikörper*" in studies describing the ability of the agent to discriminate between two immune bodies, or substances. The term *antigen* is a shortened form of "*Antisomatogen + Immunkörperbildner,*" or the substance that induces the production of an antibody (Chapter 6). Thus the definition of antibody and antigen represent a classic tautology.

In 1939, Tiselius and Kabat used electrophoresis to separate immunized serum into albumin, α-globulin, β-globulin, and γ-globulin fractions. Absorption of the serum against the antigen depleted the γ-globulin fraction; yielding the terms *gammaglobulin, immunoglobulin* (Ig), and *IgG.* Subsequently, "sizing" columns were used to separate Igs into those that were "heavy" (pentameric IgM), "regular" (IgA, IgE, IgD, IgG, monomeric IgM), and "light" (light-chain dimers), culminating with the discovery of the last major class of immunoglobulin, IgE, in 1966.

In 1949, Porter used papain to digest IgG molecules into two types of fragments, termed *Fab* (fragment antigen-binding) and *Fc* (fragment crystallizable). The constancy of the Fc fragment permitted its crystallization and thus the elucidation of its sequence and structure. The variability of the Fab fragment precluded analysis until Bence-Jones myeloma proteins were identified as clonal, isolated light chains.

In 1976, Hozumi and Tonegawa demonstrated that the variable portion of κ chains was the product of the rearrangement of variable (V) and joining (J) gene segments. In 1982, Alt and Baltimore reported that terminal deoxynucleotidyl transferase (TdT) could be used to introduce non-germline–encoded sequence between rearranging V, D (diversity), and J gene segments, potentially freeing the preimmune heavy-chain repertoire from germline constraints. In 1984, Weigert et al. determined that during affinity maturation, variable domains could undergo mutation at a rate of 10^{-3} per base pair (bp), per generation. These discoveries clarified how lymphocytes could generate an astronomically diverse antigen receptor repertoire from a handful of gene elements.

In 1982, Allison et al. raised antisera against a cell surface molecule that could uniquely identify individual T-cell clones. A year later, Kappler and a consortium of colleagues demonstrated that these surface molecules were heterodimers composed of variable and constant region domains, just like Igs. Subsequently, Davis and Mak independently cloned the β chain of the T-cell receptor (TCR). Initial confusion regarding the identity of the companion α chain led to the realization that there were two mutually exclusive forms of TCR, αβ and γδ.

PARATOPES AND EPITOPES

Igs and TCRs both belong to the eponymous Ig superfamily (IgSF).[1] The study of antibodies precedes that of TCR by decades; hence much of what we know is based on knowledge first gleaned from the study of Igs.

Ig–antigen interactions typically take place among the *paratope,* the site on the Ig at which the antigen binds, and the *epitope,* which is the site on the antigen that is bound. Thus lymphocyte antigen receptors do not recognize antigens, but they recognize the epitopes borne on those antigens. This makes it possible for the cell to discriminate between two closely related antigens, each of which can be viewed as a collection of epitopes. It also permits the same receptor to bind divergent antigens that share equivalent or similar epitopes, a phenomenon referred to as *cross-reactivity.*

Although both Igs and TCRs can recognize the same antigen, they do so in markedly different ways. Igs tend to recognize intact antigens in soluble form, and thus preferentially identify surface epitopes that are often composed of conformational structures noncontiguous in the antigen's primary sequence. In contrast, TCRs recognize fragments of antigens, both surface and internal, that have been processed by a separate antigen-presenting cell (APC) and then bound to a major histocompatibility complex (MHC) class I or class II molecule (Chapters 5, 6).

THE BCR AND TCR ANTIGEN RECOGNITION COMPLEX

Although the ability of the surface antigen receptor to recognize antigen was appreciated early on, the mechanism by which the membrane-bound receptor relayed this antigen recognition event into the cell interior was not understood, since both B-cell receptor (BCR) and TCR cytoplasmic domains are exceptionally short. This conundrum was solved when it was shown that BCR and TCR each associate noncovalently with signal transduction complexes: heterodimeric Igα:Igβ (also known as CD79a:CD79b, respectively) for B cells and multimeric CD3 for T cells. Loss of function mutations in either of these complexes leads to cell death, which becomes clinically manifest as hypogammaglobulinemia in the case of B cells (Chapter 34), or severe combined immune deficiency (SCID) in the case of T cells (Chapter 35).

IMMUNOGLOBULINS AND TCR STRUCTURES

The Ig Domain, the Basic IgSF Building Block

Igs consist of two heavy (H) and two light (L) chains (Fig. 4.1). The L chain can be either a κ or a λ chain. TCRs consist of either an αβ or a γδ heterodimer. Each component chain contains two or more IgSF domains, each of which consists of two sandwiched β pleated sheets "pinned" together by a disulfide bridge between two conserved cysteine residues.[1] Considerable variability is allowed to the amino acids that populate the external surface of the IgSF domain and the loops that link the β strands. These solvent exposed surfaces offer multiple targets for docking with other molecules.

Two types of IgSF domains, "constant" (C) and "variable" (V), are used in Igs and TCRs (see Fig. 4.1). C-type domains, which are the most compact, have seven antiparallel strands distributed as three strands in the first sheet and four strands in the second. Side chains positioned to lie between the two strands tend to be nonpolar in nature, creating a hydrophobic core of considerable stability. V-type domains add two additional antiparallel strands to the first sheet, creating a five-strand–four-strand distribution. The two additional strands, which encode framework region 2 (FR2), are used to steady the interaction between heterodimeric V domains, allowing them to create a stable antigen-binding site.[2]

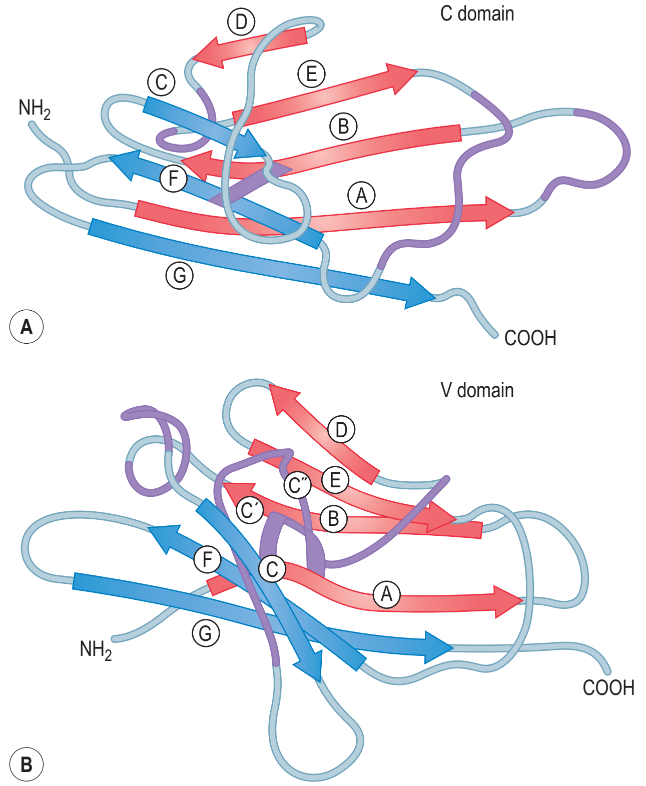


FIG 4.1 Immunoglobulin Superfamily (IgSF) Domain Structures. (**A**) A typical compact C domain structure. The β strands are labeled A through G. The sequence at the core is conserved and nonpolar. The external surface and the β-loops are available for docking and often vary in sequence. (**B**) A typical V domain structure. Two additional strands, C′ and C″, have been added. Note the projection of the C-C′ strands and loop away from the core.

Although each Ig or TCR chain contains only one amino-terminal V Ig domain, the number of carboxy-terminal C domains varies. Ig H chains contain between three and four C domains, whereas both Ig L chains and all four TCR chains contain only one C domain each. IgH chains with three C domains tend to include a spacer hinge region between the first (C_H1) and second (C_H2) domains. Each V or C domain consists of approximately 110–130 amino acids, averaging 12 000–13 000 kilodaltons (kDa). A typical L chain will mass approximately 25 kDa, whereas a three C domain Cγ H chain with its hinge and tail will mass approximately 55 kDa.

Idiotypes and Isotypes

Immunization of heterologous species with monoclonal antibodies (mAbs; or a restricted set of Igs) has shown that Igs and TCRs contain both common and individual antigenic determinants. Individual determinant(s), termed *idiotype(s)*, are contained within V domains. Common determinants, termed *isotypes*, are specific for the constant portion of the antibody and allow grouping of Igs and TCRs into recognized classes. Each class defines an individual type of C domain. Determinants common to subsets of individuals within a species, yet differing between other members of that species, are termed *allotypes* and define inherited polymorphisms that result from allelic forms of the genes.[3]

The V Domain

Early comparisons of the primary sequences of V domains identified three hypervariable intervals, termed *complementarity-determining regions* (CDRs), situated between four framework regions of stable sequence (Fig. 4.2). The current definition of these regions integrates sequence diversity with three-dimensional structure.[4] The international ImMunoGeneTics information system (IMGT) maintains an extremely useful website (http://www.imgt.org), which contains a large database of Ig and TCR sequences and numerous software tools for their analyses.

Antigen Recognition and Fab

Studies of Ig structure were facilitated by the use of papain and pepsin to fragment IgG molecules. Papain digests IgG into two antigen-binding fragments (Fab) and a single crystallizable (or constant) fragment (Fc). Pepsin splits IgG into an Fc fragment and a single dimeric $F(ab')_2$ that can cross-link as well as bind antigens. Fab contains one complete L chain in its entirety and the V and C_H1 portion of one H chain (see Fig. 4.2). Fab can be further divided into a variable fragment (Fv) composed of the V_H and V_L domains and a constant fragment (Fb) composed of the C_L and C_H1 domains. Single Fv fragments can be genetically engineered to recapitulate the monovalent antigen-binding characteristics of the original, parent antibody.[5] The extracellular domains of TCRαβ and TCRγδ correspond to Ig Fab.

Effector Function and Fc

The Fc portion (see Fig. 4.2) encodes the effector functions of the Ig. These functions are generally inflammatory reactions, which include fixation and activation of complement, and binding of antibody to Fc receptors on the surface of other cells. Each Ig class and subclass exhibits its own set of effector functions.[6]

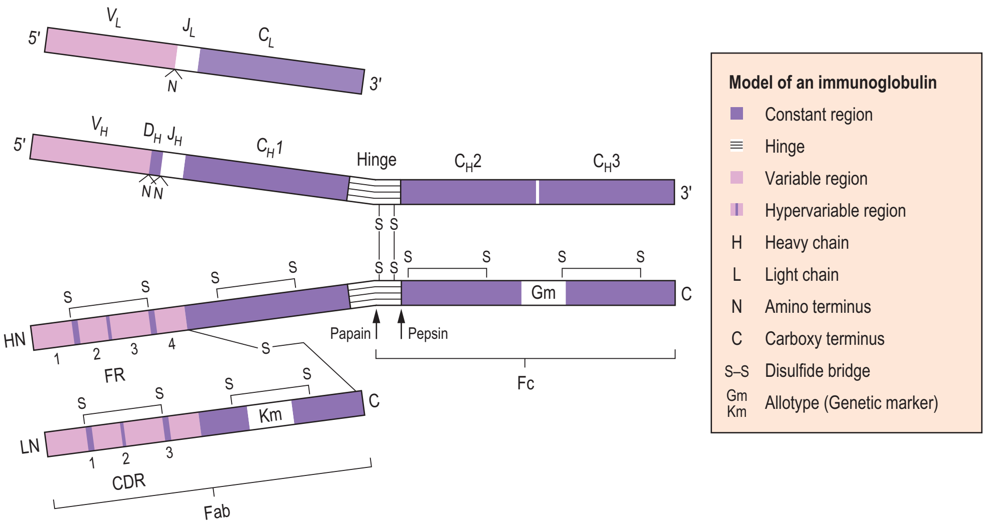


FIG 4.2 **A Two-Dimensional Model of an Immunoglobulin G (IgG) Molecule.** The top H and L chains illustrate the composition of these molecules at a nucleotide level. The bottom chains illustrate the nature of the protein sequence. See text for further details.

TABLE 4.1 Selected Properties of Immunoglobulin (Ig) Classes

	IgG	IgA	IgM	IgD	IgE
Molecular weight	160 000	170 000 or polymer	900 000	160 000	180 000
Approximate concentration in serum (mg/dL)	1000–1500	250–300	100–150	0.3–30	0.0015–0.2
Valence	2	2 (monomer)	10 (small antigen) 5 (large antigen)	2	2
Molecular formula	γ_2L_2	$(\alpha_2L_2)n$	$(\mu_2L_2)_5$	δ_2L_2	ε_2L_2
Half-life (days)	23	6	5	3	2.5
Special property	Placental passage	Secretory Ig	Primary response lymphocyte surface	Lymphocyte surface	Immediate hypersensitivity reactions

KEY CONCEPTS

Immunoglobulin (Ig) and T-Cell Receptor (TCR) Structure

- Both Igs and TCRs are heterodimeric proteins.
- Igs consist of two identical H and two L chains.
- αβ TCRs consist of one α and one β chain.
- γδ TCRs consist of one γ and one δ chain.
- Igs and TCR contain two or more immunoglobulin superfamily (IgSF) domains, which are identified by their characteristic β barrel structure.
- Each Ig and TCR chain contains a V-type IgSF domain that will form one-half of the antigen-binding site.
- Each V domain contains three hypervariable intervals known as complementarity determining regions (CDRs). The CDRs of paired heterodimers chains are juxtaposed to form the antigen-binding site.
- The C domains of Ig H chains define the Ig class or subclass.
- The two distal C IgH domains determine the effector function of the antibody.

For example, the IgG C_H2 domain plays a key role in complement fixation and in binding to class-specific Fc receptors on the surface of effector cells. Both these interactions are important in initiating the process of phagocytosis, in allowing certain subclasses to traverse the placenta, and in influencing the biological functions of lymphocytes, platelets, and other cells.

Gm Allotype System

A series of serologically defined C domain allotypes have been identified. In the case of the H chain, they are termed *Gm* for the gammaglobulin fraction of the serum in which they were first identified.[3] Allotypes have been identified for γ1, γ2, γ3, γ4, α2, and ε H chains and for the κ L chain. Associations between certain Gm allotypes and predisposition to develop certain diseases of immune function have been reported.

Ig CLASSES AND SUBCLASSES

The constant domains of the H chain define the class and subclass of the antibody. Table 4.1 lists the five major classes of Igs in humans and describes some of the physical and chemical features of these Igs. Two of the five major H chain classes, α and γ, have undergone duplication. IgG1, IgG2, IgG3, and IgG4 all have the same basic structural design and differ only in the primary sequence of their constant regions and in the location of their interchain disulfide bonds. The H chain in each of these subclasses is referred to as γ1, γ2, and so on. IgA consists of the two subclasses, α1 and α2. Table 4.2 compares the four subclasses of IgG, the two of IgA, and the classes of IgM, IgD, and IgE from the standpoint of their biological functions. In humans, the two L chain classes, κ and λ, are expressed at roughly equal frequencies. No specific effector function has been identified for either L chain class.

TABLE 4.2 **Selected Biological Properties of Classes and Subclasses of Immunoglobulins (Igs)**

	IgG				IgA				
	1	2	3	4	1	2	IgM	IgD	IgE
Percentage of total (%)	65	20	10	5	90	10			
Complement fixation	++	+	++	–	–	–	++	–	–
Complement fixation (alternative)			+	+	+/–	+/–			
Placental passage	+	+	+	+	–	–	–	–	–
Fixing to mast cells or basophils	–	–	–	–	–	–	–	–	+
Binding to:									
Lymphocytes	+	+	+	+	–	–	+	–	–
Macrophages	+	+/–	+	+/–	–	–	–	–	–
Neutrophils	+	+	+	+	+	+	–	–	–
Platelets	+	+	+	+	–	–	–	–	–
Reaction with *Staphylococcus* protein A	+	+	–	+	–	–	–	–	–
Half-life (days)	23	23	8-9	23	6	6	5	3	2.5
Synthesis mg/kg/day	25	?	3.5	?	24	?	7	0.4	0.02

IgM

IgM exists in monomeric, pentameric, and hexameric forms. The 8S monomeric 180 kDa IgM has the molecular formula μ_2L_2. It is a minor fraction in serum, but in its transmembrane form IgM plays a key role in B-cell development and function as the antigen recognition portion of the B-cell antigen receptor. The major form in serum is the 19S, 900 000 Da pentameric IgM, which contains five subunits $[(\mu_2L_2)_5]$ linked together by disulfide bridges and by one molecule of an additional polypeptide chain, the J chain, which joins two of the subunits by a disulfide bridge.[7]

IgM is the predominant Ig produced during the primary immune response. Occasionally, particularly in the case of carbohydrate antigens such as isohemagglutinins, it will remain the major or sole antibody class. IgM differs from most other Igs in having an extra C_H domain in place of a hinge.

IgM avidly fixes complement. This property is focused in CH3, the homologue of IgG CH2.[8] Although the valence of each μ_2L_2 subunit is 2, when binding to large protein antigens, five of the 10 antigen-binding sites in pentameric IgM appear blocked because of steric hindrance. As a consequence, the valence for large antigens is five.

IgG

IgG, the major Ig class, accounts for the bulk of serum antibody activity in response to most protein antigens. The four IgG subclasses are numbered in relation to their serum levels relative to each other, with IgG1 being predominant and IgG4 being the least common. IgG1 and IgG3 fix complement and bind phagocyte Fcγ receptors well, whereas IgG2 fixes complement but binds Fcγ receptors more poorly. IgG4 does not fix complement effectively in the native state but has been reported to do so after proteolytic cleavage.[9] IgG1 and IgG3 are most frequently elicited by viral antigens,[10] IgG2 by carbohydrates,[11] and IgG4 by helminthic parasites.[12]

IgG4 can attenuate allergic responses by inhibiting the activity of IgE.[13] IgG4 can function as a blocking antibody, preventing cross-linking of receptor-bound IgE. It can costimulate the inhibitory IgG receptor FcγRIIb, which can negatively regulate FcεRI signaling and thus inhibit effector cell activation. Finally, the disulfide bonds of the IgG4 hinge are easily reduced, which allows the H chains to separate and randomly reassociate to produce a mixed population of IgG4 molecules with randomized heavy-chain and light-chain pairs. This impairs the ability of IgG4 to form immune complexes and thus has an anti-inflammatory effect, facilitating immunotherapy for allergic diseases (allergy shots).

Overproduction of IgG4 is seen in a disparate group of inflammatory diseases. Fibroinflammatory masses can develop in virtually all organs except the brain and have an unexplained preference for salivary glands, lymph nodes, and the pancreas. Together, these are referred to as *IgG4-related disease* (IgG4-RD).[14]

IgA

Although IgA generally exists in a monomeric form (α_2L_2) in serum, it can interact with the J chain to form a polymer $(\alpha_2L_2)_{2,3}$-J. Second in concentration to IgG in serum, IgA functions as the predominant form of Ig in mucosal secretions.[15]

Secretory IgA (sIgA) is largely synthesized by plasma cells associated with mucosal tissues. In secretions, the molecule typically exists in polymeric form with two subunits in association with the 70 kDa secretory component $(\alpha_2L_2)_2$-SC. SC is synthesized by the epithelial cells that line the lumen of the gut. It appears to render the secretory IgA complex more resistant to proteolytic digestion, and it enhances the immune functions of SIgA.

IgE

IgE is largely found in extravascular spaces. Its plasma turnover is rapid, with a half-life of about 2 days. IgE antibodies help protect the host from parasitic infections (Chapter 31). In Westernized, affluent societies, IgE is primarily associated with allergy. Through their interaction with Fcε receptors on mast cells and basophils, IgE antibodies, in the presence of antigens, induce the release of histamine and various other vasoactive substances, which are responsible for clinical manifestations of various allergic states.[16]

IgD

Although the H chain of IgD can undergo alternative splicing to a secretory form, IgD serum antibodies in humans are uncommon and are absent in the serum of mice and primates. Instead, IgD typically is coexpressed with IgM on the surface of mature lymphocytes. The appearance of IgD is associated with the transition of a B lymphocyte from a cell that can be tolerized to the antigen to a cell that will respond to antigen with the production of antibody (Chapter 7).

TCR αβ AND γδ

TCR α, β γ, and δ chains are members of the IgSF and thus share a number of structural similarities with Igs. Each chain contains a leader peptide and extracellular, transmembrane, and intracytoplasmic components. The extracellular component can be divided into three domains: a polymorphic V domain encoded by VJ (α and γ chains) or VDJ (β and δ chains) gene segments, a C domain, and a hinge region.[17] The hinge region typically contains an extra cysteine (none in γ chains encoded by Cγ2) that forms a disulfide bond with the other partner of the heterodimer. All of the transmembrane domains include a lysine plus or minus an arginine residue that facilitates the association of the TCR heterodimer with components of the CD3 signal transduction complex, each of which has a matching negatively charged residue in its own transmembrane portion (see below). The intracytoplasmic components are tiny and play a minimal role in signal transduction.

TCR αβ

The TCR α and β chains are glycoproteins with molecular weights that vary from 42 to 45 kDa, depending on the primary amino acid sequence and the degree of glycosylation. Deglycosylated forms have a molecular mass of 30 to 32 kDa. These chains share a number of invariant residues in common with Ig heavy and light chains, in particular residues that are thought to be important for interactions between heavy and light chains. The structures of over 30 partial or full length TCRs have been solved by X-ray crystallography (Fig. 4.3).[18] In general, the structure of the TCR αβ heterodimer is similar, but not identical, to that of an Ig Fab fragment.

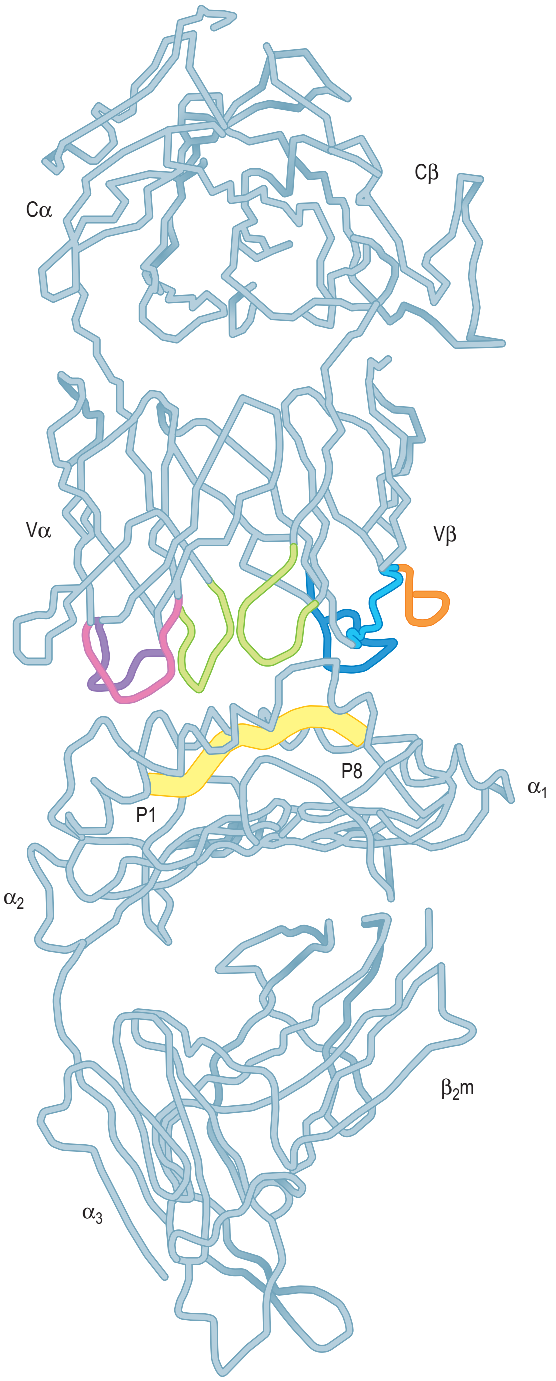


FIG 4.3 **Backbone Representation of Murine αβ T-Cell Receptor (TCR) Bound to Murine Major Histocompatibility Complex (MHC) Class I and an Octamer Peptide.** The TCR is above. The Vα CDR1 and CDR2 are depicted in magenta, Vβ CDR1 and CDR2 in blue, both CDR3s in green, and the Vβ HV4 in orange. β2M refers to β_2 microglobulin. The peptide is in yellow, and the NH_2-terminal and COOH-terminal residues are designated P1 and P8. (Reproduced with permission from Garcia KC, Degano M, Stanfield RL, Brunmark A, Jackson MR, Peterson PA, et al. An alphabeta T cell receptor structure at 2.5 A and its orientation in the TCR-MHC complex. Science. 1996;274(5285):209-19.Garcia et al.)

TCR γδ

The TCR γ and δ chains are glycoproteins with a more complex molecular size pattern than α and β chains. TCRs that use the Cγ1 gene segment, which contains a cysteine-encoding exon, are disulfide-linked (MW 36–42 kDa). TCRs that use Cγ2 exist in two non–disulfide-linked forms, one of 40–44 kDa and one of 55 kDa.[19] The differences in molecular size result from the variability of both N-linked glycosylation and primary amino acid sequence. The 55-kDa form uses a Cγ2 allele that contains three (rather than two) exons encoding the connecting piece, as well as more N-linked carbohydrate. The TCR δ chain is more straightforward, being 40–43 kDa in size and containing two sites of N-linked glycosylation. The overall architecture of the γδ TCR closely resembles that of αβ TCRs and antibodies, although the angle between the V and C domains, known as the *elbow angle,* appears more acute.

Ligand Recognition

TCR αβ T cells primarily recognize peptide-MHC complexes (pMHC) (see Fig. 4.3; Chapters 5, 6); however, other types of ligands exist. For example, some αβ TCRs can bind nonpeptidic antigens (atypical antigens) that are bound to “nonclassic” MHC class Ib molecules. The αβ TCR expressed by NKT cells recognizes lipid antigens associated with the MHC class I related CD1 surface receptor. Many γδ T cells recognize atypical antigens that may or may not be associated with an antigen-presenting molecule, although some can bind peptides. Finally, many αβ TCRs bind superantigens (SAgs) in a predominantly Vβ-dependent fashion (Chapter 6).

Binding to pMHC

TCRs recognize peptide antigens bound to the binding groove of MHC-encoded glycoproteins (see Fig. 4.3). TCR recognition of pMHC requires a trimolecular complex in which all the components (antigen, MHC, and TCR) contact one another.[18] Thus recognition is highly influenced by polymorphisms in the MHC molecule (Chapter 5). As in the case of Igs, TCR CDR1 and CDR2 are encoded in the germline V regions, whereas CDR3 is formed at the junction of the V gene with a J gene segment (TCR α and γ) or D and J gene segments (TCR β and δ chains). Vβ also has a fourth region of variability within Framework 3 that is juxtaposed to the other CDRs in the tertiary structure. This region, variously termed hypervariable region 4 (HV4) or CDR4, can participate in SAg binding.

The cocrystallization of different combinations of soluble TCR αβ interacting with MHC class I bound to antigen peptide (pMHC) has made it possible to directly address the manner in which antigen recognition occurs (see Fig. 4.3). The TCR αβ combining site is relatively flat, allowing it to interact with a rather broad surface at the point of contact with pMHC.

The TCR footprint on the pMHC complex tends to occur in a diagonal across the MHC antigen-binding groove, with TCR Vα positioned over the MHC α_2 helix and TCR Vβ overlying the MHC α_1 helix. This geometry would permit consistent access of the CD8 coreceptors to the MHC class I molecule. The CDR1 and CDR2 loops, which are entirely encoded by germline sequence, tend to interact more with the MHC molecule, whereas the CDR3 loops, which are composed of both germline and somatic (N addition) sequences, appear to dominate the interaction with the MHC-bound peptide.

The binding of TCR to pMHC appears to be driven by enthalpy—that is, binding increases the stability of the CDR loops, especially CDR3. These results have led to the suggestion that initial binding focuses on the interaction between CDRs 1 and 2 and the MHC. After this initial recognition, the CDR3s change their shape to maximize the area of contact. Conformational flexibility, or "induced fit," would allow TCRs to rapidly sample many similar pMHC complexes, stopping only when their CDR3s are able to stabilize the interaction.

TCR Binding Affinity

The affinity with which the TCR ultimately binds its ligand is a critical determinant of T-cell activation. It is, however, only one factor in determining the overall avidity of the interaction, since other cell surface molecules of the T cell (e.g., CD4, CD8, CD2, and various integrins) bind to cell surface molecules on the antigen-bearing cell to stabilize cell–cell TCR–ligand interactions. Furthermore, since both the TCR and the pMHC ligand are surface membrane proteins, each T cell can provide multiple TCRs in the same plane that can bind multiple pMHC molecules on the surface of the APC. This makes binding of TCR to pMHC functionally multivalent, enhancing the apparent affinity of the interaction.

Atypical Antigens

Some αβ T cells can recognize lipid antigens when they are complexed with members of the CD1 family.[20] The interaction of TCR αβ with CD1 resembles that of TCR αβ with MHC class I. Allelic polymorphism in CD1 is limited, which theoretically would restrict the range of lipid antigens that can be bound.

Rather than binding to a single groove on the MHC, lipids attach themselves to one of several hydrophobic pockets that can be found on the surface of CD1. Pocket volume can range from 1300 to 2200 $Å^3$. The number and length of the pockets differ between the various CD1 isoforms. For example, CD1b has three pockets that share a common portal of entry, as well as a fourth pocket that connects two of the three pockets to each other. This connecting pocket permits the insertion of lipids with a long alkyl chain, such as mycobacterial mycolic acid.

γδ T cells are activated by conserved stress-induced ligands, enabling them to rapidly produce cytokines that regulate pathogen clearance, inflammation, and tissue homeostasis in response to tissue stress.[21]

Antigen recognition by γδ TCRs resembles recognition of intact antigens by antibodies more closely than recognition of pMHC by αβ TCR.[22] γδ TCRs can recognize protein antigens, such as nonclassic MHC molecules and viral glycoproteins, as well as small, phosphate- or amine-containing compounds, such as pyrophosphomonoesters from mycobacteria and alkylamines.

Binding to non-peptide antigens plays an important role in the biology of γδ T cells. About 5% of peripheral blood T cells bear γδ TCRs, and most of these are encoded by Vγ9 JγP and Vδ2 gene segments. (In an alternative nomenclature, Vγ9 is known as *Vγ2* and JγP as *Jγ1.2.* See the IMGT database at http://www.imgt.org.) These Vγ9 JγPVδ2 TCRs recognize nonpeptide pyrophosphate- or amine-containing antigens, such as pyrophosphomonoesters from mycobacteria or isobutylamine from various sources. Other common naturally occurring small phosphorylated metabolites that stimulate γδ T cells include 2,3-diphosphoglyceric acid, glycerol-3-phosphoric acid, xylose-1-phosphate, and ribose-1-phosphate. In addition to mycobacteria, Vγ9JγPVδ2 T-cell populations are seen to expand in response to listeriosis, ehrlichiosis, leishmaniasis, brucellosis, salmonellosis, mumps meningitis, malaria, and toxoplasmosis.

Superantigens

SAgs are a special class of TCR ligands that have the ability to activate large fractions (5–20%) of the T-cell population. Activation requires simultaneous interaction between the SAg, the TCR Vβ domain, and a MHC class II molecule on the surface of an APC.[23]

Unlike conventional antigens, SAgs do not require processing to allow them to bind class II molecules or activate T cells. Instead of binding to the peptide antigen-binding groove, SAgs interact with polymorphic residues on the periphery of the class II molecule. Rather than binding to TCR β CDR3 residues, SAg can interact with polymorphic residues in CDR1, CDR2, and HV4. Soluble TCR β chains can also bind the appropriate SAg in the absence of a TCR α chain. As a consequence, although the SAg links the TCR to the MHC, the T-cell responses are not "MHC restricted" in the conventional sense, since a T cell with the appropriate Vβ will respond to a SAg bound to a variety of polymorphic class II molecules.

IMMUNOGLOBULIN GENE ORGANIZATION

Each of the component chains of Igs and TCRs is encoded by a separate multigene family.[24,25] The paradox of variability in the V region in conjunction with a nearly invariable constant region was resolved when it was shown that Ig V and C domains are encoded by independent elements, or gene segments, within each gene family. As a result, several gene elements are used to

KEY CONCEPTS

Features Common to Immunoglobulin (Ig) and T-Cell Receptor (TCR) Genes

- Ig and TCR variable domains are created by site-specific V(D)J recombination.
- Starting with a small set of individual gene segments, combinatorial gene segment rearrangement, combinatorial association of heavy and light chains, and mechanisms of junctional diversity generate a broad repertoire of antigen-binding structures.
- Each receptor is assembled in a stepwise fashion
    - *Immunoglobulins:* $D_H \rightarrow J_H$; $V_H \rightarrow D_H J_H$; cytoplasmic μ chain expression; Vκ→Jκ and, if needed, Vλ→Jλ; surface IgM expression.
    - *TCRαβ:* Dβ→Jβ; Vβ→DβJβ; cytoplasmic β-chain expression; Vα→Jα; surface αβ TCR expression.
- Complementarity determining regions (CDRs) 1 and 2 begin with exclusively germline sequence.
- CDR3 is created by the (V[D]J) joining reaction and often includes non-germline N nucleotides between the V and the D, and between the D and the J.
    - Thus CDR-H3, CDR-B3, and CDR-D3 are the most variable components of IgM, TCRαβ, and TCRγδ, respectively.
- The antigen-binding site is a product of a nested gradient of diversity. Conserved framework regions surround CDR1 and CDR2, which, in turn, surround the paired CDR3 intervals that form the center of the antigen-binding site.
- The variability of the Ig and TCR repertoires is restricted during perinatal life, limiting the immune response of the infant.

KEY CONCEPTS

Features Specific to Immunoglobulin (Ig) Genes

- Variable domain somatic hypermutation (SHM) permits affinity maturation, which further diversifies the B-cell repertoire.
- Class-switch recombination (CSR) allows replacement of an upstream C domain by a downstream one, altering effector function while maintaining antigen specificity.

encode a single polypeptide chain. For example, κ constant domains are encoded by a single Cκ exon in the κ locus on chromosome 2, whereas κ variable domains represent the joined product of Vκ and Jκ gene segments (Fig. 4.4).

V_L gene segments typically contain their own promoter, a leader exon, an intervening intron of ≈ 100 nucleotides, an exon that encodes the first three framework regions (FR 1, 2, and 3), the first two complementarity determining regions in their entirety, the amino terminal portion of CDR3, and a recombination signal sequence (RSS). A J_L (J for joining) gene segment begins with its own recombination signal, the remaining portion of CDR 3, and the complete FR 4 (see Fig. 4.2).

(Use of the same abbreviation—that is, "V"—for the complete variable domain of an Ig peptide chain as well as the gene segment that encodes only a portion of that same variable domain is the result of historic precedent. It is unfortunate that one must depend on the context of the surrounding text to determine which V region of the antibody is being discussed. The same holds true for the use of "J" to represent both the J gene segment and the J joining protein.)

The creation of a V domain is directed by the RSSs that flank the rearranging gene segments.[26] Each RSS contains a strongly conserved seven base-pair, or heptamer, sequence (e.g., CACAGTG) that is separated from a less well-conserved nine bp, or nonamer, sequence (e.g., ACAAAACCC) by either a 12- or 23- bp spacer. For example, Vκ gene segments have a 12-bp spacer and Jκ elements have a 23-bp spacer. These spacers place the heptamer and nonamer sequences on the same side of the DNA molecule, separated by either one or two turns of the DNA helix. A one-turn

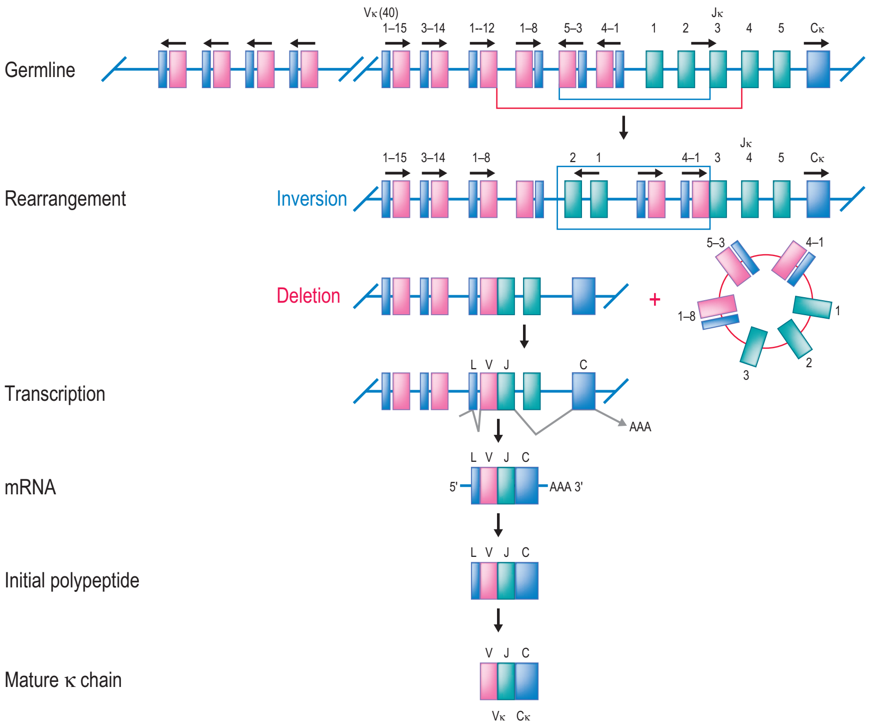


FIG 4.4 **Rearrangement Events in the Human κ Locus.** V, variable region; J, joining region; C, constant region of the κ light chain; mRNA, messenger RNA. See text for further description.

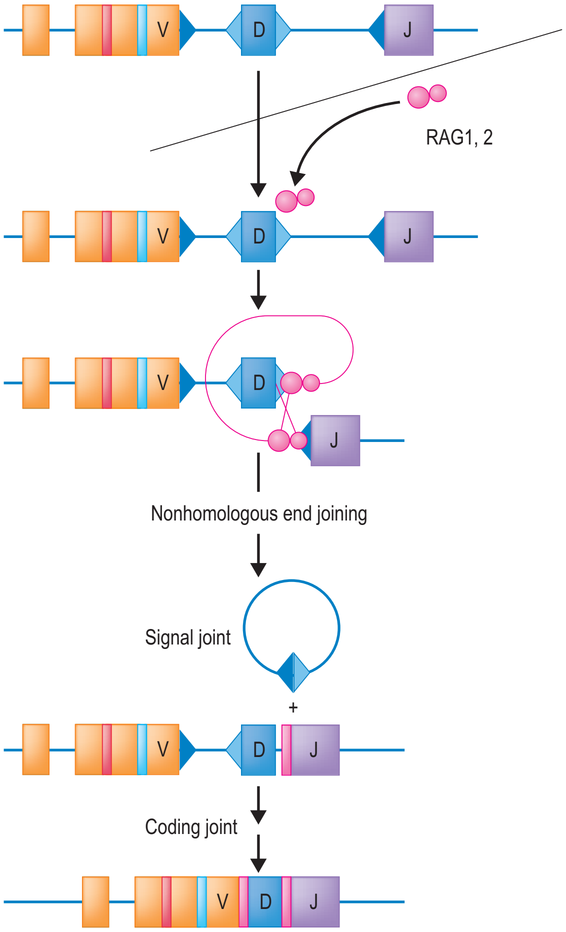


FIG 4.5 **VDJ Recombination.** Lymphoid-specific recombinase activating gene (RAG)-1 and RAG-2 bind to the recombination signal sequences (RSSs) flanking V, D, or J gene segments, juxtapose the sequences, and introduce precise cuts adjacent to the RSS. Components of the nonhomologous end joining repair pathway subsequently unite the cut RSS to form a signal joint, and the coding sequences of the rearranging gene segments to form a coding joint.

RSS (12-bp spacer) will preferentially recognize a two-turn signal sequence (23-bp spacer). This helps prevent wasteful V-V or J-J rearrangements.

Initiation of the V(D)J recombination reaction requires recombination activating genes 1 and 2 (RAG-1 and RAG-2). These genes are expressed only in developing lymphocytes.[27] The gene products RAG-1 and RAG-2 act by precisely introducing a DNA double-strand break (DSB) between the terminus of the rearranging gene segment and its adjacent RSS (Fig. 4.5). These breaks are then repaired by ubiquitously expressed components of a DNA repair process that is known as *nonhomologous end joining* (NHEJ). Lack-of-function mutations in NHEJ proteins yields susceptibility to DNA damage in all cells of the body and can lead to an SCID phenotype (Chapter 35).

The NHEJ process creates precise joins between the RSS ends and imprecise joins of the coding ends. TdT, which is expressed only in lymphocytes, adds non-germline–encoded nucleotides (N nucleotides) to the coding ends of the recombination product.

Lymphoid-specific expression of RAG-1 and RAG-2 limits V(D)J recombination to B and T lymphocytes. To ensure that TCR genes are rearranged to completion only in T cells and Ig genes are rearranged to completion only in B cells, V(D)J recombination is further regulated by limiting the accessibility of the appropriate gene segments to the specific lineage as well as to the specific stage of development. For example, H chain genes are typically assembled before L chain genes are.

The RAG-1 and RAG-2 recombinases cooperatively associate with 12-bp and 23-bp RSSs and their flanking coding gene segments to form a synaptic complex. Typically, the initial event will be recognition of the nonamer sequence of a 12-bp spacer RSS by RAG-1, which appears to function as the catalytic component of the recombinase. RAG-1 binding to the heptamer provides specificity. RAG-2 does not bind DNA independently but does make contact with the heptamer when in a synaptic complex with RAG-1. Binding of a second RAG-1 and RAG-2 complex to the 23-bp, two-turn RSS permits the interaction of the two synaptic complexes to form what is known as a *paired complex*. Creation of this paired complex is facilitated by the actions of the DNA-bending proteins HMGB1 and HMGB2 and by the presence of a divalent metal ion.

After paired complex assembly, the RAG proteins single-strand cut the DNA at the heptamer sequence. The 3' OH of the coding sequence ligates to 5' phosphate and creates a hairpin loop. The clean-cut ends of the signal sequences enable formation of precise signal joints. However, the hairpin junction created at the coding ends must be resolved by renicking the DNA, usually within 4–5 nucleotides from the end of the hairpin. This forms a 3' overhang that is amenable to further diversification. It can be filled in via DNA polymerases, nibbled back, or serve as a substrate for TdT-catalyzed N addition. DNA polymerase μ, which shares homology with TdT, appears to play a role in maintaining the integrity of the terminus of the coding sequence.

The cut ends of the coding sequence are then repaired by the NHEJ proteins. NHEJ proteins involved in V(D)J recombination include Ku70, Ku80, DNA-PKcs, Artemis, XRCC4, XLF (Cernunnos), and ligase 4.

Ku70 and Ku80 form a heterodimer (Ku) that directly associates with DNA double-strand breaks to protect the DNA ends from degradation, permit juxtaposition of the ends to facilitate coding end ligation, and help recruit other members of the repair complex. The DNA protein kinase catalytic subunit (DNA-PKcs) phosphorylates Artemis, inducing an endonuclease activity that plays a role in the opening of the coding joint hairpin. Thus absence of DNA-PKcs or Artemis inhibits proper coding joint formation. Signal joint formation is normal in Artemis deficiency, but it is impaired in the absence of DNA-PKcs. Finally, XRCC4, XLF, and ligase 4 help rejoin the ends of the broken DNA.

Depending on the transcriptional orientation of the rearranging gene segments, recombination can result in either inversion or deletion of the intervening DNA (see Fig. 4.3). The products of inversion remain in the DNA of the cell, whereas deletion leads to the loss of the intervening DNA. The increased proximity of the V promoter to the C domain enhancers promotes the subsequent transcription of the Ig gene product.

There is a price to the use of V(D)J recombination to create antigen receptor diversity. Aberrant recombination in nonreceptor genes can create deleterious genomic rearrangements that promote B-cell and T-cell neoplasias.[26] For example, deletional recombination at the SIL/SCL locus and in *Notch1*, *Izkf1*, *PTEN*, and other

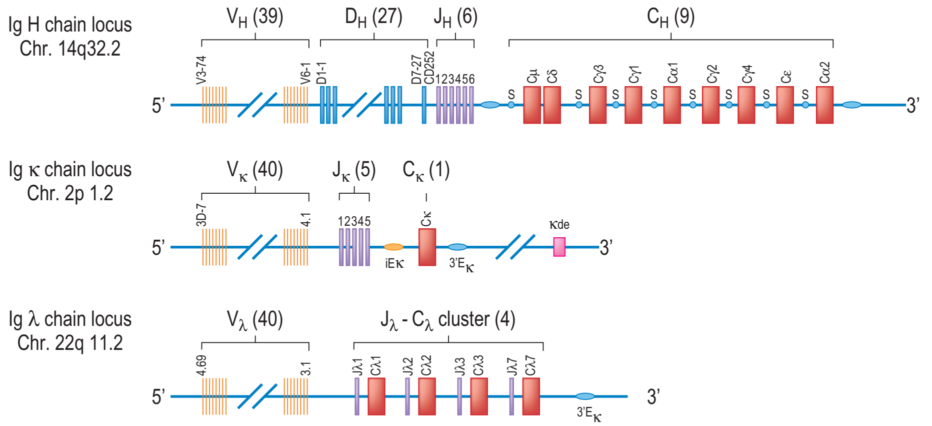


FIG 4.6 **Chromosomal Organization of the Immunoglobulin H (IgH), κ, and λ Gene Clusters.** The typical numbers of functional gene segments are shown. The κ gene cluster includes a κ deleting element that can rearrange to sequences upstream of Cκ in cells that express λ chains, reducing the likelihood of dual κ and λ light chain expression. These maps are not drawn to scale.

critical genes appear to be major drivers of lymphoid neoplasms in humans and in mice.

The κ Locus

The κ locus is located on chromosome 2p11.2. It contains 5 Jκ and 75 Vκ gene segments upstream of Cκ (Fig. 4.6). The Vκ gene segments can be grouped into six different families of various sizes.[28] Each family comprises gene segments that share extensive sequence and structural similarity.[29]

One-third of the Vκ gene segments contain frameshift mutations or stop codons that preclude them from forming functional protein, and of the remaining sequences, <30 of the Vκ gene segments have actually been found in functional Igs. Each of these active Vκ gene segments has the potential to rearrange to any one of the 5 Jκ elements, generating a potential repertoire of >140 distinct VJ combinations. Even more diversity is created at the site of gene segment joining. The terminus of each rearranging gene segment can undergo a loss of 1–5 nucleotides during the recombination process. In humans, but not in mice, N-addition can either replace some or all of the lost nucleotides or can be inserted in addition to the original germline sequence. Each codon created by N-addition increases the potential diversity of the repertoire 20-fold. Thus the focus for the diversity of the κ repertoire lies in the VJ junction that defines CDR-L3.

The λ Locus

The λ locus, on chromosome 22q11.2, contains four functional Cλ exons, each of which is associated with its own Jλ (see Fig. 4.6). The Vλ genes are arranged in three distinct clusters, each containing members of different Vλ families.[30] Depending on the individual haplotype, there are approximately 30–36 potentially functional Vλ gene segments and an equal number of pseudogenes.

In addition to normal κ and λ peptides, H chains can also form a complex with unconventional λ light chains, known as surrogate or pseudo light chains (SLC). The genes encoding the SLC proteins, λ14.1 (λ5) and VpreB, are located within the λ light-chain locus on chromosome 22 and are restricted in expression to discreet B cell developmental stages.[31] Together, these two genes create a product with considerable homology to conventional λ light chains. The λ14.1 gene contains Jλ and Cλ-like sequences and the VpreB gene includes a Vλ-like sequence. A critical difference between these unconventional ψLC genes and other L chains is that rearrangement of λ14.1 and VpreB is not required for SLC expression.

The H Chain Locus

The H chain locus, on chromosome 14q32.33, is considerably more complex than the κ and λ loci. There are ≈ 80 V_H gene segments near the telomere of the long arm of chromosome 14.[32] Of these, approximately 39 are functional (the number varies by haplotype) and can be grouped into seven different families of related gene segments. Adjacent to the most centromeric V_H, V6-1, are 27 D_H (D for diversity) gene segments (see Fig. 4.6) and 6 J_H gene segments. Each V_H and J_H gene segment is associated with a two-turn RSS, which prevents direct V → J joining. A pair of one-turn RSSs flanks each D_H. Recombination begins with the joining of a D_H to a J_H gene segment, followed by the joining of a V_H element to the amino terminal end of the DJ intermediate. The V_H gene segment contains FR1, 2, and 3, CDR1 and 2, and the amino terminal portion of CDR3; the D_H gene segment forms the middle of CDR3; and the J_H element contains the carboxy terminus of CDR3 and FR4 in its entirety (see Fig. 4.1). For example, random assortment of one of 50 active V_H and one of 27 D_H with one of the 6 J_H gene segments can generate up to 10^4 different VDJ combinations (Fig. 4.7).

Although combinatorial joining of individual V, D, and J gene segments maximizes germline-encoded diversity, the major source of variation in the preimmune repertoire is focused on the CDR-H3 interval, which is created by VDJ joining (see Fig. 4.7). First, D_H gene segments can rearrange by either inversion or deletion and thus have the potential to be read backward as well as forward. Each D_H can be spliced and translated in each of the three potential reading frames. Thus each D_H gene segment has the potential to encode six different peptide fragments. Second, the terminus of each rearranging gene segment can undergo a loss of ≥1 nucleotides during the recombination process. Third, the rearrangement process proceeds through a step that creates a hairpin ligation between the 5' and 3' termini of the rearranging gene segment. Nicking to resolve the hairpin structure leaves a 3' overhang that creates a palindromic extension, termed *P*

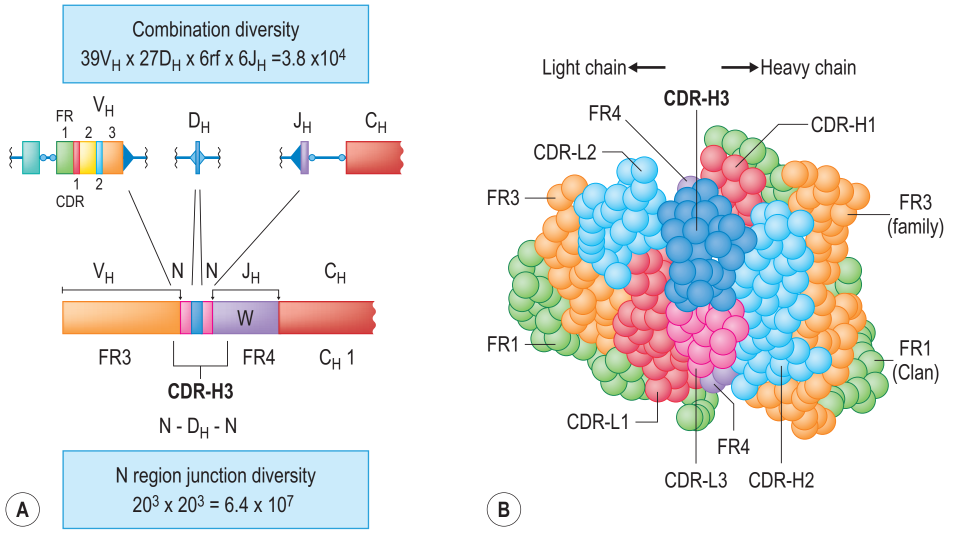

Combination diversity
$39V_H \times 27D_H \times 6rf \times 6J_H = 3.8 \times 10^4$
FR
1 2 3
V_H
1 2
CDR
D_H
J_H
C_H
V_H
N
N
J_H
C_H
W
FR3
FR4
$C_H 1$
CDR-H3
$N - D_H - N$
N region junction diversity
$20^3 \times 20^3 = 6.4 \times 10^7$
A
Light chain
Heavy chain
CDR-H3
FR4
CDR-H1
CDR-L2
FR3
FR3 (family)
FR1
FR1 (Clan)
CDR-L1
FR4
CDR-L3
CDR-H2
B

FIG 4.7 **The Antigen-Binding Site Is the Product of a Nested Gradient of Diversity.** (**A**) VDJ rearrangement yields 38 thousand different combinations. The CDR-H3 sequence contains both germline V, D, and J sequence and non-germline–encoded N-nucleotides. The addition of nine N-nucleotides on either side of the D gene segment yields 64 million different combinations. (**B**) The antigen-binding site is created by the juxtaposition of the three complementarity determining regions (CDRs) of the H chain and the three CDRs of the light chain. The view is looking into the binding site as an antigen would see the CDRs. The V_H domain is on the right side. The central location of CDR-H3, which, because of the N addition, is the focus for repertoire diversity, is readily apparent.

junction. Fourth, non–germline-encoded nucleotides (N regions or N additions) can be used to replace or add to the original germline sequence. Every codon that is added by N region addition increases the potential diversity of the repertoire 20-fold. N regions can be inserted both between the V and the D and between the D and the J. Together, the imprecision of the joining process and variation in the extent of *N* addition permits generation of CDR-H3's of varying length and structure. As a result, more than 10^{10} different H chain VDJ junctions, or CDR-H3's, can be generated at the time of gene segment rearrangement. Together, somatic variation in CDR3, combinatorial rearrangement of individual gene segments, and combinatorial association between different L and H chains yields a potential preimmune antibody repertoire of greater than 10^{16} different Igs.

Class-Switch Recombination

Located downstream of the VDJ loci are nine functional C_H gene segments (see Fig. 4.7). Each C_H contains a series of exons, each encoding a separate domain, hinge, or terminus. All C_H genes can undergo alternative splicing to generate two different types of carboxy termini: either a membrane terminus that anchors the Ig on the B lymphocyte surface or a secreted terminus that occurs in the soluble form of the Ig. With the exception of $C_H1\delta$, each C_H1 constant region is preceded by both an exon that cannot be translated (an I exon) and a region of repetitive DNA termed the *switch* (S). Through recombination between the Cμ switch region and one of the switch regions of the seven other H chain constant regions (a process termed *class switching* or *class-switch recombination* [CSR]), the same VDJ heavy chain variable domain can be juxtaposed to any of the H chain classes (Fig. 4.8).[33] Thus

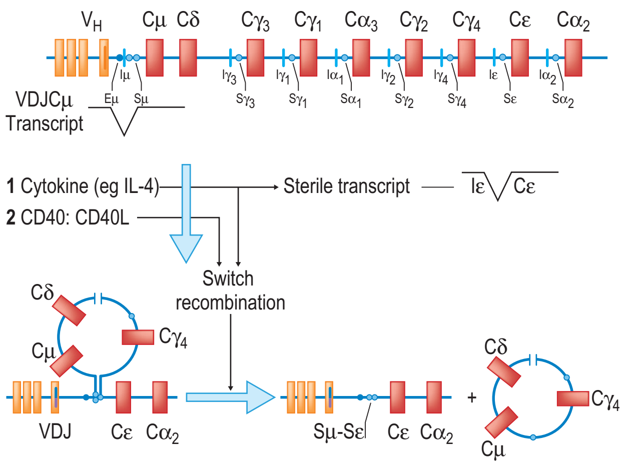

V_H
Cμ
Cδ
$C\gamma_3$
$C\gamma_1$
$C\alpha_3$
$C\gamma_2$
$C\gamma_4$
Cε
$C\alpha_2$
Iμ
$I\gamma_3$
$I\gamma_1$
$I\alpha_1$
$I\gamma_2$
$I\gamma_4$
Iε
$I\alpha_2$
Eμ
Sμ
$S\gamma_3$
$S\gamma_1$
$S\alpha_1$
$S\gamma_2$
$S\gamma_4$
Sε
$S\alpha_2$
VDJCμ Transcript
1 Cytokine (eg IL-4)
2 CD40: CD40L
Sterile transcript
Iε Cε
Switch recombination
Cδ
Cμ
$C\gamma_4$
VDJ
Cε
$C\alpha_2$
Sμ-Sε
Cε
$C\alpha_2$
+
Cδ
$C\gamma_4$
Cμ

FIG 4.8 **Immunoglobulin H (IgH) Chain Class Switching.** The molecular events involved in switching from expression of one class of Ig to another are depicted. At the top is the gene organization during μ chain synthesis. At the bottom, a class-switch recombination event has resulted in the deletion of the intervening DNA. Exposure to the appropriate cytokine or T cell–B cell interaction through the CD40–CD40L pathway results in activation of the I exon that yields a sterile epsilon transcript (Iε-Cε) (Chapter 7). The CD40–CD40L interaction is necessary for the subsequent replacement of Cμ by another constant gene (in this case, Cε). The S loci indicate switch-specific recombination signals.

the system can tailor both the receptor and the effector ends of the antibody molecule to meet a specific need.

Somatic Hypermutation

A final mechanism of Ig diversity is engaged only after exposure to antigen.[33] With T-cell help, the variable domain genes of germinal center lymphocytes undergo *somatic hypermutation* (SHM) at a rate of up to 10^{-3} changes per base pair per cell cycle. SHM is correlated with transcription of the locus, and recent studies have suggested that at least two separate mechanisms are involved. The first mechanism targets mutation hot spots with the RGYW (purine/G/pyrimidine/A) motif, and the second mechanism incorporates an error-prone DNA synthesis that can lead to a nucleotide mismatch between the original template and the mutated DNA strand. SHM allows affinity maturation of the antibody repertoire in response to repeated immunization or exposure to antigen, as B cells bearing receptors that have mutated to higher affinity for the cognate antigenic epitope are preferentially stimulated to proliferate, especially under conditions of limiting antigen concentration.

Activation-Induced Cytidine Deaminase

Activation-induced cytidine deaminase (AID) plays a key role in both CSR and SHM.[33] AID is a single-strand DNA (ssDNA) cytidine deaminase that can be expressed in activated germinal center B cells. Both SHM and CSR require transcription. Transcription helps target AID to the requisite chromosomal location and contributes to formation of requisite ssDNA substrates. For example, transcription of an Ig V domain or of the switch region upstream of the C_H1 domain opens the DNA helix to generate ssDNA that can then be deaminated by AID to form mismatched dU/dG DNA bps. Both CSR and SHM then co-opt the activities of normal cellular base excision repair (BER) and mismatch repair (MMR) to convert AID cytidine deamination lesions to mutations and/or double-strand breaks. The BER protein uracil-DNA glycosylase (UNG) removes the mismatched dU base, creating an abasic site. Cleavage of the DNA backbone at this abasic site by an apurinic/apyrimidinic (AP) endonuclease generates a ssDNA nick adjacent to the abasic site. The nick is then processed to a single-nucleotide gap. The gap is filled in by DNA polymerase β and then sealed by DNA ligase 1 or DNA ligase 3. The MMR proteins MSH2 and MSH6 can also bind and process the dU:dG mismatch. Deficiencies of AID, UNG underlie some forms of the hyper-IgM syndrome (Chapter 34). UNG and MMR double deficiency ablates CSR. It also eliminates both C/G transversion mutations and spreading of mutations, leaving only C/G transition mutations.

The benefits of diversity created by AID are balanced by the tendency of AID to target non-Ig genes. AID can create clusters of mutations in a number of genes, including *BCL6, CD95, CD79A, CD79B, PIM1, MYC, RHOH,* and paired box 5 (*PAX5*).[34] The process is termed *kataegis.* These mutations clusters can contribute to the development of lymphoproliferative disorders.

Diversity and Constraints on Immunoglobulin Sequence

In theory, combinatorial rearrangement of V(D)J gene segments, combinatorial association of H and L chains, flexibility in the site of gene segment joining, N region addition, P junctions, somatic hypermutation, and class switching can create an antibody repertoire whose diversity is limited only by the total number of B cells in circulation at any one time. In practice, constraints and biases on both the structure and sequence of the antibody repertoire are apparent.

The representation of individual V gene elements is nonrandom. Among Vκ and V_H elements, half the potentially functional V gene elements contribute minimally to the expressed repertoire. Among Vλ elements, these restrictions are even greater, with three gene segments contributing to half the expressed repertoire.

Particular patterns of amino acid composition in the sequences of the V domains create predictable canonical structures for several of the hypervariable regions. In κ chains, CDR2 is found in a single canonical structure, whereas four structures are possible for CDR1.[35] In the H chain, most germline CDR1 and CDR2 elements encode one of three or one of five distinct canonical structures, respectively.[36] Preservation of these key amino acids during affinity maturation tends to maintain the canonical structure of CDR1 and CDR2 even while they are undergoing somatic hypermutation.[37]

The enhanced sequence diversity of the CDR3 region is mirrored by its structural diversity. Few canonical structures have been defined for the H chain CDR3, and even in κ chains, 30% of the L chain CDR3 can be quite variable. However, at the sequence level, there is a preference for tyrosine and glycine residues and a bias against the use of highly charged or hydrophobic amino acids in the H chain CDR3, which reflects preferential use of only one of the six potential D_H reading frames, natural selection of reading frame content, and selection during development.[38]

The TCR αδ Chain Locus

The α and δ loci are located on chromosome 14q11-12. This region is unusual in that the gene segments encoding the two different TCR chains are actually intermixed (Fig. 4.9). There are 38–40 Vα, 5 Vα/Vδ, no Dα, and 50 Jα functional gene segments, as well as one Cα gene.[39]

The δ locus lies between the Vα and Jα gene segments. There are three committed Vδ, 5 Vα/Vδ, 3 Dδ, and 3 Jδ gene segments, as well as one Cδ gene. Vδ3 lies 3' of Cδ, and thus must rearrange by inversion. Although V region use by α and δ chains is largely independent of one another, this unusual gene organization is accompanied by sharing of 5 V gene segments. For example, Vδ1 can rearrange either to Dδ/Jδ or to Jα elements and thus can serve as the V region for both γδ and αβ TCRs.

In the large majority of αβ T cells analyzed, the α chain on both chromosomes was rearranged. This occurs by the rearrangement of the 5' RSS δRec to a pseudo-J segment, ψJα, at the beginning of the Jα cluster (see Fig. 4.9) as well as by the subsequent rearrangement of Vα to Jα on both chromosomes. Both types of rearrangement delete all of the Dδ, Jδ, and Cδ genes, thus preventing coexpression of αβ and γδ TCRs.

The TCR β Chain Locus

The β locus is positioned at chromosome 7q35.44 It contains 40–48 functional Vβ genes, two Dβ, two Jβ clusters, each containing six or seven gene segments, and two Cβ genes (see Fig. 4.9). There is one Vβ immediately downstream of Cβ2, which rearranges by inversion. Each Cβ is preceded by its own Dβ–Jβ cluster. There is no apparent preference for Vβ gene rearrangement to either Dβ–Jβ cluster. Dβ1 can rearrange to the Jβ1 cluster or the Dβ2–Jβ2 cluster. Dβ2 can only rearrange to Jβ2 gene segments. The two Cβ segments differ by only six amino acids and are functionally indistinguishable from each other.

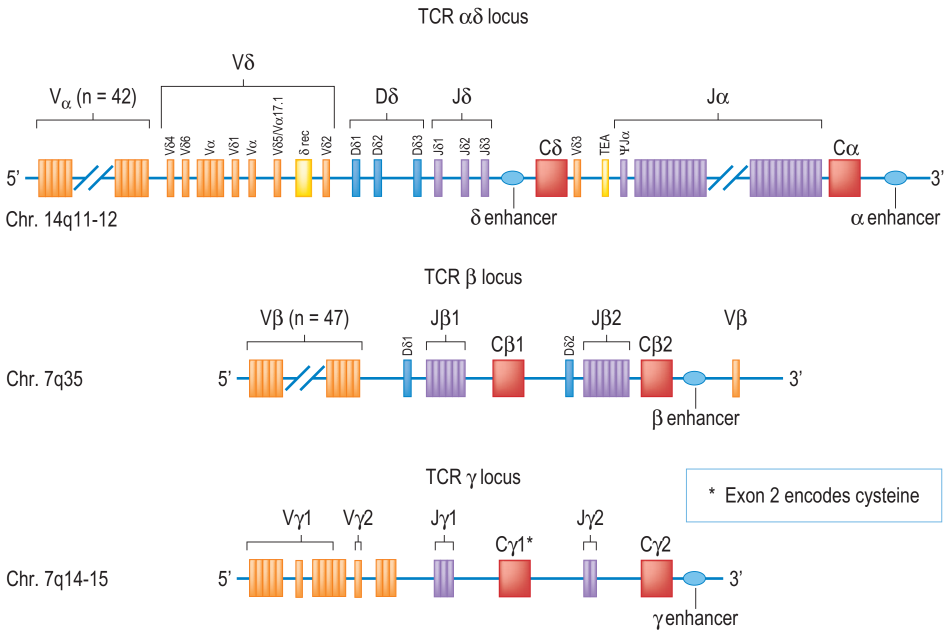


FIG 4.9 Chromosomal Organization of the T-Cell Receptor (TCR) αδ, β, and γ Gene Clusters. Typical numbers of functional gene segments are shown. These maps are not drawn to scale.

The TCR γ Chain Locus

The γ locus is located at chromosome 7p14-15.[39] There are 4–6 functional Vγ region segments intermixed with pseudogenes, no Dγ, and two J clusters with a total of 5 J segments. Each J cluster is 5' to its C region (see Fig. 4.9). The Vγ segments have been divided into six families, although only Vγ1 (nine members, five of them functional) and Vγ2 (one member) encode functional proteins. The number of Cγ gene exons varies: Cγ1 has three, whereas there are two alleles of Cγ2 that have four and five, respectively. The first Cγ exon encodes most of the extracellular portion of this region. The last Cγ exon encodes the intracytoplasmic portion of the molecule. The middle exon(s) (one for Cγ1, two or three for Cγ2) encode the connecting piece, which does (Cγ1), or does not (Cγ2), include a cysteine. Since this cysteine can form a disulfide bond with another cysteine in the δ chain, TCRs using Cγ1 contain a covalently linked γ–δ pair, whereas TCRs using Cγ2 do not.

The nomenclature of the human γ locus differs among laboratories and reports and is extensively cross-referenced on the IMGT website (http://www.imgt.org).

Allelic Exclusion

Because of the inherently imprecise nature of coding joints, only one in three V(D)J Ig or TCR rearrangements will be in-frame and capable of creating a functional protein.[40] Theoretically, one in nine cells might be expected to express two different Ig or TCR chains. However, almost all B cells express the functional products of only one IgH allele and one IgL allele, and mature αβ T cells express only one functional TCRβ gene. The process of limiting the number of receptors expressed by an individual cell is known as *allelic exclusion.*[41]

The mechanisms that ensure monoallelic expression are most commonly regulated at the level of gene rearrangement. Mechanisms that have been shown to contribute to allelic exclusion include asynchronous replication of the two alleles, with rearrangement occurring at the allele that replicates early; localization of the active allele to a more central, euchromatic region of the nucleus; and DNA demethylation of the active allele. Once a functional V domain has been generated, rearrangement terminates with the expression of a membrane-bound Ig or TCR product capable of transducing a signal. In preB cells, a functional μ H chain associates with the surrogate light chain to form the preBCR. Similarly, in developing T cell progenitors, a productive TCR β chain associates with preTα to form the preTCR. These preliminary antigen receptors signal to shut down RAG expression, promote cell division, and limit the accessibility of the IgH and TCRβ loci to further rearrangement while promoting the accessibility of the IgL and TCRα loci, respectively.

In preB cells, the κ locus is typically the first to rearrange, with λ rearrangement primarily occurring in cells that have failed to produce a proper κ chain. Surface expression of an acceptable membrane-bound IgM BCR invokes the mechanism of allelic exclusion among the L chain loci, termed *isotypic exclusion,* and promotes further maturation of the B cell.

Productive TCRα rearrangement in $CD4^+CD8^+$ T-cell progenitors allows the expression of a functional TCR αβ heterodimer (Chapter 9). Unlike IgH and TCRβ, TCRα does not undergo allelic exclusion at the level of gene rearrangement. Instead, in cells that express two functional TCRα alleles, one of the two alleles tends to preferentially pair with the one functional TCRβ chain. This is termed *phenotypic allelic exclusion.*

Allelic exclusion can be overcome by selection pressures. Developing lymphocytes that express self-reactive antigen receptors can downregulate IgH or TCR expression and reactivate gene rearrangement to replace one of the two offending chains. This process, termed *receptor editing,*[42] occurs most often in the IgL or TCRα loci, whose gene structures lend themselves to repeated rearrangement. Less commonly, the V_H in the H chain can be replaced by means of rearrangement to a cryptic RSS located at the terminus of the V_H gene segment.

KEY CONCEPTS

B Cell Receptor (BCR) and Coreceptors

- The BCR–antigen complex consists of a membrane-bound immunoglobulin (mIg) that is responsible for antigen recognition and an Igα/β heterodimer that is responsible for transducing the recognition signal into the cell.
- BCR engagement leads to the phosphorylation of tyrosines in the Igα/β immunoreceptor tyrosine-based activation motif (ITAM) motifs. This signal is then transmitted to one or more other intracellular signaling pathways.
- Recognition of antigen by B lymphocytes can also involve binding of antigen complexed with C3d and IgG to additional B-cell coreceptors.
- Binding of complexed antigen by individual coreceptors can lead to either positive or negative signals, each of which can influence the ultimate outcome of an antigen–B lymphocyte interaction.
- Deficiency of the components of the BCR antigen complex impairs B cell development and can lead to agammaglobulinemia.

B CELL RECEPTOR COMPLEX: STRUCTURE AND FUNCTION

Although the ability of surface Ig to recognize antigen was appreciated very early, the mechanism by which membrane immunoglobulin (mIg) transmitted an antigen recognition event to the cell took longer to understand. Specifically, as the predominant isotypes expressed on the surface of mature B cells, mIgM, and mIgD contain only three amino acid residues exposed to the cytoplasm, it was thought unlikely that these Ig heavy chains could function as signal transduction molecules by themselves. This presumption was eventually proved correct when it was shown that all membrane Ig isotypes associated noncovalently with a heterodimeric complex consisting of two transmembrane proteins, Igα (CD79a) and Igβ (CD79b), each of which is capable of transducing signals into the cell (Table 4.3).

TABLE 4.3 **The B-Cell Receptor (BCR) and Its Coreceptor Molecules**

Molecule	M_r	Chromosome	Function
BCR			
mIgM (μ_2L_2)	180 000	14 (IgH; 14q.32) 2 (Igκ; 2p12) 22 (Igλ; 22q11.2)	Antigen recognition
Ig-α (CD79a)	47 000	19 (19q13.2)	Signal transducer
Ig-β (CD79b)	37 000	17 (17q23)	Signal transducer
Coreceptors			
CD21	140 000	1 (1q32)	Activating coreceptor Ligand for C3d, EBV, CD23
CD19	95 000	16 (16p11.2)	Activating coreceptor Signal transducer
FcγRIIB (CD32)	40 000	1 (1q23-24)	Inhibitory coreceptor Low affinity receptor for IgG
CD22	140 000	19 (19q13.1)	Inhibitory coreceptor Adhesion molecule Signal transducer

Membrane-Bound Immunoglobulin

Igs mediate their effector functions as secreted products of plasma cells. However, as membrane-bound structures on mature B cells, Igs serve as the antigen-recognition component of the BCR complex. Although all Ig classes can be expressed at the cell surface, the vast majority of circulating mature B cells coexpress membrane-bound IgM and IgD. Appropriate activation of a naïve IgM- and IgD-expressing B cell leads to plasma cell differentiation and antibody secretion. The membrane-bound forms of IgM and IgD are the product of alternative splicing of the Ig transcript at the 3', or carboxy terminus, of the heavy chain (Fig. 4.10). The two membrane exons encode the transmembrane hydrophobic stretch of amino acids and an evolutionarily conserved cytoplasmic tail terminating in lysine, valine, and lysine.

Signal Transduction and the Ig-α/β (CD79a/CD79b) Heterodimer

The heterodimeric signal transduction component of the BCR complex that associates with mIg has been designated CD79. It is composed of an Igα (CD79a) and Igβ (CD79b) heterodimer. CD79 is responsible for transporting mIg to the cell surface and for transducing BCR signals into the cell.[43,44]

CD79a/Igα is encoded by *CD79a/MB-1* (chromosome 19q13.2) as a 226 amino acid glycoprotein of approximately 47 kDa. The exact molecular weight depends on the extent of glycosylation. *CD79b/B29* (chromosome 17q23) encodes CD79b/Igβ, which is a 229-amino acid glycoprotein of approximately 37 kDa. *CD79a* and *CD79b* share an exon–intron structure, which is similar to that of the genes that encode the CD3 TCR coreceptor molecules. These similarities suggest that both BCR and TCR coreceptors are the progeny of a common ancestral gene. Igα and Igβ both contain a single IgSF Ig domain (111 residue C-type for Igα and 129 residue V-type for Igβ). Each also contains a highly conserved transmembrane domain and a 61-(Igα) or 48-(Igβ) amino acid cytoplasmic tail that also exhibits striking amino acid evolutionary conservation.

Igα and Igβ are expressed by the earliest committed B-cell progenitors prior to Igμ H chain rearrangement. The CD79 heterodimer has been observed on the surface of early B-cell progenitors in the absence of Igμ, although neither protein is required for progenitors to commit to the B-cell lineage.[45] Later

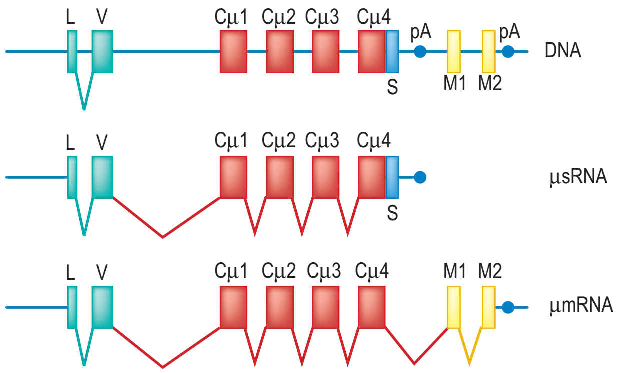


FIG 4.10 **Membrane and Secretory Immunoglobulin M (IgM) Are Created by Alternative Splicing.** Alternative splicing of the Cm carboxy-terminal exons results in messenger RNA (mRNA) transcripts encoding either secreted IgM (μ_s RNA) or membrane-bound IgM (μ_m RNA).

in development, Igα and Igβ are coexpressed together with Igs of all isotypes on the surface of B cells as a mature BCR complex.[43] The CD79 proteins are specific to the B lineage and are expressed throughout B lymphopoiesis. This has led to their use as markers for the identification of B-cell neoplasms.[46,47]

The signaling capacity of both Igα and Igβ resides within an immunoreceptor tyrosine-based activation motif (ITAM) that has the consensus sequence of D/IxxYxxL(x)7YxxL, where x is any amino acid. Similar ITAMs are also found within the cytoplasmic domain of the molecules that associate with, and signal for, the T-cell antigen receptor (CD3) and certain Fc receptors (Chapter 15). The phosphorylation of both tyrosines in both Igα/β ITAMs is considered an obligate initial step in the propagation of antigen receptor engagement to the cell nucleus.[44,48]

Tyrosine-phosphorylated ITAMs serve as efficient binding sites for Src homology 2 (SH2) domains, which are present within a large number of cytosolic signaling molecules. Whether Igα and Igβ make qualitatively different contributions toward BCR signaling or are functionally redundant remains unclear, as evidence exists to support both views. Moreover, the high degree of evolutionary conservation within the non-ITAM portion of the cytoplasmic domains suggests additional, as yet uncharacterized, signaling roles for the cytoplasmic tails of these molecules over and above positive signaling via the ITAMs.

Igα and Igβ are covalently associated by a disulfide bridge via cysteine residues that exist within the IgSF extracellular domains of both molecules. The association of the Igα/β heterodimer with membrane-bound Ig occurs through interaction within the transmembrane domains of these proteins.[43] The core BCR complex consists of a single Ig molecule associated with a single Igα/β heterodimer (H_2L_2/Igα/Igβ) (Fig. 4.11).[49]

A current model for the initiation of signals originating from the BCR (see Fig. 4.11) proposes that antigens induce the clustering of BCR complexes, increasing their local density. The increase in density leads to the transfer of phosphate groups to the tyrosine residues of the Igα/β ITAM motifs.[44,48]

Src-family tyrosine kinases, of which LYN, FYN, and BLK are most often implicated, are believed to be responsible for ITAM phosphorylation upon aggregation of Igα/β. They have been shown to physically associate with the heterodimer. It has been suggested that only a fraction of Src-family tyrosine kinases is associated with the Igα/β heterodimer and, upon aggregation, transphosphorylate juxtaposed heterodimers. However, the exact mechanism by which Igα/β undergoes initial tyrosine phosphorylation after antigen engagement remains uncertain. Regardless of mechanism, phosphorylated ITAMs subsequently serve as high-affinity docking sites for cytosolic effector molecules that harbor SH2 domains. The recruitment of the SYK tyrosine kinase, via its tandem SH2 domains, to doubly phosphorylated Ig-α/β ITAMs is thought to be a next step in propagating a BCR-mediated signal. Association of SYK with the BCR leads to its subsequent tyrosine phosphorylation by either Src-family or other Syk tyrosine kinases, further increasing kinase activity. Together, the concerted actions of the Syk and Src-family protein tyrosine kinases activate a variety of intracellular signaling pathways that can lead to the proliferation, differentiation, or death of the cell.[50]

Clinical Consequences of Disruptions in BCR Signaling

Both the development of B lymphocytes and the maintenance of the mature antigen-responsive B-cell pool demand the presence of an intact BCR and its downstream signaling pathway(s). Disruption of these pathways can present clinically with hypogammaglobulinemia and an absence of B cells.

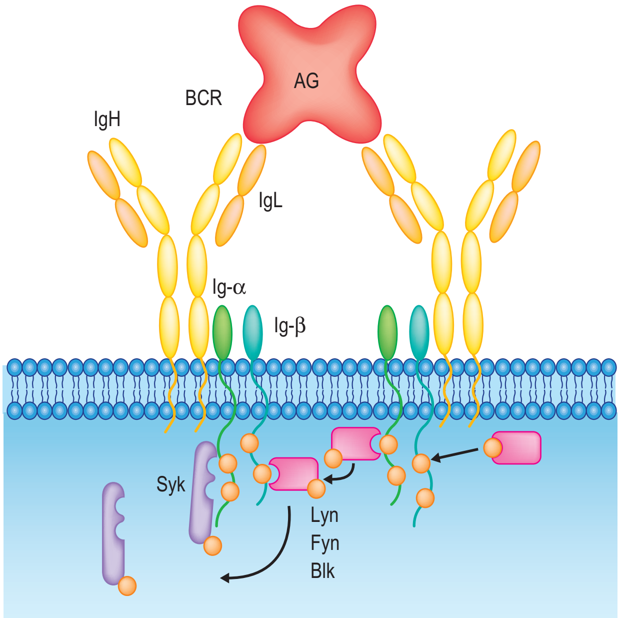


FIG 4.11 The B-Cell Receptor (BCR) Core Complex. The BCR core complex can be divided into an antigen-recognition unit fulfilled by mIgM and a noncovalently associated signal transduction unit composed of the Igα/β (CD79ab) heterodimer. Antigen engagement of mIgM oligomerizes the BCR, allowing preassociated Src-family protein tyrosine kinases to phosphorylate neighboring Igαβ immunoreceptor tyrosine-based activation motif (ITAM) tyrosines. This promotes association of the SYK tyrosine kinase with the tyrosine phosphorylated ITAMs, allowing SYK to become a substrate for other Syk or Src-family tyrosine kinases and leading to its activation.

The most common among such genetic lesions is BTK deficiency, which is an X-linked trait (Chapter 34). BTK plays an important role in BCR signaling both during development and in response to antigen. Loss of function mutations in BTK results in the arrest of human B-cell development at the preB cell stage.

BTK is intact in approximately 10–15% of patients with hypogammaglobulinemia and absence of B cells. Mouse models where BCR components or signaling pathways have been disrupted by targeted mutagenesis have provided insight into the basis of these atypical hypogammaglobulinemia disorders.[44] These studies have shown that an inability to express either a functional μ IgH chain, Igα, Igβ, or the signaling adaptor molecule, BLNK, lead to an early, severe arrest in B lymphopoiesis, with subsequent agammaglobulinemia.[50] Together, these experimental findings highlight the central role of the BCR in the generation and function of B lymphocytes. Thus mutations in any component of the antigen receptor complex or immediate downstream effectors have the potential to disrupt B-cell development and create an agammaglobulinemic state.

Besides its important role in the maturation, differentiation, and survival of B lymphocytes, the B cell antigen receptor is responsible for initiating the humoral response to foreign antigen. Some of the variables that can influence the ultimate outcome of BCR–antigen interaction include the nature of the foreign antigen, the mode of activation, the developmental stage of the B cell, and the microenvironment in which antigen encounter occurs. Exactly how these variables ultimately result in the differential activation of diverse intracellular signaling pathways with fundamentally divergent outcomes is still under study. Emerging from these studies is an appreciation of the role of BCR coreceptors, which have been shown to be capable of modulating antigen receptor signaling in response to antigen.

BCR Coreceptors

The initiation of a humoral immune response results from antigen interaction with the antigen receptors on mature peripheral lymphocytes. However, the manner in which mature B and T lymphocytes recognize antigen is fundamentally different (Chapter 6). Surface Ig, as a component of the BCR on B lymphocytes, typically recognizes an antigenic epitope in its native three-dimensional configuration which, upon engagement with mIg, is capable of transmitting a signal to the cell interior. In contrast, the antigen receptor expressed by T lymphocytes typically recognizes an antigen-derived peptide associated with an appropriate MHC structure (Chapter 5). Further, for this T-cell recognition event to be productive, a CD4 or CD8 coreceptor must also bind to the MHC structure presenting the foreign antigen.

Antigen recognition by the BCR on B lymphocytes is also influenced by coreceptors present on mature B cells (see Table 4.3). In this case, the coreceptors may also recognize antigen, but only in a form that has been modified by other components of the immune system, as described below. In general, these coreceptors and coreceptor complexes can be divided into those that regulate BCR signaling in a positive manner and those that regulate in a negative manner. Thus the ultimate outcome of signaling via the BCR depends not only on the signals transduced via the Igα/β heterodimer but also how these signals are perceived by the cell in association with the signals propagated by the various coreceptors that are concomitantly engaged.

Coreceptors That Positively Regulate BCR Signaling

CD21

Mature B lymphocytes express two receptors for complement C3 components, CD35 (CR1) and CD21 (CR2) (Chapter 21). Of these, CD21 fulfills the requirements of a BCR coreceptor (*vide infra*). The expression of CD21 is restricted to mature B cells and follicular dendritic cells (FDCs), whereas CD35 is also found on erythrocytes, monocytes, and granulocytes. CD21 is a 140 kDa surface glycoprotein encoded by the *CR2* locus on chromosome 1q32 (see Table 4.3). Expression of CD21 begins at approximately the same time as IgD during B lymphopoiesis (Chapter 7). CD21 is subsequently expressed on all mature B cells until terminal differentiation, albeit at different levels depending on B cell population. The extracellular domain of CD21 is composed of 15–16 short consensus regions (SCRs), each composed of 60–70 amino acids, and a relatively short 34-amino acid cytoplasmic tail. The two-amino terminal SCRs constitute the region that interacts with one of the third complement component (C3) cleavage products, iC3b, C3d, g, and C3d (Chapter 21).[51]

CD21 is a receptor for Epstein-Barr virus (EBV), which similarly binds the two N-terminal SCRs via its major envelope glycoprotein gp350/220. CD21, through its oligosaccharide chains, also binds CD23, the low-affinity IgE receptor (FcεRII). Whereas EBV utilization of CD21 for cell entry has clear physiological consequences in terms of infection, B-cell immortalization, and the potential for oncogenesis, the *in vivo* relevance of any CD21–CD23 interaction remains unclear.

CD19

CD19 is an IgSF surface glycoprotein of 95 kDa that is expressed from the earliest stages of B cell development until plasma cell terminal differentiation, when its expression is lost.[52] FDCs also express CD19. *CD19* maps to chromosome 16p11.2, where it encodes a 540 amino acid protein with two extracellular C-type IgSF domains as well as a large, approximately 240 residue, cytoplasmic tail that exhibits extensive conservation between mice and humans. This relatively large cytoplasmic domain includes nine conserved tyrosine residues which, upon phosphorylation, serve as docking sites for other SH2-containing effector molecules. The signaling capacity of CD19 has been shown to result from tyrosine phosphorylation, which occurs upon engagement of the BCR, CD19 or, optimally, by collication of CD19 and IgM. Known signaling effector molecules that have been identified in association with tyrosine-phosphorylated CD19 include the LYN and FYN protein tyrosine kinases, the Rho-family guanine nucleotide exchange factor, VAV, and phosphatidylinositol 3-kinase.[52] Although specific ligands for CD19 have been proposed, the physiological relevance of CD19 engagement by putative ligands has not been demonstrated.

In vitro studies using mAbs directed against CD21 or CD19 provided initial evidence that these B-cell surface antigens could influence mIg-mediated signaling.[51-53] Genetic deficiencies of CD21 (CVID7) or CD19 (CVID3) promote the development of common variable immune deficiency (CVID), which is characterized by hypogammaglobulinemia (Chapter 34). In mice, CD21 and CD19 deficiencies demonstrate impaired antibody response to T-dependent antigens. The paucity of $CD5^+$ B cells in CD19-deficient mice suggests a role for this molecule in the generation and maintenance of the B1 lineage of B cells (Chapter 7). CD19 is expressed from the earliest stages of B cell ontogeny and, accordingly, a signaling function for CD19 in B lymphopoiesis has been demonstrated.[54]

CD21–CD19 Coreceptor Complex

A mechanism by which these molecules could augment BCR-mediated signaling was provided by the identification of a CD21–CD19 coreceptor complex on mature B cells that also includes CD81 (Fig. 4.12). CD81, also known as TAPA-1, is a 26-kDa tetraspan molecule widely expressed on a number of cell types, including lymphocytes. The CD21–CD19 coreceptor model predicted that as a result of complement activation, C3d would be deposited on an antigen, thereby providing a bridge by which a CD21–CD19 receptor complex could associate with mIgM and the BCR complex.[51-53] Clustering of CD19 close to the BCR by the C3d–antigen complex would effectively recruit the signal transduction effector molecules associated with CD19 to the Igα/β heterodimer. As a consequence, the CD19-associated LYN and FYN tyrosine kinases, VAV, and PI3-kinase signaling effector molecules would be in a position to exert their activities on the Igα/β heterodimer–mediated signaling events initiated by antigen engagement of mIgM.

Strong support for CD21–CD19 coreceptor physiological function in BCR signaling was subsequently provided by experiments using a murine model of immune response. Immunization with an antigen covalently attached to C3d dramatically reduced the signaling threshold necessary for antigen to elicit an immune

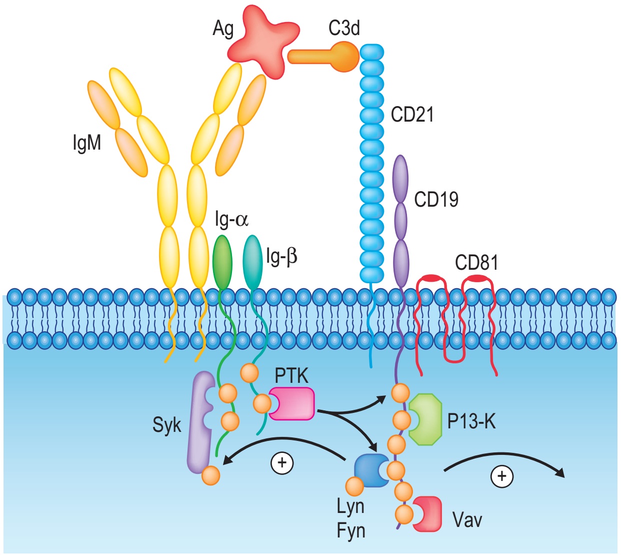


FIG 4.12 Proposed Mechanisms for the Augmentation of B-Cell Receptor (BCR) Signaling by the CD21/19 Coreceptor. Coligation of the BCR and CD21–CD19 complex by C3d–antigen complex allows a CD79-associated Src-family tyrosine kinase to phosphorylate tyrosine residues within the CD19 cytoplasmic domain. Subsequently, tyrosine-phosphorylated CD19 effectively recruits key SH2-containing signaling molecules to the BCR complex, allowing the initial BCR-mediated signal to quickly disseminate along different intracellular signaling pathways.

response.[55] Antigen bearing either two or three copies of C3d was respectively 1000 and 10 000 times more immunogenic than antigen alone. Thus the CD21–CD19 coreceptor complex provides a link between the innate and adaptive immune responses. *In vivo*, CD19-deficient mice appear to have more severely affected T-dependent immune responses compared with CD21-deficient animals, suggesting alternative roles for CD19 in regulating BCR signals beyond the CD21–CD19 coreceptor complex.

Coreceptors That Negatively Regulate BCR Signaling

FcγRIIB

Among the several receptors for the Fc portion of Ig expressed by B cells, the Fc receptor for IgG, FcγRIIB (a member of the CD32 cluster), has an important role in negatively regulating BCR-mediated signal transduction.[56] FcγRIIB is a 40 kDa single-chain molecule that is encoded by single gene located on chromosome 1q23-24. Alternative splicing of different cytoplasmic exons permits expression of three isoforms. The extracellular domain of FcγRIIB is composed of two C-type IgSF domains that can bind with low affinity to IgG. All three FcγRIIB isoforms share a common cytoplasmic region that is important for negatively regulating activation signals delivered by associated surface receptors. The region within the cytoplasmic domain of FCγRIIB responsible for the inhibitory activity of this Fc receptor toward the BCR has been identified as a sequence that contains a tyrosine residue critical for its activity.[57] Analogous to the ITAM, which provides an activation signal, this inhibitory sequence has been referred to as an *immunoreceptor tyrosine-based inhibitory motif* (ITIM). The ITIM is carried by the canonical sequence of I/L/VxYxxI/V/L (where x is any amino acid).[56] ITIMs are found in a number of other transmembrane structures, all of which share the ability to negatively regulate signaling by activating receptors.

The ability of passively administered soluble antibody to inhibit humoral responses has long been appreciated and was initially thought to occur by soluble antibody effectively masking all available antigen epitopes. The molecular mechanism accounting for this suppression is now known to be mediated by the binding of IgG to FcRγIIB and the subsequent recruitment of cytosolic phosphatases to the FcRγIIB ITIM upon tyrosine phosphorylation. Thus the inhibitory effect of IgG on BCR-mediated B cell activation is explained by the interaction of the FcγRIIB ITIM, and specifically associated phosphatases, with the BCR (Fig. 4.13). Coligation of the BCR and FcRγIIB by antigen–IgG complexes results in the tyrosine phosphorylation of the FcRγIIB ITIM, presumably by the BCR-associated tyrosine kinases. Phosphorylated FcRγIIB ITIMs then recruit two different SH2-containing phosphatases, SHIP and SHP-1, which function to remove phosphate groups from inositol lipids or tyrosines, respectively. Although both phosphatases can negatively regulate BCR-mediated signaling events, SHIP appears to be the most relevant phosphatase in FcRγIIB inhibition of BCR signaling (see Fig. 4.13).[58] Thus once the majority of antigen exists in immune complexes together with antigen-specific IgG, attenuation of an ongoing immune response occurs by the juxtaposition of FcRγIIB with the BCR.

CD22

CD22 is a 135-kDa to 140-kDa transmembrane glycoprotein that is restricted in its expression to the B lineage.[59] CD22 expression is limited to the cytoplasm of progenitor and pre-B cells in early B-cell development. Expression on the surface of the B cell occurs concomitant with the appearance of surface, or membrane, IgD. Upon B-cell activation, CD22 expression is initially transiently upregulated and subsequently downmodulated upon terminal differentiation to Ig-secreting plasma cells. Although the onset of CD22 expression follows a similar pattern during murine B lymphopoiesis, it is not restricted to the cytoplasm in early B lymphopoiesis but rather is expressed on the surface from the progenitor stage onward. The basis or function of CD22 intracellular retention in human B cell development is not understood.

CD22 maps to chromosome 19q13.1 and encodes alternatively spliced forms of CD22, CD22α, and CD22β, of which the latter is the predominant species expressed by B cells. The CD22β isoform contains seven extracellular IgSF domains, of which all but one are of the C type. The single exception is the N-terminal domain, which is of the V type. CD22α lacks the IgSF third and fourth domains, although the significance of this minority alternatively spliced product remains unclear. The CD22 murine homologue has only been found as a full-length CD22β isoform. The extracellular domain of CD22 is homologous to the carcinoembryonic antigen subfamily of adhesion molecules, which includes the myelin-associated glycoprotein (MAG) and CD33. CD22 also functions as an adhesion molecule belonging to the Siglec subfamily of the IgSF, whose members function as mammalian sialic acid–binding Ig-like lectins.[59] The two N-terminal IgSF domains have been shown to mediate adhesion to both B and T lymphocytes via the binding of structures carrying α2,6 sialic acids.

In addition to acting as an adhesion molecule, CD22 is also capable of modulating BCR signaling (see Fig. 4.13). A fraction of CD22 associates with the BCR, and CD22 is rapidly tyrosine-phosphorylated upon mIgM engagement.

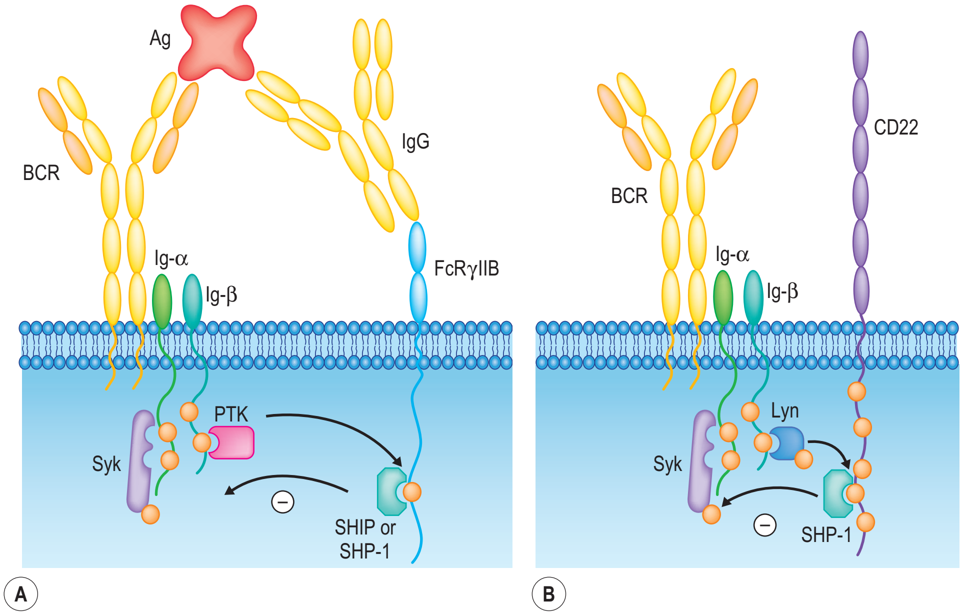


FIG 4.13 **Negative Regulation of B Cell Receptor (BCR) Signaling by FcγRIIB and CD22.** (**A**) Soluble immunoglobulin G (IgG)–antigen immune complexes juxtapose the BCR with FcγRIIB. The BCR-associated LYN tyrosine kinase subsequently tyrosine phosphorylates the FcγRIIB immunoreceptor tyrosine-based inhibitory motif (ITIM). In turn, this leads to the recruitment of the SRC homology 2 (SH2)-containing inositol phosphatase SHIP and tyrosine phosphatase SHP-1 to the phosphorylated FcRγIIB ITIM. Both of these phosphatases have demonstrable inhibitory activity on BCR-mediated signaling. Although SHIP is believed to be the major effector in the FcRγIIB-mediated inhibition of BCR signaling, the exact mechanism of its action in this context has not yet been elucidated. (**B**) CD22 associated with the BCR is tyrosine-phosphorylated upon antigen–BCR engagement. SH2-containing signaling molecules dock on tyrosine phosphorylated residues, including the SHP-1 tyrosine phosphatase that can subsequently dephosphorylate signaling molecules previously activated by a mIgM-mediated signal.

Tyrosine-phosphorylated CD22 associates with several SH2-containing signaling molecules, including the LYN and SYK tyrosine kinases, PI3-kinase, phospholipase C-γ, and SHP-1. The 140-amino acid cytoplasmic domain of CD22 includes six conserved tyrosine residues. Three of these tyrosines are located within conserved consensus ITIM sequences and possess a demonstrable capacity to bind the SH2 domain of the SHP-1 phosphatase. The presence of the multiple ITIMs and association with SHP-1 indicated that CD22 might impinge on BCR signaling in a negative manner. Physiological evidence that CD22 could act as a coreceptor to negatively regulate mIgM signaling was provided by the generation of CD22-deficient mice by targeted mutagenesis.[59] CD22-deficient B cells exhibit hyperactive B-cell responses upon BCR triggering and an increased incidence of serum autoantibodies. This suggests that B-cell tolerance is altered and that B cells are more readily activated in the absence of this negative regulator of BCR signaling.

KEY CONCEPTS

T-Cell Receptor (TCR)–CD3 Complex

- Cell-surface expression of the TCR heterodimers requires association with a complex of invariant proteins designated CD3.
- Each TCR–CD3 complex contains three CD3 dimers.
- Assembly of the TCR–CD3 complex involves interactions between TCR transmembrane basic residues and transmembrane acidic residues in each of the CD3 subunits.
- Signal transduction by the TCR involves the phosphorylation of immunoreceptor tyrosine-based activation motifs (ITAMs) in the cytoplasmic domains of CD3 proteins.
- Phosphorylated CD3 ITAMs recruit and activate the zeta chain-associated protein kinase 70 (ZAP-70) protein tyrosine kinase.
- Deficiency of CD3 proteins impairs T-cell development and can produce severe combined immunodeficiency (SCID).

THE TCR–CD3 COMPLEX

The αβ and γδ TCR heterodimers, which are responsible for the recognition of specific antigen by T lymphocytes, associate with a complex of invariant proteins designated CD3. This association is necessary for TCR cell-surface expression and enables the TCR heterodimers, which have only short cytoplasmic domains, to couple to the intracellular signaling events that lead to the activation of T-cell effector function.[60] There are four CD3 proteins: γ, δ, ε, and ζ (Fig. 4.14).

CD3 Proteins

CD3γ, CD3δ, and CD3ε are structurally similar, and the genes encoding them map to a locus in chromosome 11q23. The

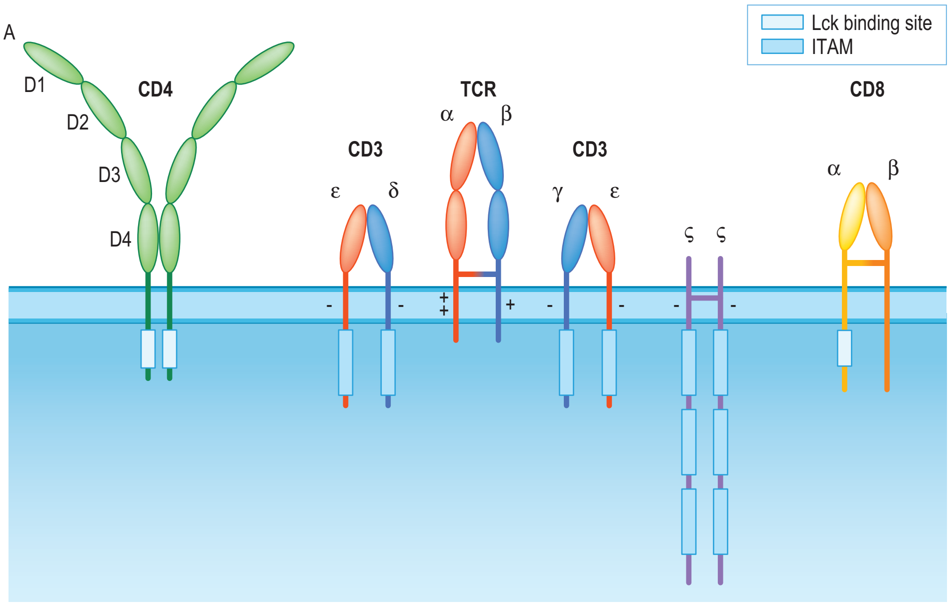


FIG 4.14 Schematic Representation of the Human T-Cell Receptor (TCR) and CD4 and CD8 Coreceptors. Immunoglobulin superfamily (IgSF) domains are represented by ovals. The four extracellular domains of CD4 are labeled D1-D4. Basic (+) and acidic (–) transmembrane charged residues are indicated, as are known and predicted sites of disulfide bonds. For schematic simplicity the cytoplasmic domains of the CD3 chains are shown as extending into the cytoplasm. The cytoplasmic domains of CD3ε and CD3ζ are positively charged and likely are associated with the inner leaflet of the plasma membrane. ITAM, immunoreceptor tyrosine-based activation motifs.

polypeptides range in size from 20 kDa to 25 kDa. Each has an extracellular C-type IgSF domain, a transmembrane region that contains an acidic residue (aspartic acid in CD3δ and CD3ε glutamic acid in CD3γ), and a cytoplasmic domain with a single ITAM. The cytoplasmic domain of CD3ε (but not of CD3δ or CD3γ) has a net positive charge and can bind to the negatively charged inner leaflet of the plasma membrane with its ITAM inserted into the lipid bilayer. The CD3 chains are present in the TCR–CD3 complex in the form of noncovalently linked CD3γε and CD3δε heterodimers; interactions between the extracellular IgSF domains lead to the formation of these CD3 heterodimers.

The 16 kDa CD3ζ differs substantially from the other CD3 proteins and is structurally homologous to the γ chain of the high-affinity IgE receptor (FcRγ chain). The extracellular domain of CD3ζ has only nine amino acids and is of unknown structure. As is the case with the other CD3 chains, the transmembrane region of CD3ζ contains an acidic residue (aspartic acid). The large cytoplasmic domain of CD3ζ has three ITAMs in tandem which, like the ITAM of CD3, also associate with the inner leaflet of the plasma membrane.[18] CD3ζ is usually present in the TCR–CD3 complex in the form of disulfide-linked CD3ζζ homodimers that form through interactions within the transmembrane domain.

Stoichiometry of the TCR–CD3 Complex

The αβTCR–CD3 complex is univalent and consists of a single αβ TCR heterodimer together with three CD3 dimers: CD3γε, CD3δε, and CD3ζζ (see Fig. 4.14). The γδTCR–CD3 complex, in contrast, lacks CD3δ. On naïve T cells, this receptor complex contains two CD3γε heterodimers and one CD3ζζ homodimer. Following activation of γδT cells, the TCR–CD3 complex incorporates the FcRγ chain, either as a homodimer or as a heterodimer with CD3ζ.[18,60,61]

Assembly and Cell-Surface Expression of the TCR–CD3 Complex

Assembly begins with formation of the individual TCRαβ, CD3δε, and CD3γε heterodimers, processes that are driven by interactions between the extracellular domains of the pairing polypeptides. The subsequent higher order assembly of the TCRαβ with the CD3 dimers depends on the interactions between the potentially charged residues within their transmembrane regions. As noted above, each of the CD3 subunits has a transmembrane acidic residue, whereas the transmembrane domains of the αβ and γδ TCRs contain basic residues. Mutation of any of these transmembrane acidic or basic residues to neutral alanine impairs formation of the TCR–CD3 complex. TCRαβ appears to associate first with CD3δε and then with CD3γε. TCRα binds CD3δε, and TCRβ likely interacts with CD3γε. The incorporation of a CD3ζζ homodimer into the complex requires the prior formation of a TCRαβ–CD3γε–CD3δε hexamer and involves interactions between the arginine residue in the transmembrane domain of TCRα and the two colocalized aspartic acids in the transmembrane domains of the CD3ζζ homodimer.[18,62]

Formation of the TCR–CD3 complex is tightly regulated. For example, when there are deficiencies of CD3γ, CD3δ, or CD3ε, TCRα and β are retained in the endoplasmic reticulum and are rapidly degraded. In the absence of CD3ζ, the TCRαβ–CD3γε–CD3δε hexamer is exported to the Golgi apparatus but then is targeted to a lysosomal degradation pathway rather than to the cell surface.[18,60-62]

Because the structures of most of the individual components of the TCR–CD3 complex are known, a model of the overall structure of the receptor has been proposed. This model envisions a compact TCR–CD3 complex, with trimeric contacts occurring within the transmembrane regions of all components (i.e., TCRα–CD3ε–CD3δ, TCRβ–CD3ε–CD3γ, and TCRα–CD3ζ–CD3ζ) and with the TCRαβ projecting further from the membrane (80 Å) than the CD3 chains (40 Å).[18,62]

Mutations in the *CD3D, CD3E, CD3G,* and *CD3Z* genes have been described in humans.[63-65] The clinical consequences of these mutations underscore the importance of the CD3 proteins for the normal development and function of T cells.

Homozygous mutations leading to complete deficiencies of either CD3δ, CD3ε, or CD3ζ protein produce a form of SCID (Chapter 35) characterized by severe T-cell lymphopenia, but in the presence of phenotypically normal B cells and NK cells ($T^-B^+NK^+$SCID).[63,64]

Mutations in *CD3G* leading to deficiency of CD3γ produce considerable clinical heterogeneity ranging from severe immunodeficiency in infants to mild forms of autoimmunity in adulthood. Homozygous deficiency in CD3γ impairs, but does not abrogate, T-cell development, leading to mild T lymphopenia, reduction in cell-surface expression of the TCR–CD3 complex on peripheral T cells by 75–80%, and impaired *in vitro* proliferative T-cell responses to lectins and to anti-CD3 mAbs. In peripheral blood, there are differential effects on phenotypically defined T-cell subsets, with very few CD8 T cells, a 10-fold reduction in $CD45RA^+$ CD4 T cells ("naïve helper" subset), and normal numbers of $CD45RO^+$ CD4 T cells ("memory" cells).[65]

Early Events in TCR–CD3 Signaling

Stimulation of the TCR–CD3 complex by pMHC leads to the phosphorylation of tyrosine residues in the CD3 ITAMs by the SRC-like protein tyrosine kinase, LCK.[66] The phosphorylated CD3 ITAMs, in turn, create high-affinity binding sites for the SH2 domains of the zeta chain-associated protein kinase 70 (ZAP-70) protein tyrosine kinase, leading to its recruitment to the TCR–CD3 complex and to its activation (Chapter 12).[66,67] The consequences of ZAP-70 deficiency (selective T-cell immunodeficiency in humans) underscore the centrality of its role in T-cell activation (Chapter 35).

The TCR appears to act as a mechanosensor to trigger the cascade of complex biochemical events leading to the activation of T-cell effector function. As the T cell migrates over the cell surface of an APC or target cell, the binding of the pMHC complex to the TCR causes the TCR to act as a lever, converting horizontal force into a vertical force that acts on the CD3 chains, exposing their ITAMs for phosphorylation. Following the initiation of signaling, sustained signaling appears to involve multimerization of TCR–CD3 complexes and engagement of coreceptors.[18,62]

T-CELL CORECEPTORS: CD4 AND CD8

Expression of CD4 and CD8 divides mature T cells into two distinct subsets: CD4 T cells (Chapter 16), which recognize peptides in the context of class II MHC molecules, and CD8 T cells (Chapter 17), which recognize antigens presented by class I MHC molecules. Indeed, CD4 binds directly to class II MHC molecules, and CD8 interacts directly with class I MHC molecules (Fig. 4.15). During antigen recognition, CD4 and CD8 are thought to bind the same pMHC complex as the TCR and thus are true coreceptors for the TCR.[68] The cytoplasmic domains of CD4 and CD8 associate with LCK and serve to bring LCK into contact with the CD3 chains of the pMHC-engaged TCR/CD3 complexes, leading to the phosphorylation of CD3 ITAMs and initiation of TCR signaling (Chapter 12).

The expression of the CD4 and CD8 coreceptors is highly regulated during T-cell development in the thymus (Chapter 8). Thymocytes initially express neither coreceptor ("double negative"). $CD4^-CD8^-$ thymocytes destined to become TCRαβ T cells progress through a $CD4^+CD8^+$ ("double-positive") stage to become mature CD4 or CD8 T cells. Positive and negative selection of thymocytes on the basis of their TCR specificities, and commitment to the CD4 or CD8 lineages occur during the double-positive stage.

CD4: Structure and Binding to MHC Class II Molecules

A member of the IgSF, CD4 is a 55 kDa glycoprotein whose relatively rigid extracellular region contains four IgSF domains (designated D1–4). Its cytoplasmic domain contains two cysteine residues that mediate a noncovalent interaction with LCK through a "zinc clasp"–like structure formed with a dicysteine motif in the N-terminal region of LCK.[66,69-71]

The N-terminal domain (D1) of CD4 binds between the membrane-proximal α_2 and β_2 domains of MHC class II. Thus CD4 interacts with pMHC class II at a distance from the α helices and peptide contacted by the TCR, enabling the TCR and CD4 to bind the same MHC class II molecule simultaneously.

Although MHC molecules are highly polymorphic, the CD4 contact sites are highly conserved. In humans, CD4 targets nonpolymorphic residues shared by all three MHC class II molecules (HLA-DR, -DP, and -DQ). The crystal structure of the TCRαβ–pMHC–CD4 ternary complex assumes a V-shape with pMHC at the apex and with TCRαβ and CD4 forming the arms of the V. There is no direct interaction between the coreceptor and the TCR heterodimer, indicating that pMHC brings the TCR and CD4 together. The approximately 70 Å of separation between the membrane-proximal domains of TCRαβ and CD4 would allow the CD3 chains to lie within the open angle between TCRαβ and CD4, promoting interactions between CD3 chains and CD4-associated LCK.[66,69,71]

Experiments using soluble forms of CD4 and pMHC have revealed that monomeric CD4 binds pMHC with very low affinity (Kd approximately 200 μM). The binding of CD4 to pMHC is of lower affinity than that of TCRαβ to pMHC (Kd 1–10 μM) and displays a far more rapid off time. Because of the low affinity and the rapid off time, it is unlikely that interactions of CD4 with MHC class II molecules initiate the interaction between a T cell and an APC (Chapter 6). Rather, these binding characteristics are more compatible with a model in which the initial event is the interaction between the TCR and pMHC, followed by the recruitment of CD4, which acts primarily to promote signaling events through the delivery of LCK.[66,69,71]

CD8: Structure and Binding to MHC Class I Molecules

There are two CD8 polypeptides, α and β, and these are expressed on the cell surface either as a disulfide-linked CD8αα homodimer or as a disulfide-linked CD8αβ heterodimer. On most αβ T cells, CD8αβ is the predominant form of CD8 while natural killer (NK) cells (Chapter 17), intestinal intraepithelial T cells, and γδ T cells exclusively express CD8αα.[66,69-71]

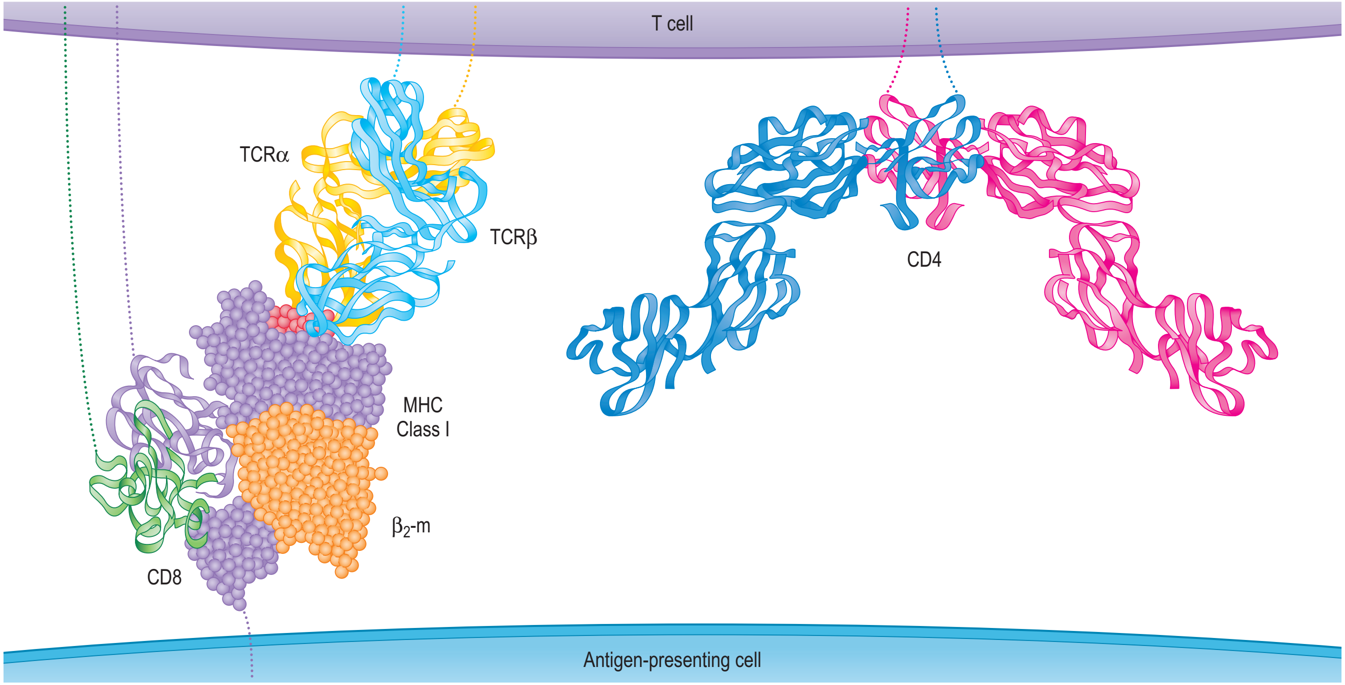


FIG 4.15 Illustration of the Interactions between the T-Cell Receptor (TCR), Peptide Major Histocompatibility Complex (pMHC), and CD8. A composite illustration of the human leukocyte antigen (HLA)-A*0201 structure in complex with a Tax peptide and its cognate TCR α and β chains (protein data bank (pdb) designation 1BD2) with the human CD8αα/HLA-A*0201 structure (pdb designation 1AKJ) was generated by superposition of the HLA moiety of the two structures. The HLA heavy chain is indicated as MHC, its light chain (β_2 microglobulin) as β_2-m, the CD8αα homodimer as CD8, the TCR α and β chains as TCRα and TCRβ. In addition, the CD4 homodimer (pdb file 1WIO) is shown to scale. Connecting peptides, transmembrane, and cytoplasmic domains are drawn by hand and indicated by dotted lines. (Figure courtesy of David H. Margulies, National Institute of Allergy and Infectious Diseases, National Institutes of Health.)

CD8α, a 34- to 37-kDa protein, and CD8β, a 32-kDa protein, share about 20% amino acid sequence homology. Both are glycoproteins and IgSF members. Although CD8 subserves a coreceptor function similar to that of CD4, in structure it differs substantially from CD4. The CD8 extracellular regions have single N-terminal IgSF V domains at the end of extended mucin-like stalk regions of 48 amino acids (CD8α) or 35–38 amino acids (CD8β). A striking difference between the two forms of CD8 lies within the cytoplasmic domain. CD8α, like CD4, contains a cysteine-based motif that enables it to interact with LCK through a "zinc clasp"–like structure. In contrast, CD8β lacks this motif and does not associate with LCK. Interestingly, CD8αβ appears to be a more effective activator of TCR signaling than CD8αα. This may reflect the palmitoylation of the cytoplasmic domain of CD8β, which allows CD8αβ to associate with lipid rafts during T-cell activation.[66,69,71,72]

The structure of CD8αα–pMHC class I complexes demonstrates that CD8αα binds to conserved residues in the α_3 domain of MHC class I (i.e., a nonpolymorphic, membrane-proximal region of the molecule distinct from the peptide-binding groove engaged by the TCR) (Chapter 5). Compared with the interaction of CD4 and MHC class II, binding is more antibody-like, with a loop of the MHC α_3 domain locked between the CDR-like loops of the two CD8α IgSF V domains. Models of the structure of the TCRαβ–pMHC–CD8 ternary complex propose a "V" shape similar to that of the crystal structure of TCRαβ–pMHC–CD4, with pMHC at the apex of the "V" and the TCR and CD8 forming the arms of the "V." CD8 binds to pMHC with lower affinity and with faster kinetics compared with the TCR. Thus the binding properties of the CD8 coreceptor, like those of CD4, are consistent with a model in which the TCR initiates pMHC binding, followed by engagement of CD8 to the same pMHC.[66,69-72]

COSTIMULATORY AND INHIBITORY T-CELL MOLECULES: THE CD28 FAMILY

Although the T-cell response to antigen requires the binding of the TCR and its coreceptors to pMHC, additional receptor–ligand interactions affect the outcome by delivering signals that promote activation (costimulation) or that inhibit it (Table 4.4). Prominent among these are the interactions of members of the CD28 family with their cell-surface ligands on APCs.[73] This family includes CD28, inducible T-cell costimulator (ICOS), cytotoxic T lymphocyte antigen-4 (CTLA-4), B- and T lymphocyte attenuator (BTLA), and programmed death 1 (PD-1). CD28 and ICOS are costimulatory receptors; the major functions of CTLA-4, PD-1, and BTLA are inhibitory. CD28 and CTLA-4 are T-cell specific, whereas BTLA and PD-1 are also expressed by B cells and ICOS by NK cells. CD28, CTLA-4, and PD-1 are the targets of therapeutic interventions in current clinical practice.

TABLE 4.4 CD28 Superfamily

Receptor	Expression	Ligand	Function on T Cells
CD28	Most CD4 T cells 50% CD8 T cells	B7-1 (CD80) B7-2 (CD86)	Costimulation of interleukin (IL)-2 production and proliferation; Promotes T cell survival
Inducible T cell costimulator (ICOS)	Activated and memory T cells Natural killer (NK) cells Not expressed by naïve T cells	ICOS ligand	Promotes T cell differentiation and effector T-cell function
Cytotoxic T lymphocyte antigen-4 (CTLA-4)	Upregulated after T-cell activation	B7-1 (CD80) B7-2 (CD86)	Inhibits IL-2 production and proliferation; Promotes peripheral T cell tolerance
Programmed death 1 (PD-1)	Upregulated after activation of T and B cells, myeloid cells	PD-L1 (B7-H1) PD-L2 (B7-DC)	Inhibits proliferations and cytokine production Promotes peripheral T cell tolerance
B and T lymphocyte attenuator (BTLA)	T and B cells, myeloid cells, dendritic cells	HVEM (herpesvirus-entry mediator)	Inhibits T cell proliferation

All members of the CD28 family have a single extracellular IgSF V domain and have, as their ligands, members of the B7 family of cell surface molecules. CD28, CTLA-4, and ICOS are disulfide-linked homodimers whose cytoplasmic domains contain the SH2-binding motif YXXM. In contrast, PD-1 and BTLA are monomers whose cytoplasmic domains each contain an ITIM and an immunoreceptor tyrosine-based switch motif (ITSM).

CD28 and CTLA-4

Half of CD8 T cells and virtually all human CD4 T cells constitutively express CD28. CD28 binds to B71 (CD80) and B7.2 (CD86) through an MYPPPYY motif in its extracellular domain. Interactions with these ligands leads to the phosphorylation of the YMNM sequence in the CD28 cytoplasmic domain and to the recruitment of phosphatidylinositol 3-kinase and Grb2. CD28 stimulation usually does not elicit a cellular response in the absence of TCR signaling. Rather CD28 signals act in concert with those of the TCR to promote cytokine production, T cell expansion, and T cell survival. TCR signaling in the absence of CD28 costimulation can induce T cell anergy (Chapter 12).

CTLA-4 inhibits the response to TCR and CD28 signals and acts to terminate peripheral T cell responses. Its importance in human immunology is underscored by observations that *CTLA4* haploinsufficiency produces a syndrome of immune dysregulation characterized by decreased numbers of regulatory T cells (Tregs), hyperactive effector T cells, hypogammaglobulinemia, and clinical autoimmunity.

The majority of CTLA-4 resides in intracellular compartments. T cell activation promotes the cell-surface expression of CTLA-4 by regulating both its transport to the surface and its subsequent internalization. CTLA-4 also binds B7.1 and B7.2 but does so with substantially greater affinity than does CD28. Moreover, the binding of CTLA-4 to these ligands is divalent, whereas that of CD28 is monovalent. Thus the inhibitory complexes formed by CTLA-4 are more stable than the costimulatory interactions involving CD28. CTLA-4 can inhibit T-cell activation by outcompeting CD28 for B7 ligands and, through trans-endocytosis, by removing B7 molecules from the APC. In addition, CTLA-4 can induce "reverse signaling" through B7.1 and B7.2 to the APC, upregulating the enzyme indoleamine 2,3-dioxygenase (IDO), which, in turn, breaks down tryptophan, a requirement for T cell proliferation.

The importance of CD28 costimulation has made it an attractive target for therapeutic intervention.[74,75] Indeed, two soluble fusion proteins composed of the extracellular domain of human CTLA-4 and the constant regions of human IgG1, abatacept and belatacept, are effective therapies for the treatment of rheumatoid arthritis (Chapter 52) and the prevention of renal allograft rejection (Chapter 81). These fusion proteins are thought to inhibit CD28 costimulation through blockade of its B7 ligands, but some of their immunosuppressive effects may be indirect through the induction of IDO and consequent local depletion of tryptophan. Conversely, inhibition of CTLA-4 by mAbs can promote durable immune responses against certain malignancies.

PD-1

PD-1 is a key inhibitory receptor that attenuates TCR signaling, promotes T-cell tolerance, and is associated with T-cell exhaustion. PD-1 is not found on resting T cells, and its expression during T-cell activation requires transcriptional activation. PD-1 binds to two ligands: programmed death ligand 1 (PDL-1), which is widely expressed, and PDL-2, which is found primarily on professional antigen presenting cells. Engagement of ligand induces tyrosine phosphorylation of the ITIM and the ITSM in the cytoplasmic domain of PD-1, leading to the recruitment of the tyrosine phosphatase SHP-2. Continued stimulation of T cell by antigen leads to sustained expression of PD-1 and differentiation into a state of hyporesponsiveness termed *T-cell exhaustion*. Blockade of PD-1 has shown considerable promise in the treatment of diverse human malignancies.

ON THE HORIZON

- Elucidation of the mechanisms that regulate which epitopes will be preferentially bound by normal and abnormal antigen receptor repertoires
- Elucidation of the mechanisms that create constraints on the diversity of antigen receptor repertoires and the role they play in diseases of immune function
- Targeting of vaccines to elicit responses to specific epitopes
- Development of additional new therapies focused on regulating signal transduction from the B cell receptor or T cell receptor to either dampen or enhance the immune response, especially at critical checkpoints

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Williams AF, Barclay AN. The immunoglobulin superfamily—domains for cell surface recognition. Ann Rev Immunol 1988;6:381–405.
2. Padlan EA. Anatomy of the antibody molecule. Mol Immunol 1994;31(3):169–217.
3. Sanchez-Mazas A, Fernandez-Vina M, Middleton D, et al. Immunogenetics as a tool in anthropological studies. Immunology 2011;133(2):143–64.
4. Ehrenmann F, Kaas Q, Lefranc MP. IMGT/3Dstructure-DB and IMGT/DomainGapAlign: a database and a tool for immunoglobulins or antibodies, T cell receptors, MHC, IgSF and MhcSF. Nucleic Acids Res 2010;38(Database issue):D301–7.
5. Nelson AL. Antibody fragments: hope and hype. MAbs 2010;2(1):77–83.
6. Schroeder HW Jr, Cavacini L. Structure and function of immunoglobulins. J Allergy Clin Immunol 2010;125(2 Suppl. 2):S41–52.
7. Johansen FE, Braathen R, Brandtzaeg P. Role of J chain in secretory immunoglobulin formation. Scand J Immunol 2000;52(3):240–8.
8. Chen FH, Arya SK, Rinfret A, et al. Domain-switched mouse IgM/IgG2b hybrids indicate individual roles for C mu 2, C mu 3, and C mu 4 domains in the regulation of the interaction of IgM with complement C1q. J Immunol 1997;159(7):3354–63.
9. Sensel MG, Kane LM, Morrison SL. Amino acid differences in the N-terminus of C(H)2 influence the relative abilities of IgG2 and IgG3 to activate complement. Mol Immunol 1997;34(14):1019–29.
10. Skvaril F, Schilt U. Characterization of the subclass and light chain types of IgG antibodies to rubella. Clin Exp Immunol 1984;55:671–6.
11. Barrett DJ, Ayoub EM. IgG2 subclass restriction of antibody to pneumococcal polysaccharides. Clin Exp Immunol 1986;63:127–34.
12. Otteson EA, Skvaril F, Tripathy SP, et al. Prominence of IgG4 in the IgG antibody response to human filiariasis. J Immunol 1985;134:2707–12.
13. James LK, Till SJ. Potential Mechanisms for IgG4 Inhibition of Immediate Hypersensitivity Reactions. Curr Allergy Asthma Rep 2016;16(3):23.
14. Martinez-Valle F, Fernandez-Codina A, Pinal-Fernandez I, et al. IgG4-related disease: evidence from six recent cohorts. Autoimmun Rev 2017;16(2):168–72.
15. Suzuki K, Ha SA, Tsuji M, et al. Intestinal IgA synthesis: a primitive form of adaptive immunity that regulates microbial communities in the gut. Semin Immunol 2007;19(2):127–35.
16. Stone KD, Prussin C, Metcalfe DD. IgE, mast cells, basophils, and eosinophils. J Allergy Clin Immunol 2010;125(2 Suppl. 2):S73–80.
17. Davis MM, Bjorkman PJ. T-cell antigen receptor genes and T-cell recognition. Nature 1988;334:395–402.
18. Rossjohn J, Gras S, Mile JJ, et al. T cell antigen receptor recognition of antigen-presenting molecules. Annu Rev Immunol 2015;33:169–200.
19. O'Brien RL, Roark CL, Jin N, et al. γδT-cell receptors: functional correlations. Immunol Rev 2007;215:77–88.
20. Cohen NR, Garg S, Brenner MB. Antigen Presentation by CD1 Lipids, T Cells, and NKT Cells in Microbial Immunity. Adv Immunol 2009;102:1–94.
21. Bonneville M, O'Brien RL, Born WK. Gammadelta T cell effector functions: a blend of innate programming and acquired plasticity. Nat Rev Immunol 2010;10(7):467–78.
22. Allison TJ, Garboczi DN. Structure of gammadelta T cell receptors and their recognition of non-peptide antigens. Mol Immunol 2002;38(14):1051–61.
23. Li H, Llera A, Malchiodi EL, et al. The structural basis of T cell activation by superantigens. Ann Rev Immunol 1999;17:435–66.
24. Tonegawa S. Somatic generation of antibody diversity. Nature 1983;302:575–81.
25. Krangel MS. Mechanics of T cell receptor gene rearrangement. Curr Opin Immunol 2009;21(2):133–9.
26. Roth DB. V(D)J recombination: mechanism, errors, and fidelity. Microbiol Spectr 2014;2(6).
27. Carmona LM, Schatz DG. New insights into the evolutionary origins of the recombination-activating gene proteins and V(D)J recombination. FEBS J 2016;doi:10.1111/febs.13990. [Epub ahead of print].
28. Zachau HG. The immunoglobulin kappa gene families of human and mouse: a cottage industry approach. Biol Chem 2000;381(9-10):951–4.
29. Kirkham PM, Schroeder HW Jr. Antibody structure and the evolution of immunoglobulin V gene segments. Semin Immunol 1994;6(6):347–60.
30. Kawasaki K, Minoshima S, Nakato E, et al. One-megabase sequence analysis of the human immunoglobulin lambda gene locus. PCR Meth App 1997;7(3):250–61.
31. Vettermann C, Jack HM. The pre-B cell receptor: turning autoreactivity into self-defense. Trends Immunol 2010;31(5):176–83.
32. Matsuda F, Ishii K, Bourvagnet P, et al. The complete nucleotide sequence of the human immunoglobulin heavy chain variable region locus. J Exp Med 1998;188(11):2151–62.
33. Hwang JK, Alt FW, Yeap LS. Related Mechanisms of Antibody Somatic Hypermutation and Class Switch Recombination. Microbiol Spectr 2015;3(1):MDNA3-0037-2014.
34. Casellas R, Basu U, Yewdell WT, et al. Mutations, kataegis and translocations in B cells: understanding AID promiscuous activity. Nat Rev Immunol 2016;16(3):164–76.
35. Tomlinson IM, Cox JP, Gherardi E, et al. The structural repertoire of the human V kappa domain. EMBO J 1995;14(18):4628–38.
36. Chothia C, Lesk AM, Gherardi E, et al. Structural repertoire of the human VH segments. J Mol Biol 1992;227:799–817.
37. Tomlinson IM, Walter G, Jones PT, et al. The imprint of somatic hypermutation on the repertoire of human germline V genes. J Mol Biol 1996;256(5):813–17.
38. Khass M, Blackburn T, Burrows PD, et al. A new role for VpreB: an invariant surrogate antigen that selects Ig antigen binding sites. Sci Immunol 2016;1(1).
39. Arden B, Clark SP, Kabelitz D, et al. Human T-cell receptor variable gene segment families. Immunogenet 1995;42(6):455–500.
40. Jung D, Giallourakis C, Mostoslavsky R, et al. Mechanism and control of V(D)J recombination at the immunoglobulin heavy chain locus. Annu Rev Immunol 2006;24:541–70.
41. Levin-Klein R, Bergman Y. Epigenetic regulation of monoallelic rearrangement (allelic exclusion) of antigen receptor genes. Front Immunol 2014;5:625.
42. Nemazee D, Weigert M. Revising B cell receptors. J Exp Med 2000;191(11):1813–17.
43. Neuberger MS, Patel KJ, Dariavach P, et al. The mouse B-cell antigen receptor: definition and assembly of the core receptor of the five immunoglobulin isotypes. Immunol Rev 1993;132:147–61.
44. Wang LD, Clark MR. B-cell antigen-receptor signalling in lymphocyte development. Immunology 2003;110(4):411–20.
45. Pelanda R, Braun U, Hobeika E, et al. B cell progenitors are arrested in maturation but have intact VDJ recombination in the absence of Ig-alpha and Ig-beta. J Immunol 2002;169(2):865–72.
46. Mason DY, Cordell JL, Brown MH, et al. CD79a: a novel marker for B-cell neoplasms in routinely processed tissue samples. Blood 1995;86(4):1453–9.
47. Tsuganezawa K, Kiyokawa N, Matsuo Y, et al. Flow cytometric diagnosis of the cell lineage and developmental stage of acute lymphoblastic leukemia by novel monoclonal antibodies specific to human pre-B-cell receptor. Blood 1998;92(11):4317–24.
48. Gauld SB, Cambier JC. Src-family kinases in B-cell development and signaling. Oncogene 2004;23(48):8001–6.
49. Schamel WW, Reth M. Monomeric and oligomeric complexes of the B cell antigen receptor. Immunity 2000;13(1):5–14.
50. Conley ME. Genetics of hypogammaglobulinemia: what do we really know? Curr Opin Immunol 2009;21(5):466–71.
51. Carroll MC. The role of complement and complement receptors in induction and regulation of immunity. Ann Rev Immunol 1998;16:545–68.
52. Fujimoto M, Poe JC, Inaoki M, et al. CD19 regulates B lymphocyte responses to transmembrane signals. Semin Immunol 1998;10(4):267–77.
53. Fearon DT, Carroll MC. Regulation of B lymphocyte responses to foreign and self-antigens by the CD19/CD21 complex. Ann Rev Immunol 2000;18:393–422.
54. Otero DC, Anzelon AN, Rickert RC. CD19 function in early and late B cell development: I. Maintenance of follicular and marginal zone B cells requires CD19-dependent survival signals. J Immunol 2003;170(1):73–83.

55. Dempsey PW, Allison ME, Akkaraju S, et al. C3d of complement as a molecular adjuvant: bridging innate and acquired immunity. Science 1996;271(5247):348–50.
56. Bournazos S, Wang TT, Dahan R, et al. Signaling by antibodies: recent progress. Annu Rev Immunol 2017;35:285–311.
57. Nimmerjahn F, Ravetch JV. Fcγ receptors: old friends and new family members. Immunity 2006;24(1):19–28.
58. Ono M, Okada H, Bolland S, et al. Deletion of SHIP or SHP-1 reveals two distinct pathways for inhibitory signaling. Cell 1997;90(2):293–301.
59. Nitschke L. CD22 and Siglec-G: B-cell inhibitory receptors with distinct functions. Immunol Rev 2009;230(1):128–43.
60. Brazin KN, Mallis RJ, Das DK, et al. Structural features of the αβ TCR mechanotransduction apparatus that promote PMHC discrimination. Front Immunol 2015;6:441.
61. Rudolph MG, Stanfield RL, Wilson IA. How TCRs bind MHCs, peptides, and co-receptors. Ann Rev Immunol 2006;24:419–66.
62. Kuhns MS, Davis MM, Garcia KC. Deconstructing the form and function of the TCR/CD3 complex. Immunity 2006;24(2):133–9.
63. de Saint-Basile G, Geissmann F, Flori E, et al. Severe combined immunodeficiency caused by deficiency in either the δ or the ε subunit of CD3. J Clin Invest 2004;114(10):1512–17.
64. Roberts JL, Lauritsen JP, Cooney M, et al. T-B+NK+ severe combined immunodeficiency caused by complete deficiency of the CD3ζ subunit of the T-cell antigen receptor complex. Blood 2007;109(8):3198–206.
65. Recio MJ, Moreno-Pelayo MA, Kilic SS, et al. Differential biological role of CD3 chains revealed by human immunodeficiencies. J Immunol 2007;178(4):2556–64.
66. Gao GF, Rao Z, Bell JI. Molecular coordination of αβ T-cell receptors and coreceptors CD8 and CD4 in their recognition of peptide-MHC ligands. Trends Immunol 2002;23(8):408–13.
67. Palacios EH, Weiss A. Function of the Src-family kinases, Lck and Fyn, in T-cell development and activation. Oncogene 2004;23(48):7990–8000.
68. Li Y, Yin Y, Mariuzza RA. Structural and biophysical insights into the role of CD4 and CD8 in T cell activation. Front Immunol 2013;4:206.
69. van der Merwe PA, Davis SJ. Molecular interactions mediating T cell antigen recognition. Ann Rev Immunol 2003;21:659–84.
70. Artyomov MN, Lis M, Devadas S, et al. CD4 and CD8 binding to MHC molecules primarily acts to enhance Lck delivery. Proc Natl Acad Sci USA 2010;107(39):16916–21.
71. Wang Z, Yang X, Chu X, et al. The structural basis for the oligomerization of the N-terminal domain of SATB1. Nucleic Acids Res 2012;40(9):4193–202.
72. Chang HC, Tan K, Ouyang J, et al. Structural and mutational analyses of a CD8αβ heterodimer and comparison with the CD8αα homodimer. Immunity 2005;23(6):661–71.
73. Murakami N, Riella LV. Co-inhibitory pathways and their importance in immune regulation. Transplantation 2014;98(1):3–14.
74. Baumeister SH, Freeman GJ, Dranoff G, et al. Coinhibitory pathways in immunotherapy for cancer. Annu Rev Immunol 2016;34:539–73.
75. Topalian SL, Drake CG, Pardoll DM. Immune checkpoint blockade: a common denominator approach to cancer therapy. Cancer Cell 2015;27(4):450–61.

MULTIPLE-CHOICE QUESTIONS

1. The mechanism that generates the greatest diversity in immunoglobulins (Igs) is:
   A. Combinatorial V(D)J rearrangement
   B. Combinatorial H and L chain association
   C. N nucleotide addition
   D. Somatic hypermutation (SHM)
   E. Class-switch recombination (CSR)

2. The mechanism that generates the greatest diversity in T-cell receptors (TCRs) is:
   A. Combinatorial V(D)J rearrangement
   B. Combinatorial H and L chain association
   C. N-nucleotide addition
   D. SHM
   E. CSR

3. Signal transduction through the B-cell receptor (BCR) is inhibited by:
   A. CD3ζ
   B. CD19
   C. CD21
   D. CD22
   E. CD81

4. Signal transduction through the TCR is enhanced by:
   A. FcγRII
   B. CD28
   C. Cytotoxic T lymphocyte antigen-4 (CTLA-4)
   D. Programmed death 1 (PD-1)
   E. CD79a

5

The Major Histocompatibility Complex

Dimitrios S. Monos, Robert J. Winchester

A primary objective of the immune system is to protect our bodies against pathogens. The major histocompatibility complex (MHC) comprises a genomic region that has evolved to include many genes responsible for coordinating the immune response. It is named the *histocompatibility complex* because it was first identified as the site of numerous genes that determined whether transplanted tissue would be accepted or rejected. We now know that this region coordinates immunological functions far beyond those related to histocompatibility. The MHC region includes genes that determine both innate and adaptive immunities and thus influences responses to pathogens (viruses, bacteria, fungi, and parasites), transplantation, autoimmunity, cancer biology, vaccinations, responses to drugs, and possibly other functionalities presently unknown. This chapter describes some aspects of the genomic organization of the MHC region and its immunological importance, focusing primarily on the human leukocyte antigen (HLA) genes and their encoded molecules. These HLA molecules play a central role in adaptive immunity. They serve as the structures that present self and foreign peptides to T cells (Chapter 6). They also participate in aspects of innate immunity by interacting with receptors on the surface of natural killer (NK) cells (Chapter 17). The terms HLA and MHC are often used interchangeably. However, in this chapter, "MHC" is reserved for the broader genomic region, and "HLA" is used to indicate the genes and their respective encoded protein products in humans.

CLINICAL RELEVANCE

- HLA molecules regulate antigen-specific immune responses by binding pathogen-derived peptides and then presenting them to either CD4 or CD8 T cells.
  - CD4 T cells activated by MHC class II peptide complexes can stimulate B cells to produce antigen-specific antibodies.
  - CD8 T cells activated by MHC class I peptide complexes become cytotoxic and can kill the cells that present the pathogen-derived peptides.
- Certain HLA alleles are the major genetic determinants of susceptibility to many autoimmune diseases or drug hypersensitivity reactions.
  - The process involves a combination of self peptides or small molecules (drugs) bound to specific HLA alleles and then recognized by T-cell receptors.
- HLA molecules play a key role in governing transplant rejection and appear to regulate placental development in pregnancy.
- Cancerous cells modify expression of their HLA genes to avoid recognition by the immune response.

GENOMIC ORGANIZATION OF THE MHC

The human MHC region includes approximately 3.8 million base pairs (Mbps) of DNA on the short arm of chromosome 6 (6p21.3) and is defined as the region spanning from the gamma-aminobutyric acid receptor (*GABBR1*) gene on the telomeric side of the region to the kinesin family member C1 (*KIFC1*) gene toward the centromere (ENSEMBL 86 GRCh38.p7 coordinates chr6: 29555629-33409924).[1] More recent work has suggested that the functional MHC region may include additional downstream and upstream sequences totaling seven or more Mbps.

The classic 3.8Mbp MHC region is among the most gene-dense segments of the human genome. It includes 158 protein-coding genes and 86 pseudogenes of unknown functionality (ENSEMBL 86 GRCh38.p7).[2] At least 65 (41%) of the coding genes are involved in innate and adaptive immunities.[2] The MHC is one of the most studied regions of the genome because it includes the genes that encode the highly polymorphic HLA proteins, which play a pivotal role in immune recognition. The MHC is divided into three regions: class I, class II, and class III (Fig. 5.1). The class I region is at the telomeric end and includes the classic HLA class I genes (*HLA-A, HLA-B,* and *HLA-C*), the class I–related (like) genes (*MICA, MICB*), the nonclassic HLA class I genes (*HLA-E, HLA-F,* and *HLA-G*), and four pseudogenes (*HLA-H, HLA-K, HLA-J,* and *HLA-L*). The class II region occupies the centromeric end and contains the *DRA* and *DRB1* genes and, depending on the DR haplotype, one or none of the *DRB3, DRB4, DRB5* genes that code for the DR, DR52, DR53, or DR51 molecules, respectively; the *DQA1* and *DQB1* genes that encode the DQ molecule; and the *DPA1* and *DPB1* genes that encode the DP molecule. It also includes the *DM* and *DO* genes encoding the antigen-processing molecules DM and DO involved in the class II antigen presentation pathway (Chapter 6), as well as the *TAP* and *LMP* genes encoding proteins involved in the classic class I antigen presentation pathway (Chapter 6). The class III region, interposed between class I and II regions, is a very gene-dense region and contains many immune and non-immune related genes, including those encoding some complement components, lymphotoxin α and β, tumor necrosis factor (TNF), heat shock proteins (HSPs), *NFKB, NOTCH4,* and 21-hydroxylase (*CYP21*). The genes within the HLA class I and class II regions demonstrate the sequence and structural homology that marks evolution through mechanisms of gene conversion, gene duplications, insertions, deletions, and subsequent mutations, resulting in divergence of function. Although the genomic organization of class I and class II genes is quite distinct, the

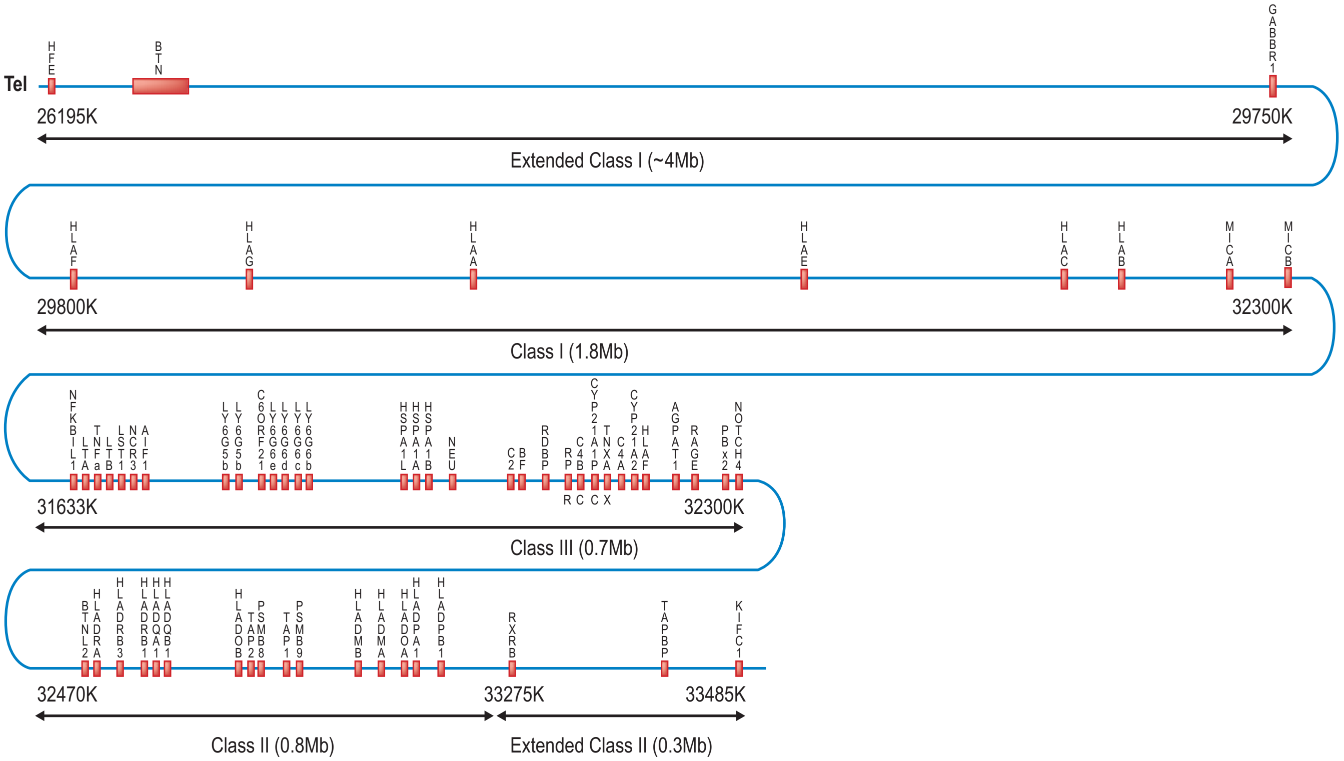

Tel
HFE
BTN
GABBR1
26195K
29750K
Extended Class I (~4Mb)
HLAF
HLAG
HLAA
HLAE
HLAC
HLAB
MICA
MICB
29800K
32300K
Class I (1.8Mb)
NFKBIL1
LTA
TNFa
LTB
LST1
NCR3
AIF1
LY6G5b
LY6G5b
C6ORF21
LY6G6e
LY6G6d
LY6G6c
LY6G6b
HSPA1L
HSPA1A
HSPA1B
NEU
C2
BF
RDBP
RP
C4B
CYP21A1P
TNXA
C4A
CYP21A2
HLAF
AGPAT1
RAGE
PBx2
NOTCH4
RCCX
31633K
32300K
Class III (0.7Mb)
BTNL2
HLADRA
HLADRB3
HLADRB1
HLADQA1
HLADQB1
HLADOB
TAP2
PSMB8
TAP1
PSMB9
HLADMB
HLADMA
HLADOA
HLADPA1
HLADPB1
RXRB
TAPBP
KIFC1
32470K
33275K
33485K
Class II (0.8Mb)
Extended Class II (0.3Mb)

FIG 5.1 Gene Map of the Extended Major Histocompatibility Complex (MHC) Complex. The core of the MHC complex consists of three major regions: class I, class III, and class II. The extended regions of the complex fare denoted as extended class I and extended class II. Sequence numbering begins at the telomere. The map depicts immune-related expressed genes as well as certain reference genes. The approximate locations of these selected genes near the start or end of the regions are indicated. [Modified from Beck S, Trowsdale J. The human major histocompatibility complex: lessons from the DNA sequence. Annu Rev Genomics Hum Genet 2000; 1:117–37.]

structures of the derived molecules have a striking resemblance that presents a unique example of convergent evolution, perhaps driven by the shared role of these molecules in presenting peptide to the T-cell receptor (TCR).

Another unique characteristic of the MHC is the extensive linkage disequilibrium (LD) observed among the very distant genomic regions of *HLA-A, HLA-B, HLA-C, HLA-DR,* and *HLA-DQ* genes, but not the *HLA-DP* gene. LD is the phenomenon whereby particular alleles of gene loci on the same strand of DNA are inherited together more often that would be expected by chance. Anthropological population studies have suggested that the particular combinations of alleles of the different genes, as distant as they may be, provide a survival advantage, perhaps reflecting functional interdependence in antigen-specific immune responses.

A particular combination of alleles of different loci in LD on the same strand of DNA is called a *haplotype*. The frequency of a given haplotype varies among different populations, reflecting distant selection by pathogens, ethnic admixture, and drastic population reductions (genetic bottlenecks). LD is strongest between *HLA-B* and *HLA-C* and between *HLA-DR* and *HLA-DQ*, most likely because of their physical proximity. However, there is no LD between *DP* and the rest of the haplotype because of a hot spot of recombination between *DQ* and *DP*, even though these two loci are relatively proximal to each other.

The haplotype is the unit of inheritance of the MHC from either parent. Each parent shares one haplotype with each of the children and typically differs from a child by one haplotype. Two siblings may share two, one, or no haplotypes and thus range from being HLA identical, through haploidentical, to HLA disparate. Parents are usually only haploidentical with their children. Exceptions to this rule may occur in inbred populations, where both parents may share an identical HLA allele by descent. The HLA alleles originating from the maternal and paternal haplotypes are both expressed.

Ten years of genome-wide association studies (GWAS) have revealed a large number (884) of single nucleotide polymorphisms (SNPs) within the MHC that are associated with many (479) traits and diseases, establishing the MHC as the only region in the genome with this high density of SNPs that is associated with so many diseases.[3] The complexity of the region, with its many insertions, deletions, duplications, and LD, does not allow for an easy dissection of disease-causing variants. However, recent developments in next-generation sequencing (NGS) of the entire MHC[4] will most likely advance our understanding of the principles underlying this complex genomic organization, how this complexity results in so many biological interdependencies, and how it contributes to the pathophysiology of the diseases associated with the MHC.

STRUCTURE AND FUNCTION OF THE HLA MOLECULES

The main function of both class I and class II HLA molecules is to bind peptides derived from self or nonself antigens and

KEY CONCEPTS

Genomic Organization of the MHC

- The major histocompatibility complex (MHC) is the most complex genomic region in the whole human genome. It is associated with more diseases than any other genomic region of comparable size.
- The class I region contains the polymorphic human leukocyte antigen (*HLA*)*-A, -B,* and *-C* genes; the less polymorphic nonclassic class I *HLA-E, HLA-F,* and *HLA-G* genes; and the class I-related *MICA* and *MICB.*
- The class II region contains the *HLA-DR A* and *B, DQ A* and *B,* and *DP A* and *B* genes. It also contains the *TAP, LMP, DM,* and *DO* genes, which encode molecules that help process antigens into peptides that can bind class I and class II molecules.
- Genes within the MHC demonstrate extensive linkage disequilibrium. A string of alleles of polymorphic MHC genes that commonly exist in linkage disequilibrium within a given population is termed an MHC haplotype.
  - Haplotypes are preserved by means of natural selection, which acts to secure survival advantages for reproductive fitness within a given environment.
  - Common haplotypes within a given population appears to reflect functional interdependencies of the MHC gene alleles.
  - Haplotypes can be different in different populations.
  - The HLA genes of the two chromosomes (haplotypes) are coexpressed.

then traffic to the cell surface, where these peptides can be displayed, or presented, for recognition by the appropriate T cells. Their structure has evolved to satisfy this particular requirement.

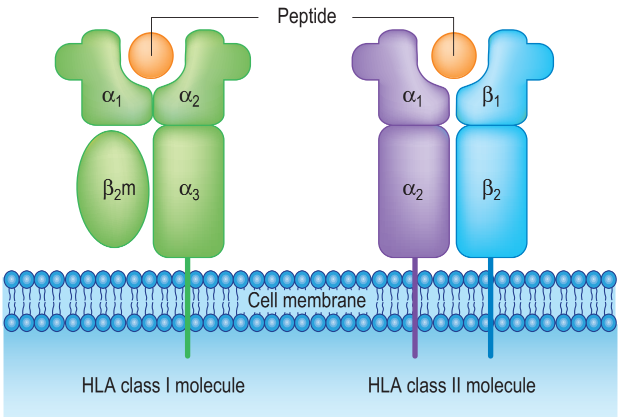


FIG 5.2 Human Leukocyte Antigen (HLA) Class I and II Domain Organization. Although HLA class I and class II proteins have a different chain structure, the organization of their domains is extremely similar. Both class I and II molecules are expressed on the cell surface, where they are accessible to T cells. Both have an outermost domain that contains a cleft where antigenic peptides are displayed. Two of the three class I α domains fold to create a domain with a peptide-binding cleft. The remaining α_3 domain helps support the peptide binding domain and anchors the molecular to the cell membrane. The class I molecule also contains an extrinsic β chain, β_2 microglobulin, which is encoded by a separate, invariant gene. β_2 microglobulin associates with the α_3 domain to support the antigen-binding domain created by the α_1 and α_2 domains. Class II molecules share a similar overall structure but are the product of two genes of variable sequence, each of which contains one α and one β domain.

Classic HLA Class I Molecules

The classic HLA-A, HLA-B, and HLA-C class I molecules consist of an α and a β chains. The α chain masses 45 kilodalton (kDa) and is 362-366 amino acids long. It is encoded by the respective class I genes within the MHC. The β chain, β_2 microglobulin (12 kDa), is encoded by its respective gene on chromosome 15. The α chain has three ≈90 amino acid extracellular domains encoded by exons 2, 3, and 4, respectively, a transmembrane segment (≈ 25 amino acids) encoded by exon 5 and a C-terminal cytoplasmic end (≈30 amino acids) encoded by exons 6 and 7. β_2 microglobulin, which is invariant, comprises the fourth domain (Fig. 5.2). The first two α domains (α_1 and α_2) are the most distal to the cell membrane. They combine to form a domain with a peptide-binding groove, or cleft, that is flanked by a surface that interacts with a TCR or a NK cell killer immunoglobulin-like receptor (KIR). The ends of the peptide-binding cleft are closed and fix the peptide's orientation. The sides of the peptide-binding cleft are composed of α helices, and the floor is composed of symmetric strands of β pleated sheet (Fig. 5.3). The α_3 domain and β_2 microglobulin are both members of the immunoglobulin superfamily (IgSF). Together they create a structure that supports the peptide-binding domain and, with the transmembrane domain of the α chain, attaches the molecule to the cell surface. Class I HLA molecules are ubiquitously expressed in all nucleated cells and in platelets. Expression of class I molecules is significantly reduced on red blood cells (RBCs) and absent on sperm cells.

HLA class I molecules bind peptides derived from the processed proteins of a pathogen or other self/nonself peptides (Chapter 6). These peptides average nine amino acids in length. Two or more of the amino acid side chains are used to anchor the peptide to pockets on the surface of the HLA class I molecule, with the NH_2 terminus to the "left" of the HLA-binding groove (see Fig. 5.3). Individual HLA class I alleles are generally distinguished by their own distinct pattern of peptide binding, as illustrated for selected HLA-B molecules in Table 5.1.[5] Among class I molecules, one or a few amino acid changes may considerably alter the binding properties of a binding pocket. In a healthy nonendocytosing cell, HLA molecules are filled with a variety of peptides from self molecules. The bound peptides are selected according to the binding motif of the particular allele. Even during viral infection or upon pathogen phagocytosis, the number of nonself peptides may not be high. Together, the MHC class I and its peptide create a complex ligand that serves as the target of the TCR on the T-cell surface. The expression of class I molecules is upregulated by the interferons (IFNs) IFN-α, IFN-β, and IFN-γ, granulocyte macrophage–colony-stimulating factor (GM-CSF) and certain other cytokines (Chapter 9). Class I expression is governed by a regulatory element that is located ≈160 nucleotides upstream from the initiation site of the class I gene. This site binds a number of regulatory factors, including those induced by IFNs.

Intact HLA-A, HLA-B, or HLA-C molecules are also ligands for KIR (Chapter 17).[6,7] KIR genes are located on chromosome 19. Their independent segregation from HLA genes located on chromosome 6 produces a wide diversity in the number and type of inherited HLA–KIR combinations. These combinations eventually influence immune competency and adaptability. KIRs do not interact with the whole top area of the HLA molecule that is normally recognized by the TCR. Instead, they interact with one end of the top of the molecule. All HLA-C alleles are

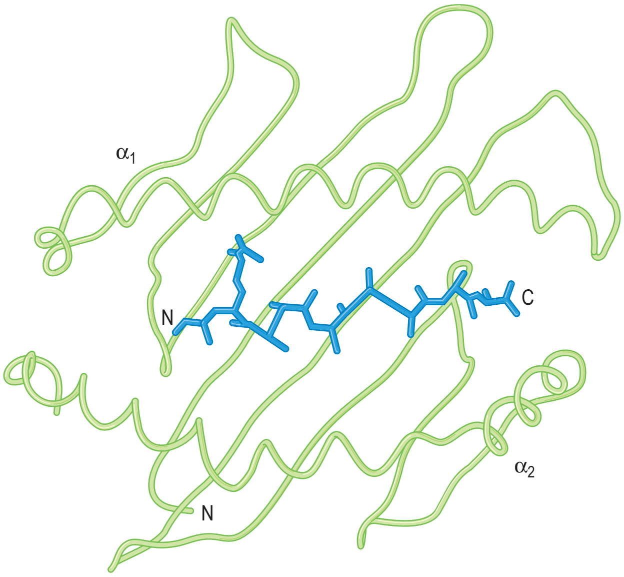


FIG 5.3 **The Three-Dimensional Structure of Human Leukocyte Antigen (HLA)-B27.** The α-helical margins of the peptide-binding cleft contain the bound peptide RRIKAITLK, which is oriented with its amino terminus to the left. There are extensive contacts at the ends of the cleft between peptide main-chain atoms and conserved HLA side chains. The peptide amino and carboxyl termini are tethered to the cleft by hydrogen bonds and charge interactions. The peptide reciprocally stabilizes the three-dimensional fold of HLA-B27. The positively charged side chain of arginine in the P2 position of the peptide inserts into the B pocket, which contains a complementary negatively charged glutamic acid at its base. The resulting salt bridge is the dominant anchor for the peptide. Side chains P4, P6, and P8 make minor contributions to the interaction of the peptide with the HLA-B27 molecule. The central region of the peptide is left free to interact with a T-cell receptor. [Modified from Madden DR, Gorga JC, Strominger JL, Wiley DC. The three-dimensional structure of HLA-B27 at 2.1 A resolution suggests a general mechanism for tight peptide binding to MHC. Cell 1992;70:1035.]

able to interact with their C1 (lysine) or C2 (asparagine) epitope (position 80 of HLA β chain). However, not all HLA-A and HLA-B alleles interact with KIRs. Only the few HLA-As and some HLA-B alleles that carry the Bw4 epitope can do so.

Upon interacting with their HLA ligands, the inhibitory KIR dampens NK-cell reactivity. In the presence of normal class I expression, NK cells will be inhibited from killing and the cell can function normally. However, if a cell loses classic class I expression, the NK cells are then controlled by their activating receptors that interact with their ligands (largely unknown) and perform their cell-killing function. This makes for a complicated dance involving both inhibitory and activating signals.

MICA and MICB

Within the class I region are *MICA* and *MICB* (HLA class I-related polypeptide sequence A and B). The products of these genes are more distantly related members of the class I family that neither associate with β_2 microglobulins nor bind peptides.[8] These molecules are expressed as “danger signals” by virus-infected or otherwise stressed cells. MICA and MICB are ligands for the activating NKG2D molecule (*KLRK1*), another member of the killer cell lectin-like receptor complex,[9] and appear on memory-effector T cells or NK cells, providing a signal to help activate their effector cytolytic response.

TABLE 5.1 **Peptide-Binding Motifs Encoded by Different HLA Alleles Influence the Number of Peptides in a Protein That can be Recognized by a HLA Molecule (*e.g.,* HIV Envelope Protein)**

Allele designation	HLA-B*27:05	HLA-B*35:01	HLA-B*07:02
Peptide-binding motif	XRXXXXXX[KRYL]	XPXXXXXXY	XPXXXXXXL
Peptides from the HIV envelope protein able to bind to each allotype	IRGKVQKEY IRPVVSTQL TRPNNNTRK IRIQRGPGR SRAKWNNTL LREQFGNNK FRPGGGDMR WRSELYKYK KRRVVQREK ARILAVERY ERDRDRSIR LRSLCLFSY TRIVELLGR CRAIRHIPR IRQGLERIL	None	DPNPQEVVL KPCVKLTPL RPVVSTQLL SPLSFQTHL IPRRIRQGL
Number of peptides bound	15	0	5

Single-letter amino acid codes are used. X denotes any amino acid; R, arginine; K, lysine; Y, tyrosine; L, leucine; P, proline, etc.

Nonclassic HLA-E, HLA-F, and HLA-G

The HLA nonclassic molecules E, F, and G are less polymorphic and have different functions and more limited tissue distribution compared with their classic HLA class I counterparts.[10]

HLA-E primarily presents self peptides to the TCR of CD8 T cells. The diversity of these self peptides is limited and includes the leader peptide of classic HLA class I molecules. The binding of HLA-E to inhibitory receptors, such as CD94/NKG2A, is an important part of the surveillance mechanism for missing self. In tumor cells, loss of class I expression results in a survival advantage for the particular tumor cell. In the absence of class I expression, HLA-E molecules will no longer be able to form a complex with the intracellular leader peptides of class I. As a result, HLA-E molecules are not expressed on the cell surface, and the inhibitory signals to the NK cells are removed. This licenses NK cells to kill the tumor target. Thus by selectively expressing self peptides derived from classic HLA class I molecules (leader sequences), HLA-E has evolved to be at the interface of innate and adaptive immunities.

HLA-F has a small binding cleft that does not contain peptide, and its functions are not well understood. It has limited polymorphism. It mainly resides intracellularly and rarely reaches the cell surface.

HLA-G has limited tissue distribution and is primarily expressed by placental trophoblast cells, the thymus, the cornea, and some erythroid and endothelial precursor cells. HLA-G has a peptide groove, binds a nonamer peptide, and is recognized as an MHC–peptide complex ligand by the leukocyte Ig-like inhibitory receptors (LIR-1 and LIR-2) and KIR receptors. In melanoma, HLA-G expression can be used by the tumor cells to avoid immunosurveillance by flooding the local microenvironment with soluble HLA-G and compromising the function of immune cells. The expression of HLA-G in chorionic villi suggests a role in the maintenance of pregnancy. The mechanism appears to involve production of soluble forms of HLA-G. They appear to have an inhibitory role on the immune cells of the mother. Uniquely among HLA molecules, HLA-G exists in different isoforms. Of these, four are expressed on the cell membrane, and three others exist as soluble forms. The functional significance of these isoforms is not known.

Classic Class II HLA Molecules

Classic class II HLA molecules are selectively expressed in cells of the immune system, similar to B cells, activated T cells, macrophages, dendritic cells (DCs), and activated T cells. The overall structure of class II HLA molecules is very similar to that of class I HLA molecules. The HLA class II molecules are also heterodimers that consist of two transmembrane glycoprotein α (34 kDa) and β (29 kDa) chains. Unlike class I, however, both the α and β chains are encoded by genes within the MHC. Each of the two chains is composed of two extracellular domains. *DR, DQ,* or *DP A1* include α_1 and α_2 domains that are encoded by exons 2 and 3 of the gene. *DR, DQ,* or *DP B1* include β_1 and β_2 domains that are encoded by exons 2 and 3 of the gene. The α_1 and β_1 domains form the binding groove of the class II HLA molecule and are highly variable. The single exception is the α_1 domain of DR, which not polymorphic. The α_2 and β_2 domains proximal to the membrane are members of the IgSF and have limited polymorphisms (see Fig. 5.2). Unlike class I, where the peptide-binding domain is encoded by α_1 and α_2 domains in the same gene, trans-arrangement of α and β chains derived from the two different haplotypes of the same or even different isotypes permit combinatorial polymorphism in class II.

Although the structure of the peptide-binding cleft in class II is homologous to that of class I, there are several distinct differences that have major functional consequences. Among the most important of these differences are those in length and cleft structure. The majority of peptides interacting with class II molecules have a length of >13 amino acids, whereas class I prefers peptides of nine amino acids. This is permitted in class II because, unlike class I, the binding cleft is open at the ends and the ends of the peptide can extend on both sides of the HLA molecule.

The peptide is bound to the class II molecule through the side chains of the peptide amino acids, which interact with five different polymorphic pockets within the cleft. Loading of the HLA class II molecules with peptides takes place primarily within the endosomes, where the HLA molecule interacts with endocytosed and phagocytosed extracellular antigens (Chapter 6). To prevent binding of intracellular peptides in the class II pocket, it first interacts with a protein called *invariant chain* (Ii) while the MHC molecule traffics through the endoplasmic reticulum (ER). The invariant chain is a trimer, and each of its subunits binds noncovalently with an HLA class II molecule. The MHC–invariant chain complex also interacts with another chaperone protein called *calnexin*. Upon release of calnexin, the class II molecule moves either directly into the late endosomal MHC class II compartment (MIIC) or is cycled to the cell surface, where it is then internalized into the MIIC. Once in the endosomal environment, invariant chain is degraded by proteases, including cathepsin S and L. It then leaves a fragment of peptide known as the class II–associated invariant chain peptide (CLIP). Upon dissociation of the CLIP peptide from the class II binding cleft within the endosome, the relevant exogenous peptide is associated with the class II molecule, as assisted by HLA-DM, prior to transport of the stable HLA class II–peptide complex to the cell surface.

Nonclassic HLA-DM and HLA-DO

The nonpolymorphic nonclassic class II molecules HLA-DM and HLA-DO are exclusively expressed in endosomes, and they regulate peptide binding to the classic HLA class II molecules. HLA-DM, a peptide editor, plays a central role in peptide loading of MHC class II molecules.[11] HLA-DO interacts with HLA-DM, but its expression is more restricted.

Proteosome Elements Within the Class II Region

The products of four genes in the class II region are involved with processing and loading peptides onto class I molecules (see Fig. 5.1). PSMB8 and PSMB9 are proteasome subunits generating peptides from the breaking down of proteins. TAP1 and TAP2 transport the peptides from the cytoplasm to the ER. The presence of these genes, which are related to the functioning of HLA class I molecules, in the midst of genes encoding the HLA class II molecules, is probably the reason we observe strong LD within the MHC. It appears that allelic forms of genes in the class I region require the presence of allelic forms in the class II region, indicating functional interdependencies developed throughout the evolutionary process and therefore the need for being transmitted together from generation to generation.

Principles of Peptide Presentation

The mechanism by which HLA class I and class II molecules present peptides became clear when the structures of these two molecules were determined. A simplified diagram of the domain structure of MHC class I and class II proteins is depicted in Fig. 5.2. A more intricate ribbon structure of the actual class I molecule interacting with the TCR is presented in Chapter 4. For both class I and class II, the peptide-binding structure takes the shape of a β pleated floor with two α helix walls. The peptide lies within the groove created by these structures (see Fig. 5.3; Fig. 5.4).

Each HLA molecule, whether class I or class II, binds a single peptide; but the same HLA molecule has a significant degree of promiscuity and can bind thousands of different peptides. Each of the binding grooves is composed of individual polymorphic pockets that dictate the binding of different peptides. Although the mode of TCR docking on HLA molecules is globally conserved, the shapes and chemical properties of the interacting surfaces found in these complexes are so diverse that no fixed pattern of contact has been recognized even between conserved TCR residues and conserved side chains of the HLA α helices.[12] Indeed, of the amino acid side chains not bound to the HLA, only two or three are typically bound to the clonotypic TCR. This limited contact yields considerable TCR plasticity, which has the important evolutionary implication of freeing the HLA molecule and the peptide–HLA complex from the strict stereochemical constraints that are usually imposed in receptor–ligand interactions. The consequence of TCR plasticity and this unusual receptor–ligand interaction has been the evolutionary development of a uniquely large number of different genes that encode various HLA structures, each of which is able to bind and present a different range of peptides to the same clonotypic TCR.

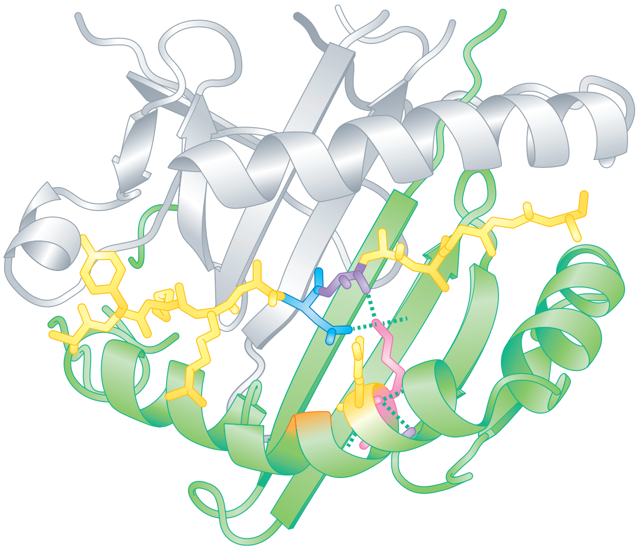


FIG 5.4 Structure of a Human Leukocyte Antigen (HLA) Class–II Peptide Complex. The structure was prepared using PyMol from published coordinates. The HLA molecule is largely shown as a ribbon, while the peptide is a stick diagram. The peptide-binding groove is delimited by α helices. The upper helix is encoded by the α chain, and the lower helix by the β chain. β pleated sheets form the saddle-like floor. Side chains are depicted on the β chain at positions 70 and 71, a region involved in specifying the side chain pocket P4. This pocket binds the fourth side chain of the peptide contained within the HLA molecule. The side chains shown are respectively glutamine and lysine, which form part of the "shared epitope" structure associated with susceptibility to rheumatoid arthritis. The lysine is shown forming hydrogen bonds with the peptide antigen.

KEY CONCEPTS

Structure of the Human Leukocyte Antigen (HLA) Molecules

- HLA class I molecules are involved in both innate and adaptive immunity.
  - For innate immunity, natural killer (NK) cell functions are influenced by the binding of killer immunoglobulin-like receptor (KIR) to class I genes.
  - For adaptive immunity, they enable T cells to identify antigenic peptides.
- Class I and class II HLA molecules have similar structures that enable them to bind peptides and present them to T-cell receptors (TCRs).
- Peptide binding to the HLA molecule is influenced by allele-specific pockets within the binding cleft of the HLA molecule, which interact with the amino acid side chains of the peptide antigen.
  - The HLA–peptide complex is recognized by the TCR.

Selection by Self Peptides in the Thymus

Peptides derived from external antigens, including pathogens, are typically absent during the formation of the T-cell repertoire in the thymus (Chapter 8). Thus self peptide–HLA complexes are used as surrogates for selecting, or training, individual T cells to recognize nonself pathogen peptides.[13] For T cells, "immunological self" is the set of self peptides and self MHC molecules that select the TCR repertoire in the thymus. This process of selection occurs primarily in two steps. The first step is a positive selection process, whereby only those cells with TCRs interacting with the self peptide–HLA complex survive. T cells with receptors that do not recognize any self peptide–HLA complex are eliminated. The second step is a negative selection process, whereby among the selected T cells with self recognition, those with high affinity interactions with the self peptides–HLA complex are eliminated allowing the rest with lower affinity interactions to survive and be released in the periphery. These self peptides constitute the T-cell recognition component of an individual's adaptive immune system. This patterning of TCR recognition on self peptides presented by self MHC molecules is critical to the development of autoimmunity and allorecognition.

Evolutionary Considerations Driving the Separate Functions of Class I and Class II

One basic task of the T cell is to protect the body from two major types of pathogens: *viruses,* which would commandeer the replicative machinery of a cell, and *bacteria,* which replicate autonomously and often extracellularly. These two types of pathogens present very different challenges to the immune system. To terminate viral infection, a cell harboring a virus has to be killed by a cytotoxic CD8 T cell, whereas a bacterium can be eliminated by being phagocytized by a macrophage that has been selectively activated by a CD4 T helper (Th) cell. The need to determine whether the presence of a pathogen peptide should elicit a killer-cell response or a Th-cell response is presumed to be the evolutionary drive that resulted in the creation of two specialized forms of HLA molecules, class I and class II (see Fig. 5.2).[12,14]

The specialized antigen processing and presentation intracellular machinery used to load class I molecules offers a means for the cell to reflect at the cell surface the molecular profile of antigens within the cell. This allows class I to sample for the presence of an intracellular viral infection. Recognition of the HLA class I–peptide complex is through the TCR of a CD8 T cell, which primarily reacts to the detection of an inappropriate intracellular antigen (*i.e.,* a virus) by cytotoxic activity. Class II peptide loading occurs in coordination with phagocytosis and lysosomes. Class II thus offers a means by which the immune system can be informed of the presence of extracellular antigens, such as bacteria. The recognition of the HLA class II–peptide complex is through the TCR of a CD4 T cell, which leads to the activation of Th cells and then to activation of an immune response.

Among the evolutionary strategies used for viral survival, some virally encoded genes decrease the expression of the HLA class I surveillance system, which would otherwise alert the immune system to the presence of an infected cell (Chapters 17 and 25).[15] This attempt to escape surveillance by downregulation of HLA class I is countered by the extensive interaction of class I molecules with various NK receptors expressed on NK cells or T-cell subsets. These interactions provide a mechanism for detecting decreases in HLA class I expression, which is termed recognition of "missing self."[16] There are two principal types of NK receptor used to detect missing self, members of the IgSF, such as KIRs, which bind directly to intact class I MHC molecules, and members of the C-type lectin family, such as CD94/NKG2C, which recognize the leader of class I molecules that selectively binds to HLA-E molecules. Considering that NK cells are cellular components of the innate immune response, this means that HLA class I molecules play a role in both adaptive and innate immune responses.

KEY CONCEPTS

Function of the Human Leukocyte Antigen (HLA) Molecules

- The HLA molecule is a receptor that binds a peptide. The peptide–HLA complex is a ligand that binds to the clonotypic T-cell receptor (TCR). The trimolecular Peptide–HLA–TCR complex triggers the activation and proliferation of the T cell in an adaptive immune response.
  - HLA class I A, B, and C molecules are expressed on the surface of virtually all nucleated cells.
  - HLA class II DQ, DR, and DP molecules are constitutively expressed on B cells, professional antigen-presenting cells (APCs), thymic epithelial cells, and activated T cells.
- The *immunological self* is the set of self peptides and self HLA molecules that select the TCR repertoire in the thymus and that constitute the T-cell recognition component of an individual's adaptive immune system.
- Nonself peptide–HLA complexes are recognized by T cells during an adaptive immune response, and they become activated to either initiate an immune response (CD4 helper T cells) or recognize a target (CD8 cytotoxic T cells).
- Through thymic selection, the TCR can adapt to recognize a very large variety of peptide–HLA structures.
  - Because of the plasticity of this recognition, genes encoding the HLA molecules are free to evolve a large number of genes encoding duplicated or alternative peptide-presenting molecules with specificity to bind different peptides.
  - The diversification of peptide-presenting structures fosters the development of different T-cell repertoires with completely different recognition properties. This thwarts the possibility that a pathogen will be able to evolve a way to bypass recognition.

GENERATION AND SELECTION OF POLYMORPHISMS: BIOLOGICAL CONSEQUENCES

The hallmark of the classic HLA molecules, both class I and class II, is their extensive polymorphism. HLA polymorphism observed in different human populations is far greater than any other polymorphism observed in any other part of the human genome. This is a direct reflection of their role in the immune response. Pathogens characterized by different proteins and peptides, either in different epidemics or endemic to regions, account for much of the evolutionary drive responsible for the large number of alternative gene forms and their regional diversity across the human race. An individual with an adaptive immune system based on HLA molecules that effectively bind peptides derived from common pathogens is much more likely to have an effective response against that common pathogen. This results in selection of individuals with a particular allele, encoded by an HLA gene.

Genetic polymorphism implies that alleles of a gene are present at a frequency greater than expected from random mutation as a result of selection for diversity. In the case of the HLA genes, there is no preponderant wild-type allele, which would be an example of *balancing selection.* Instead, virtually all alleles qualify as genetically polymorphic. These reflect prior successful selection events. HLA polymorphisms provide a major evolutionary survival benefit, since they equip the species with a large number of very specific, but alternative, HLA molecules that differ in their binding pockets, are most efficient in presenting different peptides, and selecting different T-cell repertoires. A polymorphism that offers survival advantage would eventually increase in frequency. This illustrates *frequency-dependent selection,* where the fitness of individuals bearing a particular allele increases, if they can manage an effective immune response to the particular pathogen. Moreover, selection also operates on the pathogen, encouraging peptide variation. Variation in peptides drawn from common pathogens and the introduction of novel pathogens with novel peptides result in pressure on the species to create variation in HLA molecules among individual members of that species. The remarkably different frequency of the HLA alleles in different ethnic subsets tells the history of the successful adaptation of our ancestors' adaptive immune systems to a new environment with different pathogens, as well as bottlenecks resulting from migration and perhaps survival during periods of massive epidemics.

The evolutionary consequence of the diversification of genes encoding HLA molecules is seen at two levels. The first is at the level of the *individual* and is characterized by the presence of different HLA class I and class II loci, each of which codes for one or two different peptide-presenting HLA molecules for each locus. The second is at the level of the *population* and is evidenced by the development of a very large number of alleles at each locus, with each allele coding for alternative polymorphic gene forms and thus for various peptide-presenting allotypes, each of which has the potential to bind a different set of peptides. Duplication of HLA genes involved in peptide presentation is a genetic strategy that increases the range of peptide-presenting structures available to the individual, thus enhancing the variety of presented peptides that can be recognized and bound.[8]

KEY CONCEPTS

The Biological Significance of Polymorphisms: Why so Many?

- Human leukocyte antigen (HLA) class I and class II genes are extremely polymorphic.
  - Each HLA allele encodes molecules with different peptide-binding properties that influence the particular peptides recognized by the T cells.
  - The sequence of the HLA gene thereby determines the peptide recognition features of the adaptive immune response.
- HLA allelic polymorphisms are maintained by frequency-dependent selection, where the fitness of an individual bearing a novel allele increases because it can respond more effectively to certain pathogens.
- The multiple loci and numerous alleles per locus serve both the fitness of an individual and the survival of the species.
  - The polymorphism of the HLA system reflects the environmental/pathogen challenges to which a particular population has been exposed over evolutionary time.

HLAS IN INFECTIONS, TRANSPLANTATION, AUTOIMMUNITY, AND CANCER

HLA in Infections

The first line of defense during infection by a pathogen is the triggering of innate immunity (Chapter 3). The infectious agent and the foreign peptides generated from that agent then initiate an immune response involving immune cells and signals that subsequently induce adaptive immunity.

During the course of an infection, specialized antigen-presenting cells (APCs; DCs and macrophages) are activated to take up antigen. Increasing synthesis of class II coupled with presentation of and pathogen's peptides by class II to the immune system of the host drives CD4 T cells that recognize the HLA class II–peptide complex and become activated. This event triggers adaptive immunity. Eventually CD8 T cells recognize target cells infected with the pathogen by interacting with the HLA class I–peptide complex on their cell surface and the targets are eliminated, thus containing the infection (Chapter 6).

In response, pathogens have evolved mechanisms to overcome the specific attack by the host's immune cells. The first of these mechanisms is "antigenic drift" or "antigen shift," whereby the pathogen, through minor (drift) or more substantive changes (shift), evades both humoral and cellular responses. These changes make the pathogen unrecognizable, as some of these new peptides do not form recognizable complexes with the HLA molecules of the host and therefore evade the T-cell responses. Another mechanism frequently adopted by viruses is to persist *in vivo* by not replicating until the immunity of the host is compromised. By not replicating, they avoid detection and exist in a dormant state (latency). It therefore becomes evident that the infectivity of a microorganism reflects the interplay between several complex processes. These include the ability of the pathogen to create new molecular forms unrecognizable by the host and thus evade detection. These efforts by the pathogen to avoid immunity are then counterbalanced by molecular polymorphisms between HLA molecules that enable recognition of new molecular forms of the pathogen.[17]

HLA in Transplantation

The large number of different HLA alleles greatly reduces the probability that two unrelated individuals will inherit an identical set of HLA alleles. Two basic mechanisms of responses in transplantation have been described. The first involves the "direct" recognition of the peptide–HLA complex of the donor tissue by the T cells of the recipient. This is possibly through structural similarities of the HLA molecules of the donor that allow the TCR of the recipient to interact with the peptide–HLA complex. The second involves the "indirect" presentation of donor's HLA antigens processed by the recipient's APCs, generating peptides presented by the recipient's HLA molecules to the recipient's T cells. This indirect mechanism operates the same way as the presentation of a foreign antigen, whereby the HLA molecule is now the foreign antigen processed by the antigen-processing mechanisms of the recipient.

By using appropriate immunosuppressive agents and therapies, T-cell activation by the donor's HLA molecules after clinical transplant can be controlled (Chapter 81). However, the major long-term problem is the presence of donor-specific antibodies that develop against mismatched HLA antigens. Controlling the antibody responses to mismatched HLA molecules has been very challenging, and there is a need for continuous monitoring of their development. An approach holding some promise for the future is the utilization of regulatory T cells (Tregs), which have important immunoregulatory role in all immune responses and can possibly induce transplant-specific tolerance (Chapter 18).

HLA in Autoimmunity

Selection on the self peptide presented by self HLA allotypes in the thymus can predispose to autoimmunity because the T-cell portion of the adaptive immune system is entirely selected on the self peptide. The inherent autoreactivity of the T-cell system can thus set the stage for the development of autoimmune diseases associated with the recognition of particular self peptides, or peptides from external antigens that mimic these self peptides and are effectively presented by self HLA.[18] Certain alleles encode HLA molecules that bind peptides from molecules expressed in sites favoring autoimmune recognition by T cells. These molecules become the target of the adaptive immune response. Together, features specific to certain sets of self peptides and to certain self HLA molecules can contribute to the progressive development of autoimmunity and, ultimately, autoimmune disease (Chapter 50).

HLA in Cancer

Immune evasion is a critical process in tumor biology. It is enabled by several mechanisms that include immune editing, downregulation of HLA expression, secretion of immunosuppressive mediators, and expression of proteins that modulate immune checkpoints. Most recently, somatic mutation of HLA genes was revealed to be a significantly frequent process in some tumor types. The strategies of immune evasion by cancer cells also include the silencing or aberrant expression of HLA class I and class II molecules, events that have often been associated with high-grade malignancy and metastatic potential in a variety of human cancers.[19]

In patients with solid tumors, HLA-G can contribute to a tumor-escape mechanism that favors cancer progression, and blocking strategies have been proposed to counteract it. Conversely, HLA-G can inhibit proliferation of malignant B cells as a result of the interaction between HLA-G and its receptor ILT2, which mediates negative signaling on B-cell proliferation. Thus treatment of some malignancies can benefit by blocking HLA-G, whereas in others HLA-G induction can counteract tumor progression.[20]

The concept of developing cancer-specific immunotherapies involving tumor-specific antigens presented by HLA molecules to T cells has been successfully tested in a number of tumors (Chapter 77), including testicular cancer and melanoma. These T-cell immunotherapies require adoptive transfer of T cells that have been expanded *ex vivo* and transferred back to the patient. Another approach is the use of retroviral vectors to transfer tumor-specific TCR genes into the patient's T cells before reinfusion.[21] Even though HLA molecules are involved in these processes, histocompatibility testing is not necessary in these therapies because the original T cells are derived from the patient. However, if the mechanism of immunotherapy involves neoantigens (epitopes of mutated proteins) from tumors presented by specific HLA alleles, such individualized therapy needs to take HLA alleles into account.

HLA AND DISEASE ASSOCIATIONS

Over the last several decades, a large number of studies have established strong associations between certain diseases and individuals carrying particular HLA alleles. In spite of extensive study, the mechanisms underlying HLA–disease associations remain unclear.

Hypotheses generated to explain these associations can be grouped into two general categories. The first category invokes linkage disequilibrium between a particular HLA allele that is associated with a given disease and another neighboring genomic element on the haplotype that is actually causative of the disease and does not involve HLA molecules directly. This can occur because the genes within the MHC are in extensive LD among themselves. Examples of this type of associations include hereditary hemochromatosis, where an apparent association with HLA-A alleles results from mutations in a nonclassic HLA class I gene,

KEY CONCEPTS

Human Leukocyte Antigen (HLA) in Infections, Transplantation, Autoimmunity, and Cancer

- HLA and infections by pathogens is an interplay of balancing acts, whereby the pathogens try to avoid the immune response and HLA alleles adapt to secure a robust immune response.
- Transplantation is an artificial system, and the transplant is perceived by the immune response to be a foreign element.
  - Induction of tolerance is the objective.
  - For the host to tolerate the graft, physicians attempt to engender immune nonresponsiveness to the transplant by manipulating the immune response pharmacologically.
  - The presence of donor-specific antibodies is a major problem in the field of transplantation because these antibodies are primarily responsible for chronic rejection reactions we observe.
- There are three features of the adaptive immune system that can set the stage for pathogenic autoimmunity.
  - First, the individual's adaptive immune system is determined by the set of self peptides and self HLA molecules that select the T-cell receptor (TCR) repertoire.
  - Second, the drive to genetic polymorphism that generates many different alternative forms of peptide-binding HLA molecules influences patterns of self and nonself reactivity.
  - Third, certain HLA allotypes bind particular self peptides from critical target antigenic molecules that can predispose to autoimmune responses and disease.
- Oncogenesis is associated with the modification of patterns of antigen presentation by the cancerous cells to immune cells and by modification of immune cell responses to the cancerous cells.
  - This enables the cancer to avoid immune surveillance and detection by the immune response.

HFE, which is in LD with *HLA-A*; and congenital adrenal hyperplasia resulting from an allele of the gene *CYP 21B,* which causes 21-hydroxylase deficiency and is located within the class III region of the MHC and thus in LD with *HLA-B*.

A second category implicates antigen presentation by the HLA allele. This category deals with diseases that have a strong immunological component. It has been hypothesized that inappropriate immune reactivity to some self antigens can reflect aberrant T-cell repertoire selection, immune cross-reactivity with foreign antigens, immune attack of "altered self" antigens, or differences in the expression levels of certain HLA alleles that secondarily influence the course of infections or cancer. The MHC cusp theory represents another hypothesis. In this case, the MHC codes for allele-specific ligands in the cusp region of the molecule, which interact with non-MHC receptors and activate various pathways. Aberrations in these pathways could cause MHC-associated diseases. According to this hypothesis, the cusp region has a peculiar three-dimensional shape that has been preserved on both class I and class II molecules through evolution, not dependent on antigen presentation, and is a hub for signal transduction ligands that interact with a variety of receptors and activate important biological functions. The MHC cusp theory posits that HLA molecules promote disease because of their auxiliary allele-specific, yet antigen presentation–independent, biological effect.[22]

Although many of these associations lie within the highly polymorphic HLA genes,[1] GWAS using SNP markers have established that not only HLA genes but also the MHC region as a whole harbors many SNPs associated with a large number of traits or diseases. Indeed, up to 90% of autoimmune disease variants have been located within the noncoding regions of the genome.[23] It is therefore possible that disease association elements may not only lie within the HLA genes but also be dispersed within the rest of the MHC. One possible genomic element that can be located within the noncoding regions of the MHC and yet have a significant regulatory role is microRNA (miRNA). A search for functional genomic elements within the noncoding regions of the MHC genes revealed 12 miRNAs, including hsa-miR-6891 (miR-6891), which is encoded by intron 4 of HLA-B.[24] Thus some, and perhaps many, diseases associated with specific MHC elements, whether HLA alleles or not, may involve noncoding RNAs (miRNAs or long noncoding RNAs) with important biological functions of a regulatory nature.

Below is a compilation of selected diseases with strong HLA allele associations in different populations. A more extensive list of diseases and reference materials can be found in other works.[25-27]

Ankylosing Spondylitis

One of the more extraordinary observations in the MHC field was made in 1973 when the frequency of HLA specificity HLA-B27 was found to be 95% in those with the disease ankylosing spondylitis (AS) (Chapter 57). This impressive observation implicated HLA-B*27 in the pathogenesis of AS and propelled the field of HLA and disease associations.[28] Different B27 alleles have different strengths of association with the disease, making genetic testing useful over and above serological testing. Even though the association of AS to B27 is among the strongest genetic associations with a common disease, the mechanism of action remains uncertain. AS has been hypothesized to be triggered by exposure to a common environmental pathogen. HLA-B*27:02 and B*27:05 demonstrate the highest degree of association. AS is characterized by arthritis affecting the spine and the pelvis. Twin studies have confirmed that susceptibility to AS is genetically determined. Family studies suggest that <50% of the overall genetic risk is caused by HLA-B27. HLA-B27 is found in 8–10% of the population, with only a minority of carriers progressing to development of the disease. It is likely that other genes, both within and outside the MHC, are involved. A number of GWAS have demonstrated that non-HLA genes are also associated with AS. These include the interleukin-23 receptor (IL-23R) and the protein-cleaving enzyme ER aminopeptidase 1 (ERAP1).

B27 testing can be an instructive component of the diagnostic work up of AS. Because of the chronic nature of the disease and its gradual debilitating nature, the value of B27 testing is that a presumptive diagnosis allows institution of treatment early in the disease when patients may have minimal symptoms.

Narcolepsy

Narcolepsy is a long-term neurological disorder characterized by irresistible daytime sleep attacks. People with narcolepsy suffer from episodes in which they fall asleep unexpectedly during the day. These "sleep attacks" can occur at any time and during any activity. Narcolepsy affects approximately 1 in 2000 people. Often those affected have low levels of the neurotransmitter *hypocretin* (also known as *orexin*). Hypocretin is a neuropeptide hormone that is responsible for controlling appetite and sleep patterns. Even though the cause of narcolepsy is unknown, the disease is believed to be of autoimmune nature.

Family studies have shown that genetic heritability plays a role in narcolepsy. However, twin studies show that only 25–30% of twins are concordant for the disease, again implicating environmental or other epigenetic events. The HLA-DQB1*06:02 allele on the DRB1*15:01–DQA1*01:02–DQB1*06:02 haplotype has been shown to be one of the most important predisposing genetic factors, with 85–95% of patients with narcolepsy carrying this haplotype.[29] Conversely DQB1*05:01 and DQB1*06:01 have a protective effect. The protective associations of these two DQB1 alleles with narcolepsy may provide an insight to the molecular mechanisms for the differential associations of DQB1*06:01 and DQB1*06:02, as the size of P4 pocket of DQB1*06:02 is larger than the DQB1*06:01. This difference possibly influences the binding of larger residues in the DQB1*06:02 allele, which may explain the opposite effect these two alleles have on narcolepsy. Homozygosity for HLA-DQB1*06:02 increases the risk for narcolepsy compared with heterozygosity, as does heterozygosity for HLA-DQB1*03:01/DQB1*06:02.

HLA testing for DQB1*06:02 in narcolepsy is a useful aid to diagnosis. However, as instructive as the association may be, it is not specific as there are many patients with narcolepsy without HLA-DQB1*06:02 and many individuals with HLA-DQB1*06:02 who do not have narcolepsy.

Type 1 Diabetes

Type 1 diabetes (T1D) is also known as *insulin-dependent diabetes mellitus* (IDDM; Chapter 71). This is a disease in which the body fails to maintain normal glucose levels because of the destruction of insulin-producing pancreatic islet cells. The disease is characterized by infiltration of immune cells (CD4 and CD8 T cells) into the islets of the pancreas and by autoantibody production. When >90% of an individual's β cells are destroyed, clinical symptoms ensue.

Twin studies have shown that the concordance rate for the disease is 30–50%. This suggests that other factors, including environmental triggers (*e.g.,* diet and viral infections) and epigenetic changes, may be involved. The major heritable risk of T1D comes from the HLA system (about 50%). More than 90% of Caucasian patients with T1D carry the haplotypes DRB1*03:01, DQA1*05:01, DQB1*02:01 or DRB1*04:01, DQA1*03:01, DQB1*03:02. Patients heterozygous for these haplotypes carry a greater susceptibility risk. The critical residues are thought to be position 52 on the DQα chain and position 57 on the DQβ chain. The presence of arginine of 52 at DQα and the absence of aspartate at DQβ are strongly associated with T1D. Conversely, in Caucasian populations, resistance to T1D is conferred by DQA1*01:02, DQB1*06:02. Besides the contribution of the HLA, GWAS have identified a number of other genomic regions associated with the development of T1D.[30]

HLA typing is useful as an aid to diagnosis of T1D. Considering that the islet destruction by the autoimmune processes is progressive, associated with the presence of autoantibodies, HLA typing of siblings of patients with T1D may be useful information in assessing risk for the nonsymptomatic siblings.

Rheumatoid Arthritis

Rheumatoid arthritis (RA) is a chronic disease characterized by inflammation of the synovial lined joints leading to joint deformation and disabilities (Chapter 52). The presence of autoantibodies, such as rheumatoid factor and anticitrullinated protein antibodies (ACPAs), are largely responsible for the classification of RA as an autoimmune disease. It is a multifactorial disease that involves both environmental and genetic factors. RA in the general population has a prevalence of <1%. Studies with monozygotic twins show a 12–15% concordance rate for the disease, underlying both genetic heritability, but also other factors, such as environmental triggers or epigenetic components. Multiple genetic loci have been shown to contribute to the risk of developing RA. Of these, the HLA class II DRB1 is the most important and contributes 30–50% of the overall genetic susceptibility risk.

The HLA DRB1 alleles associated with RA share common sequences at positions 70–74 of the β chain.[31,32] This has led to the shared epitope hypothesis. Amino acids in these positions influence both peptide binding and contact between the HLA and the TCR. HLA-DRB1 alleles associated with RA have any of the following sequences: QKRAA, QRRAA, RKRAA, and RRRAA. In Fig. 5.4, the yellow-colored residue in the α-helical ribbon is glutamine, and the magenta residue is positively charged lysine. Hydrogen bonding to two side chains of the peptide is shown. The region around position 70 is involved in the formation of a peptide side chain–binding pocket that binds the fourth amino acid side chain contained within the HLA molecule. The presence of a negatively charged residue at position 71 or 74 removes susceptibility for RA. The presence of two alleles of this group increases susceptibility and favors development of more severe disease.[33]

In addition to the MHC, GWAS have led to the identification of over 100 loci associated with RA. Among these is the protein tyrosine phosphatase, nonreceptor type-22 (*PTPN22*) gene, which codes for an inhibitor of T-cell activation. The majority of these additional loci are expression quantitative trait loci (eQTLs), in which genetic variants regulate the level of transcription.

Multiple Sclerosis

Multiple sclerosis (MS) is a complex neurodegenerative disease in which myelin sheath degradation is caused by the immune system (Chapter 66). On the basis of family and twin studies, the disease has been shown to have a large genetic component. HLA-DRB1*15:01, DQA1*01:02, DQB1*06:02 is a disease susceptibility haplotype and accounts for up to 35% of the risk for developing the disease. The primary association is with the DRB1 locus, and homozygosity is associated with increased risk. A number of GWAS have identified more than 100 additional candidate genomic regions conferring risk, including cell adhesion, leukocyte activation, apoptosis, Janus kinase (JAK)–signal transducer and activator of transcription (STAT) signaling, nuclear factor (NF)-κB activation, and T-cell activation and proliferation.[34] The exact mechanism by which HLA genes increase susceptibility to MS is as yet unknown. Although HLA typing is not of any diagnostic value for MS, genetic testing is still of value in providing insights into the mechanism of the disease.

Celiac Disease

Celiac disease (CD) is an autoimmune disorder of the small intestine caused by a combination of genetic and environmental factors (Chapter 75). The disease is characterized by diarrhea and weight loss, among other symptoms. The strong genetic component has been confirmed with monozygotic twins demonstrating 90% concordance.[35] A significant portion of the genetic predisposition (about 40% of genetic risk) comes from HLA genes. GWAS also implicate additional genomic regions. The disease-triggering environmental factor comes from a component of wheat gluten, the protein gliadin (family of closely related proline- and glutamine-rich proteins). CD is a lifelong condition, with the only effective treatment being a gluten-free diet.

The implicated HLA molecules are the class II antigens DQ2 and DQ8. The DQ2 molecule mostly associated with CD is encoded by the HLA-DQA1*05:01–DQB1*02:01 alleles, with a small proportion encoded by the DQA1*02:01–DQB1*02:02 genotype. The DQ8 molecule associated with CD is DQA1*03–DQB1*03:02. Approximately 90% of patients with CD express the HLA-DQ2 molecules, with the remaining 10% mostly expressing the HLA-DQ8 molecule. Deaminated by transglutaminase, negatively charged gluten peptides bind strongly to HLA-DQ2 and -DQ8 to present an HLA–gluten peptide complex that activates CD4 T cells. The immune response also includes the development of antibodies against gluten and autoantibodies to endogenous tissue transglutaminase.

Genetic testing for HLA-DQ as a complement to histology can help confirm the diagnosis in patients not known to be positive for tissue transglutaminase antibody.

DRUG HYPERSENSITIVITY AND PHARMACOGENOMICS

Severe cutaneous adverse reactions to drugs include syndromes, such as Stevens-Johnson syndrome/toxic epidermal necrolysis and drug reaction with eosinophilia and systemic symptoms or drug-induced hypersensitivity syndrome[36] (Chapter 48). Although their incidence is very low, they are severe, life-threatening adverse drug reactions with mortality rates as high as 5–12.5%. The associations reported between drug hypersensitivity and specific HLA alleles has been a recent finding and has led to the possibility that hypersensitivity reactions may be predictable and preventable. Drugs associated with immunologically mediated drug-induced hypersensitivity include the anticonvulsant carbamazepine and the antiretroviral agents nevirapine and abacavir. Regulatory agencies, such as the US Food and Drug Administration (FDA), have issued relevant and informative pharmacogenomics guidelines: (http://www.fda.gov/Drugs/ScienceResearch/ResearchAreas/Pharmacogenetics/ucm083378.htm).

Carbamazepine

Carbamazepine is an aromatic amine anticonvulsant, used for the treatment of epilepsy and other seizure disorders, trigeminal neuralgia, and bipolar disorder. Approximately 10% of patients develop mild cutaneous adverse reactions. Carbamazepine has been shown to be associated with the HLA class I allele HLA-B*15:02 or A*31:01.

Nevirapine

Nevirapine is a nonnucleoside reverse transcriptase inhibitor, widely prescribed for human immunodeficiency virus 1 (HIV-1) infection. Hypersensitivity HLA class I and II associations have been described for DRB1*01:01, B*35:05, Cw8, and B*14:02.

Abacavir

Abacavir belongs to the family of nucleoside reverse transcriptase inhibitors and is used for the treatment of HIV-1 infection. Two recent abacavir studies have shown that 100% of patients who develop abacavir drug hypersensitivity carry the HLA-B*57:01 allele. This predictive value supports the use of HLA-B*57:01 typing of patients prior to initiating treatment with abacavir, even though not all of HLA-B57:01-positive patients develop hypersensitivity.

KEY CONCEPTS

Human Leukocyte Antigen (HLA) and Disease Associations

- HLA molecules are associated with many diseases.
  - HLA alleles frequently confer a higher risk for a number of immune related diseases than other genomic factors.
  - Most associations reflect situations where the HLA molecules are directly involved in the disease process.
  - Some associations reflect linkage disequilibrium with other non-HLA genes that are directly involved and responsible for the disease phenotype.
  - In some cases, the HLA molecule, the associated peptide, and the T-cell receptor (TCR) are sufficient for the development of disease.
  - In others, the HLA molecule may be necessary, but not sufficient for the development of the disease.
- Twin studies have demonstrated that genetics is not the only component of many of HLA-associated diseases and that environmental of metagenomic modifications are also likely involved in the disease process.

METHODS OF DETECTING HLA POLYMORPHISMS: HLA TYPING

Since the discovery of the HLA genes over 50 years ago, there has been a concerted effort to properly categorize and characterize these very polymorphic genes. Our understanding of the complexity and polymorphic nature of the HLA genes has been substantially improved as the technologies for characterizing these genes have improved (Fig. 5.5).

Initial serological and cellular testing in the 1960s (antibody and mixed lymphocyte culture [MLC]), supplemented by two-dimensional electrophoresis and restriction fragment length polymorphism (RFLP) analysis in the 1970s to the 1980s, made us aware of the high degree of polymorphism that was not properly revealed by the technologies used earlier. The development of the polymerase chain reaction (PCR) in the mid-1980s revolutionized our understanding of these genes at the molecular level. From PCR, methods utilizing sequence-specific oligonucleotide probes (SSOPs, or SSO) and sequence-specific primers (SSPs) provided the means for more directly evaluating the highly variable sequence motifs within the HLA genes. Subsequently, Sanger sequence-based typing (SBT) in the 1990s significantly advanced tissue typing and transplantation genetics by providing an unprecedented molecular view of HLA polymorphism in the context of exonic variation. Most recently, NGS appears to have revolutionized the field by addressing the HLA typing complexity in a very definitive way. NGS provides entire HLA gene characterization and haploid sequence determination.[37]

To meet the growing demand, clinical HLA typing over the past decade has transitioned from a combination of serological and DNA-based methods to more direct, faster, more affordable, and more informative DNA-based techniques. Even though serological typing may continue to have some clinical or research-based testing in determining the expression of the HLA molecule at the cell surface (a function that DNA-based testing cannot always verify), direct DNA-based typing techniques have all but replaced serological methods in routine HLA typing.

DNA-Based Typing Techniques: SSO, SSP, and SBT

The techniques primarily in use today in clinical immunogenetics laboratories are SSO, SSP, and SBT. The genomic regions analyzed

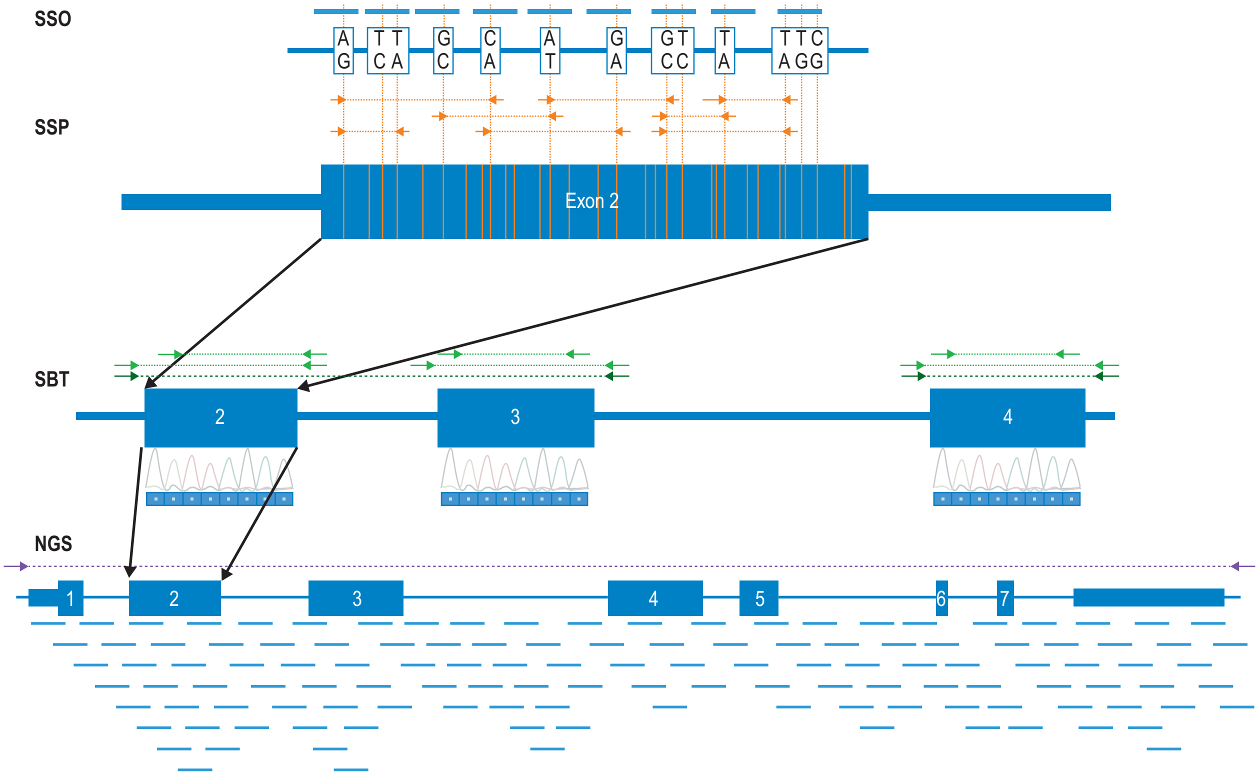


FIG 5.5 **Examples of Molecular Human Leukocyte Antigen (HLA) Typing Techniques and Their Methods of Interrogating the HLA Genes.** For any given HLA gene (dark blue rectangles), SSOs of ≈20 bp (light blue lines) can provide single-nucleotide resolution of haplotype differences (polymorphic differences, red lines in exon 2). This requires a complex panel of oligonucleotide probes to discern differences between HLA alleles. This probe set is static and therefore cannot adjust to novel alleles. SSPs (orange arrows) can provide haplotype-specific or allele-specific resolution of nucleotide differences and additionally provide some level of phasing between polymorphic sites. As with SSOs, these oligonucleotide sets are complex and static, limiting their flexibility. SBT provides whole-exon information on the polymorphic content of the HLA allele (amplification primers [dark green] and sequencing primers [light green arrows]) but cannot discern phasing, as this method generally does not rely on allele-specific primers for amplification as a first step. Next-generation sequencing (NGS) provides whole-gene amplification (amplification primers, purple arrows) and detection of polymorphic content for any HLA allele (known or unknown) and provides significant phasing between polymorphic sites that are within the read lengths of the system being used (usually between 200 and 1000 bp). This is accomplished through the alignment of thousands of short overlapping reads that are combined to form a single consensus sequence (blue lines).

are usually exon 2 and 3 of class I and exon 2 of class II genes. However, this rather limited genomic characterization generates many typing ambiguities.

SSO interrogates polymorphic differences using panels of individual DNA oligo probes that differentially hybridize to the target of interest. The probe either perfectly matches or mismatches the target's polymorphic sites. The hybridization pattern of the oligos is compared with an expected pattern, based on the sequence database of HLA alleles, and is interpreted as an HLA type.

SSP uses panels of specific primer sets that overlap polymorphic sites. Perfectly matched primers produce an amplification product, whereas mismatched primers do not. The pattern of amplification from multiple primer sets determines the HLA allele.

In **SBT**, specific gene regions, usually exons, are amplified and sequenced through a process of polymerase-based extension of specific-sequencing primers using fluorescently labeled nucleotides, indicating allelic differences base by base. Whether SSO, SSP, or SBT; the genomic regions analyzed are usually exon 2 and 3 of class I and exon 2 of class II genes. The result of this rather limited genomic characterization is the generation of many typing ambiguities.

Next-Generation Sequencing

Protocols utilizing NGS technology are on the rise because they provide the means for the complete characterization of these genes and the elimination of ambiguities in a cost-effective manner.[38] Regardless of the platform (Illumina MiSeq; ThermoFischer Ion Torrent; Pacific Biosciences, Oxford Nanopore), these systems resolve a number of technological barriers and limitations that continue to hamper existing molecular techniques, such as inflexibility in typing practices, discovery of novel alleles, and the inability to easily resolve phase ambiguities. Although

HLA typing by NGS has been introduced only recently, it is likely that this new method will transform the way HLA typing is performed in the coming years.

KEY CONCEPTS

The Resolution of the Human Leukocyte Antigen (HLA) Typing Problem

- Currently HLA typing is performed primarily through DNA-based methodologies.
- It appears that the dominant methodology will soon be the single molecule DNA sequencing (next-generation sequencing [NSG]) and most likely with platforms sequencing the length of the whole gene, whether class I or class II.

HLA NOMENCLATURE

The HLA genes/alleles are very polymorphic (about 15,000 have been named, and close to 30,000 sequence entries have not been named yet), and their numbers are expected to further increase, approaching hundreds of thousands, possibly millions. This has led to the development of comprehensive systems for their naming.

The WHO Nomenclature Committee for factors of the HLA system undertook the first systematic approach for the naming of HLA alleles in 1968. The HLA naming convention has undergone a substantial number of iterations because the earlier naming conventions were unable to address the growing numbers and complexity of alleles (*i.e.,* A*02 and B*15 have more than 100 alleles). The most recent nomenclature was introduced in 2010 to address the growing numbers of new alleles being discovered and to reduce naming complexity and confusion. The changes added colons (:) into the allele names to act as delimiters of the separate fields (field separator). Thus each HLA allele name has a unique number corresponding to up to four sets of digits separated by colons.

The first field following the asterisk in the allele name (XX:xx:xx:xx) describes the allele family and generally corresponds to the serological assignment carried by the allele. HLA typing defined only at the first field is often referred to as "low-resolution typing." The second field following the first colon (xx:XX:xx:xx) is assigned sequentially as new alleles are determined (*e.g.,* 01, 02, 03….101, etc.). Together, these two fields (XX:XX) indicate one or more nucleotide substitutions that change the HLA protein coding sequence and are often referred to as "high-resolution typing." Indeed, the Harmonization of Histocompatibility Typing Terms Working Group recently defined a high-resolution typing result "as a set of alleles that encode the same protein sequence for the region of the HLA molecule called the antigen-binding site and that exclude alleles that are not expressed as cell-surface proteins."[39] The third field (xx:xx:XX:xx) is for designating synonymous nucleotide substitutions within the coding sequence that do not change the amino acids of the protein, and the fourth field (xx:xx:xx:XX) identifies sequence polymorphisms in introns, or in the 5' and 3' untranslated regions (Fig. 5.6).

All alleles receive a name that includes at least the first two fields. At the end of the allele name, specific characters have been added (N = null; L = low expression; S = secreted; C = cytoplasm; A = aberrant; Q = questionable) to designate unique characteristics for an allele, such as whether a protein is not expressed (*i.e.,* HLA-A*24:09N) or whether the expression of the protein is unclear (*i.e.,* HLA-A*32:11Q).[40] For ambiguous allele strings, the codes "P" and "G" were introduced. A group of alleles having nucleotide sequences that encode the same protein sequence for the peptide binding domains (exon 2 and 3 for HLA class I and exon 2 only for HLA class II alleles) will be designated by an upper case "P," which follows the two-field allele designation of the lowest-numbered allele in the group. For example, HLA-A*01:01:01:01, HLA-A*01:01:01:03, or HLA-A*01:37 could be named HLA-A*01:01P.

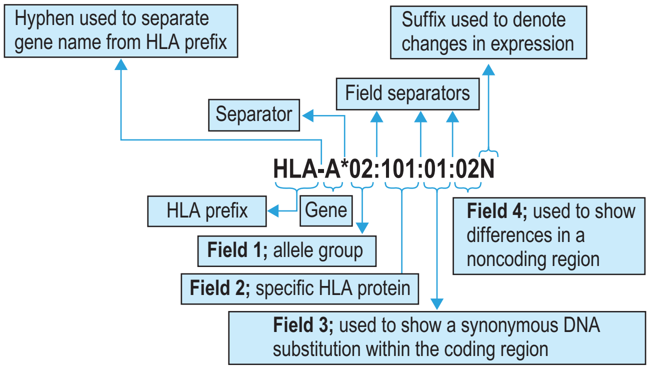

Hyphen used to separate gene name from HLA prefix
Suffix used to denote changes in expression
Field separators
Separator
HLA-A*02:101:01:02N
HLA prefix
Gene
Field 1; allele group
Field 2; specific HLA protein
Field 3; used to show a synonymous DNA substitution within the coding region
Field 4; used to show differences in a noncoding region

FIG 5.6 Human Leukocyte Antigen (HLA) Nomenclature. [Courtesy of Steven G. E. Marsh, Anthony Nolan Research Institute, London, United Kingdom.]

A group of alleles that have identical nucleotide sequences across the exons encoding the peptide binding domains (exons 2 and 3 for HLA class I and exon 2 for HLA class II) were named after the first allele in the sequence and given the code "G" as a suffix. The upper case "G" follows the first three fields of the allele designation. For example, HLA-A*01:01:01:01, HLA-A*01:01:01:03, or HLA-A*01:37 could be named HLA-A*01:01:01G. More details regarding HLA nomenclature can be found on the website: http://hla.alleles.org.

To manage and to have access to the sequences of the ever-growing number of alleles, the IMGT/HLA Database project was initiated in 1997 as part of a European collaboration. The database is an invaluable resource, as it provides detailed DNA sequences and protein sequences for all known HLA alleles. It is also interactive and incorporates tools for data retrieval and analysis so that the user can select what segments of the gene/molecule to examine and make comparisons among different alleles. It can also be used for new data submission.

ON THE HORIZON

- Advanced technologies for the sequencing and detailed characterization of human leukocyte antigens (HLAs) and of the whole major histocompatibility complex (MHC) should allow a clearer elucidation of the functional interrelationships of the different genes within the MHC and the genomic elements responsible for many MHC-associated diseases.
- Computational approaches for the accurate definition of peptide-binding properties of individual HLA alleles, whether class I or class II, will influence our ability to control the tri-molecular complex of HLA–peptide–TCR and therefore control some processes, such as responses to infectious diseases, autoimmunity, transplantation, vaccine design, and cancer.
- Immunotherapies for cancer involving neoantigens will require individualized approaches when HLA alleles play a critical role.
- Understanding the genomic organization of the MHC, the most complex genomic region in the human genome, will most likely reveal and teach us important lessons relevant to the organization and operation of the rest of the genome as well.

FUTURE LEARNING AND RESOURCES

This chapter provides only a limited sketch of this fascinating, but complex, topic. The reader is referred to the *HLA Facts Book* for a more detailed and very accessible presentation, although slightly out of date. There are also a number of websites with extremely useful information. Four stand out in terms of their utility and the curated quality of the information. *(i)* The IMGT/HLA Database contains all MHC sequences and has a variety of sequence alignments of different alleles as well as specialized sequence searches (http://www.ebi.ac.uk/imgt/hla/index.html and http://hla.alleles.org). *(ii)* The NCBI maintains dbMHC, which includes several components of the international histocompatibility working group (IHWG) that are of interest. Among these is the anthropology database, which contains HLA class I and class II allele and haplotype frequencies in various human populations (http://www.ncbi.nlm.nih.gov/projects/mhc/). *(iii)* Information about the genes and the genetic organization of the MHC is contained in several sites, but perhaps the most comprehensive and comprehensible is that using the Entrez search engine (http://www.ncbi.nlm.nih.gov). *(iv)* A comprehensive database of MHC ligands and peptide motifs is located at http://www.syfpeithi.de.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Shiina T, et al. The HLA genomic loci map: expression, interaction, diversity and disease. J Hum Genet 2009;54(1):15–39.
2. Aken BL, et al. The Ensembl gene annotation system. Database (Oxford) 2016;2016.
3. Clark PM, Kunkel M, Monos DS. The dichotomy between disease phenotype databases and the implications for understanding complex diseases involving the major histocompatibility complex. Int J Immunogenet 2015;42(6):413–22.
4. Dapprich J, et al. The next generation of target capture technologies - large DNA fragment enrichment and sequencing determines regional genomic variation of high complexity. BMC Genomics 2016;17(1):486.
5. Lund O, et al. Definition of supertypes for HLA molecules using clustering of specificity matrices. Immunogenetics 2004;55(12):797–810.
6. Kelley J, Walter L, Trowsdale J. Comparative genomics of natural killer cell receptor gene clusters. PLoS Genet 2005;1(2):129–39.
7. Parham P. Immunogenetics of killer cell immunoglobulin-like receptors. Mol Immunol 2005;42(4):459–62.
8. Beck S, Trowsdale J. The human major histocompatability complex: lessons from the DNA sequence. Annu Rev Genomics Hum Genet 2000;1:117–37.
9. Lanier LL. On guard–activating NK cell receptors. Nat Immunol 2001;2(1):23–7.
10. Pablo Gomez-Prieto CP-L, Diego Rey Enrique Moreno, Antonio Arnaiz-Villena HLA-G, -F and -E. Mehra NK, editor. Polymorphism, Function and Evolution in The HLA Complex in Biology and Medicine A Resource Book. Jaypee Brothers Medical Publishers Ltd; 2010. p. 159–74.
11. Karlsson L. DM and DO shape the repertoire of peptide-MHC-class-II complexes. Curr Opin Immunol 2005;17(1):65–70.
12. Housset D, Malissen B. What do TCR-pMHC crystal structures teach us about MHC restriction and alloreactivity? Trends Immunol 2003;24(8):429–37.
13. Stefanova I, et al. On the role of self-recognition in T cell responses to foreign antigen. Immunol Rev 2003;191:97–106.
14. Trowsdale J, Parham P. Mini-review: defense strategies and immunity-related genes. Eur J Immunol 2004;34(1):7–17.
15. Lilley BN, Ploegh HL. Viral modulation of antigen presentation: manipulation of cellular targets in the ER and beyond. Immunol Rev 2005;207:126–44.
16. Raulet DH. Missing self recognition and self tolerance of natural killer (NK) cells. Semin Immunol 2006;18(3):145–50.
17. Murphy K, et al. Janeway's immunobiology. 8 ed. New York: Garland Science; 2012.
18. Winchester R. The genetics of autoimmune-mediated rheumatic diseases: clinical and biologic implications. Rheum Dis Clin North Am 2004;30(1):213–27, viii.
19. Vinay DS, et al. Immune evasion in cancer: Mechanistic basis and therapeutic strategies. Semin Cancer Biol 2015;35(Suppl.):S185–98.
20. Rouas-Freiss N, et al. The dual role of HLA-G in cancer. J Immunol Res 2014;2014:359748.
21. Wurz GT, Kao CJ, DeGregorio MW. Novel cancer antigens for personalized immunotherapies: latest evidence and clinical potential. Ther Adv Med Oncol 2016;8(1):4–31.
22. Holoshitz J. The quest for better understanding of HLA-disease association: scenes from a road less travelled by. Discov Med 2013;16(87):93–101.
23. Farh KK, et al. Genetic and epigenetic fine mapping of causal autoimmune disease variants. Nature 2015;518(7539):337–43.
24. Ladewig E, et al. Discovery of hundreds of mirtrons in mouse and human small RNA data. Genome Res 2012;22(9):1634–45.
25. Mehra NK, et al. The HLA complex in biology and medicine : a resource book. New Delhi: Jaypee Bros. Medical Publishers; 2010.
26. Warrens A, Lechler R. HLA in health and disease. San Diego, Calif.; London: Academic; 2000.
27. Tiwari JL, Hla and disease associations. 2012, [Place of publication not identified]: Springer.
28. Schlosstein L, et al. High association of an HL-A antigen, W27, with ankylosing spondylitis. N Engl J Med 1973;288(14):704–6.
29. Mignot E. Genetics of narcolepsy and other sleep disorders. Am J Hum Genet 1997;60(6):1289–302.
30. Bradfield JP, et al. A genome-wide meta-analysis of six type 1 diabetes cohorts identifies multiple associated loci. PLoS Genet 2011;7(9):e1002293.
31. Winchester R. The molecular basis of susceptibility to rheumatoid arthritis. Adv Immunol 1994;56:389–466.
32. Goronzy JJ, Weyand CM. Rheumatoid arthritis. Immunol Rev 2005;204:55–73.
33. Michou L, et al. Validation of the reshaped shared epitope HLA-DRB1 classification in rheumatoid arthritis. Arthritis Res Ther 2006;8(3):R79.
34. Hussman JP, et al. GWAS analysis implicates NF-kappaB-mediated induction of inflammatory T cells in multiple sclerosis. Genes Immun 2016;17(5):305–12.
35. Greco L, et al. The first large population based twin study of coeliac disease. Gut 2002;50(5):624–8.
36. Cheng CY, et al. HLA associations and clinical implications in T-cell mediated drug hypersensitivity reactions: an updated review. J Immunol Res 2014;2014:565320.
37. Lind C, Ferriola D, Mackiewicz K, et al. Next-generation sequencing: the solution for high-resolution unambiguous human leukocyte antigen typing. Hum Immunol 2010;71:1033–42.
38. Duke JL, et al. Determining performance characteristics of an NGS-based HLA typing method for clinical applications. HLA 2016;87(3):141–52.
39. Nunes E, et al. Definitions of histocompatibility typing terms: Harmonization of Histocompatibility Typing Terms Working Group. Hum Immunol 2011;72(12):1214–16.
40. Detrick B, Schmitz JL, Hamilton RG, Manual of molecular and clinical laboratory immunology. [Chapter 113. Molecular Methods for Human Leukocyte Antigen Typing: Current Practices and Future Directions. ASM Press June 10, 2016; p 1069–1090] 2016.

MULTIPLE-CHOICE QUESTIONS

1. Genome-wide association studies (GWAS) reveal a large number of single nucleotide polymorphisms (SNPs) within the MHC to be associated with many diseases. This finding could have resulted from:
   - **A.** Extensive linkage disequilibrium within the major histocompatibility complex (MHC)
   - **B.** Numerous genomic elements within the MHC may contribute independently to different diseases
   - **C.** Many genes work together for certain physiological functionalities, and in particular diseases there may be multiple loci affected simultaneously
   - **D.** All of the above

2. Human leukocyte antigen (HLA) genes are very polymorphic because:
   - **A.** They need to interact with a wide range of environmental and intracellular molecular entities to which they elicit an appropriate immune response.
   - **B.** They need to interact with the T-cell receptor (TCR), which is also polymorphic.
   - **C.** It is necessary for thymic selection.
   - **D.** Through evolution they have mutated, but their functionality is not affected.
   - **E.** All of the above.

3. Single-molecule DNA sequencing or next-generation sequencing (NSG) is the right technology for the characterization of the HLA polymorphisms because:
   - **A.** It allows the cost-efficient and complete sequencing of the HLA genes
   - **B.** It provides sequencing of haploid DNA (single DNA molecule), and by doing so the polymorphisms are on phase
   - **C.** Eliminates essentially all ambiguities generated when other legacy methods are used
   - **D.** All of the above
   - **E.** None of the above

4. The reported associations of diseases with HLAs can result from:
   - **A.** Linkage disequilibrium with another nearby locus that is actually responsible for the disease
   - **B.** The HLA is involved in the disease process directly
   - **C.** The HLA is necessary but not sufficient for the disease
   - **D.** All of the above
   - **E.** None of the above

5. HLA molecules are involved in many immune-related responses and processes; which one(s) are true?
   - **A.** HLA class I genes encode products involved in both innate and adaptive immunity.
   - **B.** HLA class II are involved in adaptive immunity.
   - **C.** HLA class I bind peptides originating primarily from outside the cell.
   - **D.** HLA class II bind peptides anchored with their amino and carboxyl ends in the binding groove of the HLA molecule.
   - **E.** CD8 reacts with the domain proximal to the membrane of a HLA class II β chain.

6

Overview of T-Cell Recognition: Making Pathogens Visible to the Immune System

Andrea J. Sant

ANTIGENS

Antibodies and T-Cell Receptors Recognize Antigens

By the late nineteenth century, "antibodies" were the hypothesized to be molecular entities that mediated specific immune memory and could neutralize toxins and whose presence resulted in the formation of precipitates when mixed with the molecular species that induced their formation. In almost all cases, evidence for the presence of such antibodies required the prior exposure of responding animals to the very substances (or ones closely related, as in the case of toxoids) with which the antibodies reacted. This *specific* relationship of inducing agent and antibody led to the concept of an *antigen*—the molecular entity that could induce the formation of antibodies specific for it and that could be recovered in the blood of exposed animals. By developing the concept of the specific *receptor,* with a specificity analogous to the lock-and-key model of enzymes, Paul Ehrlich could explain the specificity of antibodies in molecular terms of a reciprocal interaction between a receptor and its binding partner (*ligand*).[1] Thus an "antigen" is any molecule that, in whole or in part, binds specifically to the antigen-binding domain of an "antigen receptor" (antibody or T-cell receptor [TCR]) (Fig. 6.1).

Ehrlich proposed several tantalizing, but unsatisfying, explanations for the other critical property of antigens—that they induce the formation of their own antibodies. This view of antigen is that of the vaccinologist, who wants to induce effective immunity to an organism expressing that antigen, or of a clinician, who is wondering why a patient does or does not respond to a particular allergen, self antigen, or tumor antigen. More than a century later, the explanation for the antigenicity of antigens remains an important and poorly understood issue—why humans fail to respond adequately to some pathogen or tumor antigens and how vaccines can be improved; why humans respond to their own self antigens (autoantigens) or antigens present in tissue grafts (alloantigens); and how autoimmune and tissue graft-related diseases (graft-versus-host disease [GvHD] and graft rejection) can be prevented or treated. The cellular mechanisms governing how and when humans respond to antigens remains at the cutting edges of both laboratory science and clinical medicine and are discussed later in this chapter. At the core of this central issue is the defining feature of the immune system—the distinction between "self" and "nonself." The host must remain tolerant to its own macromolecules and yet have the capacity to respond to its nonself. The molecular components of the immune system, including cells that display TCRs (T cells) and B-cell receptors (BCRs; antibodies expressed by B cells), undergo developmental events that promote self-tolerance—elimination of reactivity to self proteins, leaving cells with the potential to respond to a diverse set of antigens that may be expressed on pathogens or other "foreign"molecules.

Antigens can belong to many different chemical classes and can derive from viral or bacterial proteins, lipids, carbohydrates, or combinations of these, such as lipoproteins or glycoproteins. Antigens can also be small chemical compounds, termed *haptens* (Fig. 6.2), made synthetically in the laboratory, such as nitrophenyl (NP), or be a natural compound introduced into the host, such as urushiol, the toxin found in poison ivy, which becomes modified and antigenic when introduced into the host (Fig. 6.3). Haptens generally require linkage to a larger host protein or foreign protein to become immunogenic.

Innate Receptors Recognize Pathogen-Associated Molecular Patterns or Danger Signal Ligands

As a result of advances in the domain of innate immunity, it has become challenging, but all the more important, to distinguish between antigens and the many ligands for innate immune receptors (Chapter 3). Innate receptor ligands are often described as exhibiting patterns or motifs characteristic of a microbial class or physiological condition and the host proteins that recognize them as "pattern recognition receptors."[2] Many of these innate ligands facilitate the host immune system's recognition of a pathogen. Included in these innate ligands are Toll-like receptors (TLRs), which recognize such ligands as bacterial lipopolysaccharide (endotoxin), or viral single- or double-stranded RNA. Included among the innate activators are cytosolic DNA sensors and nucleotide-binding oligomerization domain (NOD)–like receptors (NLRs), which recognize bacterial wall peptidoglycans and intracellular metabolites induced by cellular damage.

The conceptual difference between antigens and innate receptor ligands lies in the diversity of ligands and the receptors that can engage them. The innate immune response recognizes predictable ligands through binding of a limited set of receptors that recognize these pathogen-derived molecules. The result of this recognition is often a rapid production of cytokines or induction of cell-surface proteins, including host human leukocyte antigen (HLA) proteins and costimulatory proteins that help activate the adaptive immune system. In contrast, the adaptive immune response collectively recognizes a wide variety of antigens through a tremendously diverse, but clonally distributed, set of BCRs and TCRs. The function of the specific antigen receptors in the adaptive immune responses and the cells that bear them is to promote activation and proliferation of the antigen-specific cells, leaving the host with memory B cells or T cells that are in higher abundance and often in a poised state after the first encounter.

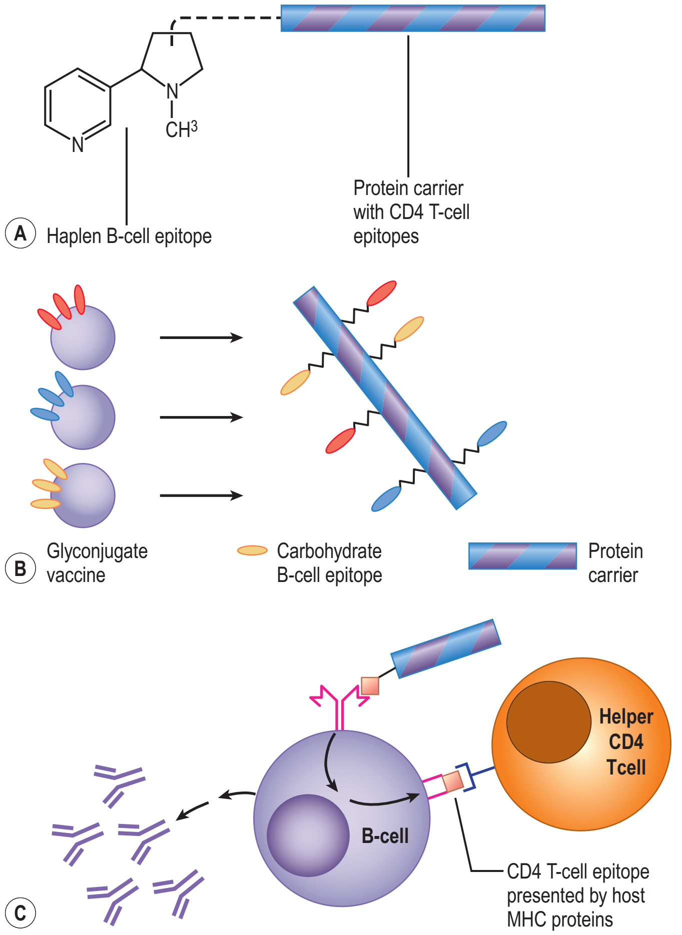

N
N
CH3
A
Haplen B-cell epitope
Protein carrier with CD4 T-cell epitopes
B
Glyconjugate vaccine
Carbohydrate B-cell epitope
Protein carrier
Helper CD4 Tcell
B-cell
CD4 T-cell epitope presented by host MHC proteins
C

FIG 6.1 Hapten, Carriers, and Two Kinds of Antigens. The antigen-binding site of an antibody binds an antigen through the latter's epitope: this is the biochemical sense of antigen used in ELISA, flow-cytometry and western blot analysis. Haptens are self-conjugating antigen moieties that can modify epitopes and provide new binding specificities. Haptens and many antigens by themselves are not immunogens, the second sense of "antigen". Immunogens (complete antigens) are processed by antigen presenting cells to reveal T-cell epitopes presented by MHC molecules. MHC, Major histocompatibility complex.

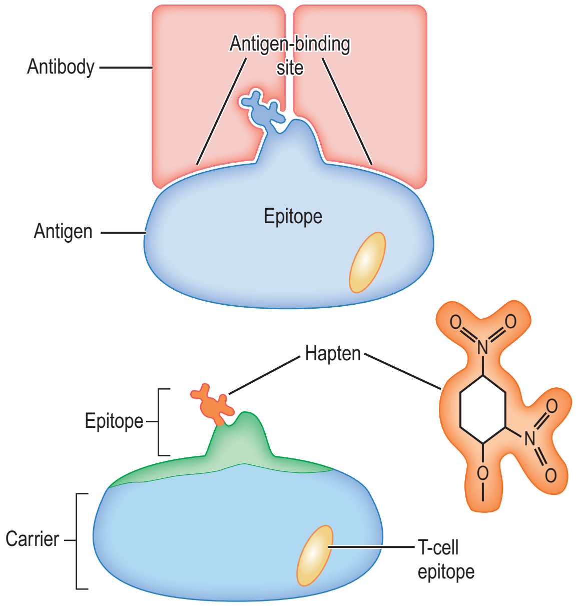


FIG 6.2 Haptens can modify the epitope of the antigen that is recognized by immunoglobulin, while the T-cell epitope remains unchanged.

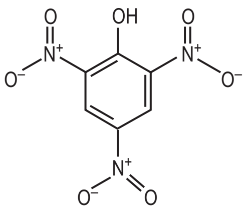
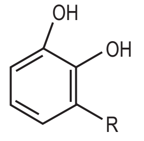
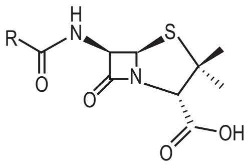
Trinitrophenol (Picric acid)	Urushiol (poison ivy) R=aliphatic chain	Penicillin
		

FIG 6.3 Examples of small chemical haptens that can be recognized by antigen specific immunoglobulins in the host, which are produced B cells.

This memory state allows for a more rapid response to a later confrontation with the same or related antigen or pathogen. This is the basis of induced immunological memory, which is the goal of many vaccines.

THE NATURE OF ANTIGEN RECOGNITION BY IMMUNOGLOBULIN AND T-CELL RECEPTOR DIFFERS

The receptors for B cells and T cells share similarities in development, structural organization, and function (Chapter 4). Both are derived from gene rearrangement events during development, and both combine variable and constant regions within their protein structure. Both are heterodimers that can be expressed at the cell surface of their respective lymphocytes and convey signals via partner chains within the transmembrane and intracytoplasmic segments that ultimately promote activation and proliferation of the responding lymphocyte. Despite these similarities, the nature of recognition by B cells and T cells is quite distinct. Immunoglobulins (Igs) tend to recognize solvent-exposed regions on a conformationally intact molecule (*e.g.,* a viral glycoprotein or a bacterial toxin). In contrast, TCRs typically recognize peptide fragments of protein antigens presented by host cell-surface molecules. In the case of TCRαβ, major histocompatibility complex (MHC) class I or II molecules are the presenters and lead to *MHC-restricted recognition* of antigen.

Antigens for B Cells

Vaccines can be used to target toxin binding and neutralization. For example, when ingested, *Vibrio cholerae* produces the secreted cholera toxin (CT) that can cause life-threatening diarrhea.[3,4] Both inactivated or attenuated forms of *V. cholerae* and recombinant forms of the cholera toxin B subunit protein have been used as vaccines. When introduced orally, these vaccines can elicit antibodies that neutralize the activity of the toxin within the gut. In contrast, protective antibodies elicited by viruses and virus-derived vaccines most often act by preventing binding and entry of the virus into host cells. Typically, vaccination strategies for these types of pathogens, such as human immunodeficiency virus (HIV) or influenza virus, employ semipurified or recombinant

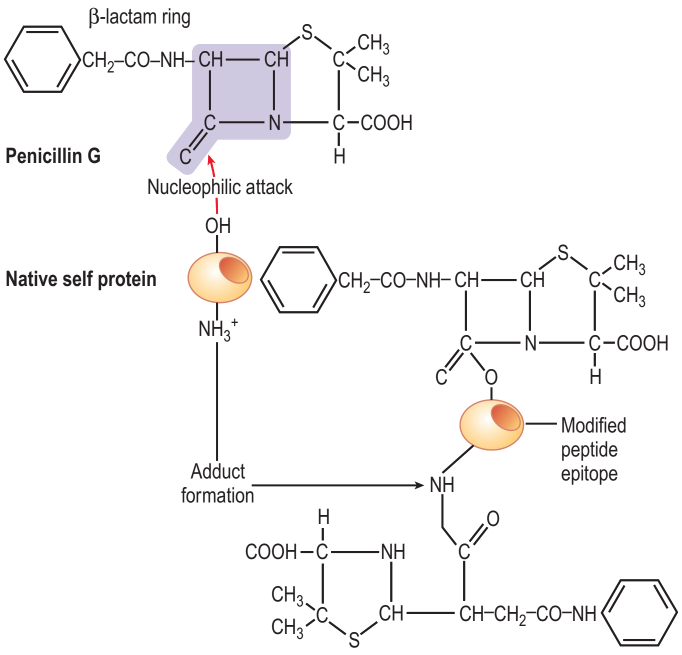


FIG 6.4 Penicillin Creates Neoantigens by Forming Covalent Adducts With Self Proteins. Penicillin allergies involve both antibodies against penicillin and with T-cell responses to penicillin-modified self proteins. The same chemical reaction that allows penicillin to inhibit peptidoglycan formation in bacteria leads to adduct formation of cellular proteins. Nucleophilic attack by penicillin G (upper left) on the β-lactam ring (shaded) opens the ring and creates an adduct (lower left) with serines and lysines. The lactam adducts can be presented to B cells as modified self proteins or processed for presentation by major histocompatibility complex (MHC) molecules to T cells as lactam-conjugated self peptides.

proteins that elicit neutralizing antibodies in the host against surface displayed proteins on the virus.[5-8] Antigens recognized by B cells can also be damaging to the host, as in the case of allergens such as penicillin and the chemical urushiol found in poison ivy (see Fig. 6.4).

KEY CONCEPTS

Antigens for B Cells

- Immunogens contain:
  - Epitopes that bind to the antigen-binding sites of antibodies
  - Class II epitopes for T helper cells
- Haptens can have almost any chemical nature.
- The amino acids that comprise epitopes on native proteins can be scattered throughout the primary sequence, but juxtaposed and found on the surface of the folded molecule.

Coupling of B-Cell and T-Cell Epitopes Permits Highly Focused Adaptive Responses

Encapsulated bacteria often express lipids or carbohydrate moieties on their cell surfaces that can be targeted by vaccines because they can promote pathogen clearance or neutralization. These polysaccharides by themselves are not immunogenic (*e.g.,* able to elicit an immune response) and therefore are often chemically conjugated to a complex protein.[9,10] Coupling promotes T-cell help and thus production of high-affinity antibodies (Fig. 6.1).

Coupling of a pathogen-derived carbohydrate to another foreign protein forms the basis of conjugate vaccines, used to elicit neutralizing antibody responses to bacteria (Chapter 90). Because related bacterial strains can display distinct polysaccharide structures that can be recognized by the B cells in the host (*serotypes*), some conjugate vaccines couple multiple distinct polysaccharides to a single carrier protein. A classic example is Prevnar® 13, which is composed of polysaccharides from 13 different serotypes of *Streptococcus pneumoniae* that are conjugated to a highly immunogenic inactivated diphtheria toxin protein. The conjugate vaccine offers distinct epitopes for the responding B-cell and CD4 T-cell helpers (see Fig. 6.1). The clonally distributed B-cell Ig receptor binds to the polysaccharides, whereas the protein carrier provides peptides to be presented on the surface of the priming dendritic cell (DC) and the antigen-reactive B cell. This allows focused help by the antigen-specific CD4 T cell.

For viral glycoproteins, B cells and CD4 T cells often recognize distinct sites on the same protein, with the B cell focusing on the surface regions of the intact protein structure and the T cell recognizing peptides produced after internalization and degradation.

Superantigens

Superantigens (SAgs) are microbial proteins that have the capacity to bind both class II MHC molecules and the TCR, thereby activating T cells via TCR signaling.[11,12] Included among the SAgs are bacterial toxins derived from *Staphylococcus aureus* and *Streptococcus pyogenes,* some gram-negative bacteria, *Mycoplasma arthritidis,* and *Yersinia pseudotuberculosis*, as well as proteins made by endogenous retroviruses. These SAgs are abbreviated as a three- or four-letter acronym that relates to the organism from which they are derived. These include *S. aureus* enterotoxin A or B ("SEA" or "SEB", respectively), *S. pyogenes* enterotoxin A or C ("SPE-A or SPE-C", respectively), and *Y. pseudotuberculosis* A (YPM-A").

KEY CONCEPTS

Properties of Superantigens

Defining Properties

- Presented and recognized as an unprocessed, native protein
- Contact T-cell receptor (TCR) and major histocompatibility complex (MHC) class II molecules outside the traditional antigen-binding groove

Specific Properties

- Selectively stimulate T cells expressing certain TCR Vβ chains
- TCR recognition not MHC allele-restricted
- Stimulate both CD4 and CD8 T cells in an MHC class II-dependent manner

SAgs bind class II MHC molecules (Chapter 5) on APCs as intact macromolecules. Although most bind outside of the peptide-antigen binding groove[13] some SAgs, such as the *S. aureus* toxic shock protein, can contact the bound peptides. Although class II binding by SAgs is largely independent of allelic polymorphism, it often displays a preference for particular class II isotypes. For example, HLA-DR tends to be preferred over HLA-DP and HLA-DQ in the binding of SAgs to human T cells. Key to the biological activity of SAgs is the ability to dock simultaneously with the HLA class II molecule expressed by APCs and the TCRs expressed by host T cells. TCR recognition of SAgs typically involves a characteristic and large subset of the various TCR Vβ families, since binding occurs outside of the TCR peptide-binding groove. Consequently, the binding of SAgs to class II and TCRs can recruit as many as 20% of the total TCR repertoire, leading to polyclonal activation of T cells. Subsequent production of proinflammatory cytokines by the activated T cells can lead to a profound systemic inflammatory syndrome. Most notable among these is toxic shock syndrome, caused by *S. aureus* (toxic shock syndrome toxin [TSST]) and streptococcal toxic shock syndrome (STSS), caused by *S. pyogenes.* Bacterial SAgs have also been implicated in the pathogenesis of acute rheumatic fever and Kawasaki disease.

ANTIGEN-PRESENTING CELLS

Cells That Present Antigens to B Cells: Follicular Dendritic Cells[14]

In absolute terms, B cells do not need another cell to present antigen to them. They express antigen Ig receptors (BCRs) that can interact with intact proteins expressed by pathogenic organisms or within protein vaccines. Physiologically, B cells typically recognize their antigen as a multivalent array displayed by other cell types.[15] The most potent and sustained B-cell responses involve BCR recognition of antigen in a cellular context in conjunction with provision of cognate CD4 T-cell help (*vide infra*). However, a polyvalent antigen, such as that displayed as highly repetitive structures expressed on encapsulated bacteria, can elicit B-cell antibody responses in the absence of specific T-cell help. These antigens are designated T-independent antigens, or "TI" antigens. Although they do not require specific CD4 T-helper (Th) cells, their responses can be enhanced by cytokines produced by other cell types.

Pathogen or vaccine-encoded proteins that enter the host and access secondary lymphoid tissue (*e.g.,* spleen or lymph nodes) are presented on the surface of specialized cell types via high-affinity binding to cell-surface molecules, such as complement receptors or Ig Fc receptors. Antigens accessing the lymphoid tissue can initially access and be bound via these receptors on subcapsular macrophages. Ultimately antigen accesses follicular dendritic cells (FDCs). FDCs bind opsonized antigen via complement receptors (CR1 and CR2) and thus can display the intact antigen directly to B cells. FDCs can maintain this antigen store throughout the extended course of the immune responses, as Ig affinity maturation takes place within the germinal center. It is thought that binding of opsonized antigen by FDCs leads to recycling of the antigen complex into nondegradative, internal, endosomal compartments. This allows intermittent display of the antigen at the surface as the B-cell Ig receptor repertoire matures into highly selected B cells. These highly selected B cells can differentiate into long-lived memory B cells and into plasma cells that secrete high-affinity antibodies.

Cells That Present Antigens to T Cells

T cells recognize antigen in the form of peptide fragments presented by host MHC proteins, which are termed *H-2* in the mouse and *HLA* in humans (Chapter 5). APCs mediate this process of T-cell recognition. Antigen accesses the draining lymph node or spleen, either directly, by drainage from a peripheral site or after being carried by tissue resident APCs. Peripherally distributed APCs (*e.g.,* DCs) act as sentinels in tissue sites of infection (*e.g.,* skin or lung). Here, DCs pick up and carry pathogen-derived antigens to the draining lymph node, where recognition by and priming of naïve T cells can take place. Resident DCs in the lymph node can also internalize antigen from the lymph drainage or from other antigen-bearing cells. In the secondary lymphoid tissue, there are limited sites in which the contact between circulating naïve T cells and APCs take place. These sites are dictated by cytokines, chemokines, and a network of stromal cells that collectively controls cellular trafficking and localization. In addition to the restrictions imposed by APCs and T cells locating one another, naïve T cells have a high threshold for activation. Active APCs require a high density of antigenic peptide–MHC ligands, and they must express accessory ligands that "costimulate" the T cells that engage their TCR.

The most important type of APCs for priming CD4 and CD8 T cells are DCs that gain access to the antigen and express costimulatory proteins that promote T-cell activation, such as B7 (CD80) and CD40. These costimulatory proteins are upregulated upon encounter with pathogen-derived pattern recognition receptors. After priming, other types of "professional" APCs can be recognized by T-cells, including B cells and macrophages, where delivery of help for antibody responses or pathogen elimination can occur, respectively. For CD4 T cells, T-cell interactions are typically limited to cell types that express MHC class II molecules, which are 0only expressed on a subset of cells. B cells, macrophages, and some epithelial and endothelial cells can express class II molecules, particularly after activation or in inflammatory cytokine milieus. In some experimental systems, more atypical APCs, such as mast cells and basophils, have been found to upregulate class II molecules and thus be targets of CD4 T cell recognition. CD8 T cells can recognize a wide variety of host cells because of the almost ubiquitous expression of MHC class I molecules. Almost all nucleated cells are thus potential targets for CD8 T-cell recognition, a function that is critical for elimination of infected cells at many distal sites by cytolytic CD8 T cells.

MHC-RESTRICTED RECOGNITION OF ANTIGEN

TCR, the antigen-specific receptor on T cells, recognizes a physical complex between host MHC proteins and small peptide fragments derived from protein antigens (Chapter 5). The interaction between the peptide and the MHC is highly specific. Because the genetically polymorphic regions of host MHC proteins are the decisive features in the MHC that determine which peptides are presented, this event is termed *MHC-restricted presentation of antigen* (Fig. 6.5).

Class I MHC

MHC class I proteins consist of a genetically polymorphic MHC-encoded heavy chain and a nonpolymorphic 12-kD light chain (β_2M). MHC class I molecules *present* peptide antigens to CD8 T cells, whose most common protective function is cytolysis of pathogen-infected cells or transformed tumor cells. Each molecule contains a series of pockets, controlled by genetically polymorphic amino acid residues, which allow high-affinity binding of the peptide's side chains. The binding cleft is flanked by hydrogen bonding residues at the periphery of the pocket so that the cleft is *closed* at its end. This peripheral closure of the binding pocket limits the size of peptides that can be bound and presented to CD8 T cells to 8–10 amino acids (Fig. 6.6, Top). Class I molecules are synthesized and expressed on all nucleated cells. Thus the

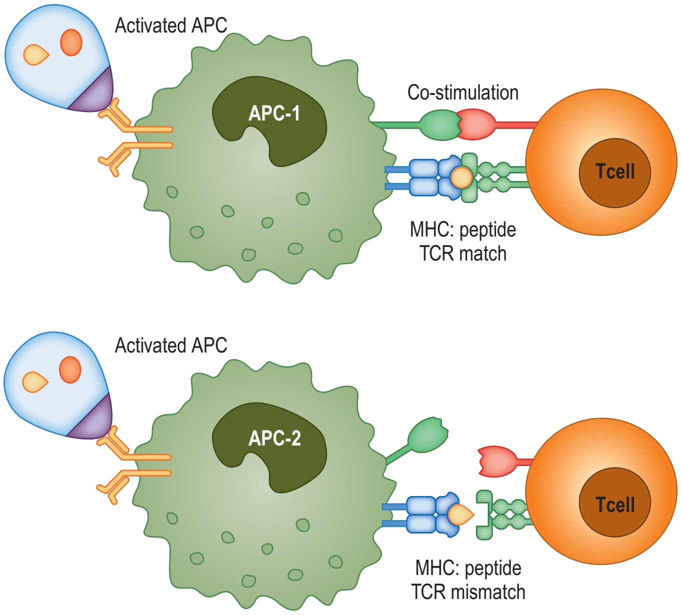


FIG 6.5 Major Histocompatibility Complex (MHC) Restriction Carries Out a Critical Function. Naive T cells (purple) respond to cognate epitopes only when presented in association with an MHC molecule on the surface of an antigen-presenting cell (APC). APCs can be activated by antigen uptake, cytokines, or pathogen-derived molecules, and this makes them more effective at activation T cells. Experimentally observed MHC restriction reflects the need for the APC to bear the correct MHC that matches the T cell (top). In the absence of a matched MHC, the T cell will be unable to respond to the peptide (bottom).

repertoire of presented peptides is widely available throughout the body for scanning by circulating CD8 T cells.

Class II MHC

MHC class II proteins consist of a heterodimer of MHC-encoded, genetically polymorphic α and β chains. Peptide binding is controlled by a series of 4–5 *pockets* into which the amino acid side chain of peptides selectively dock. Allele-dependent amino acid diversity within these pockets permits varied binding of peptide fragments across individuals. Coupled with a network of hydrogen-bonding residues between the main chain of the peptide and the α helices of the MHC class II α and β chains that stretches along the entire binding cleft, the interaction between peptide and MHC class II can be exceptionally stable. The class II binding cleft is open and thus can bind peptides ranging in length from 13 to 20 amino acids (see Fig. 6.6, Bottom).

Class II molecules are expressed on a restricted, but diverse, set of cells that includes B cells, macrophages, monocytes, DCs, and, in humans, activated T cells. Class II–bearing cells can activate CD4 T cells. CD4 Th cells promote expansion of B cells and their production of isotype-switched high-affinity antibodies. Interleukin-2 (IL-2)–secreting CD4 T cells facilitate CD8 T-cell expansion and generation of long-lived CD8 T memory. CD4 T cells can mediate protective immunity through secretion of cytokines, such as interferon-γ (IFN-γ), which leads to killing of intracellular pathogens. Finally, CD4 T cells can directly kill infected host cells or class II–positive tumors via perforin- or granzyme-mediated cytotoxicity.

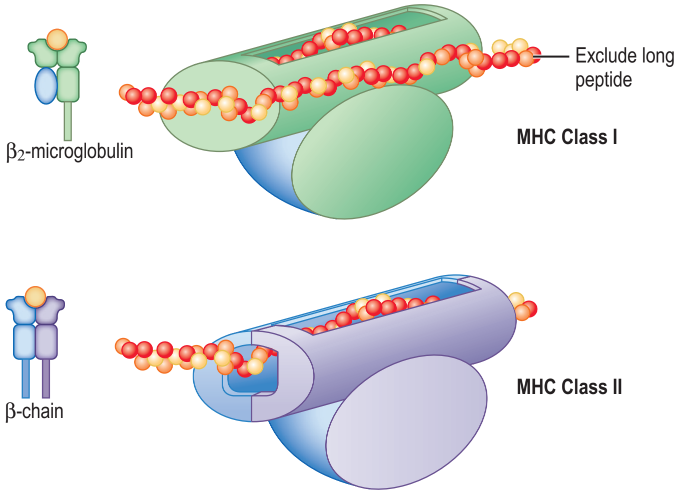


FIG 6.6 Peptide Binding by Major Histocompatibility Complex (MHC) Class I and II Molecules. MHC class I molecules (top) are typically "closed" at both ends. Peptides that are too long must be cleaved prior to entry into the binding site. The clefts of class II molecules (bottom) are "open" at the ends, and thus permit the binding of long peptides. Both molecules interact with peptides through both hydrogen bonding residues that are directed by nonpolymorphic residues in the MHC molecule and that are thus conserved among different complexes, and anchor-pocket interactions that typically employ polymorphic residues in the MHC molecule and thus vary for different MHC alleles and different peptides.

KEY CONCEPTS

Antigen Uptake for Presentation to CD4 T Cells

- B cells acquire antigen through uptake by the immunoglobulin receptor.
- C-type lectins expressed on the surface of many types of antigen-presenting cell (APC) receptors promote the binding and internalization of antigen by the APCs.
- Vaccines can be targeted to dendritic cells for CD4 T-cell priming by conjugation to C-type lectins and thus enhance antibody responses.

MHC Class II Antigen Presentation

MHC class II–restricted presentation of antigen to CD4 T cells is sensitive to inhibitors of endosomal proteolysis. Inhibitors include pharmacological reagents that neutralize endosomes, such as chloroquine or ammonium chloride, or specific protease inhibitors, such as leupeptin. These data have helped establish the paradigm that MHC class II molecules acquire peptides generated by endosomal proteolysis of internalized protein antigens. This process involves a number of cofactors and checkpoints that serve to regulate MHC class II molecule trafficking to endosomal/lysosomal compartments, acquisition of peptide, and trafficking to the cell surface. These pathways and mediators control the ability of class II molecules and their peptides to be recognized by CD4 T cells. Central features include the mechanisms that direct sorting and localization of class II molecules to the intracellular endosomal compartment(s) containing antigen, proteolytic enzymes that efficiently degrade antigen into peptides, and cofactors that promote rapid binding of the appropriate peptides[16-19] (Fig. 6.7).

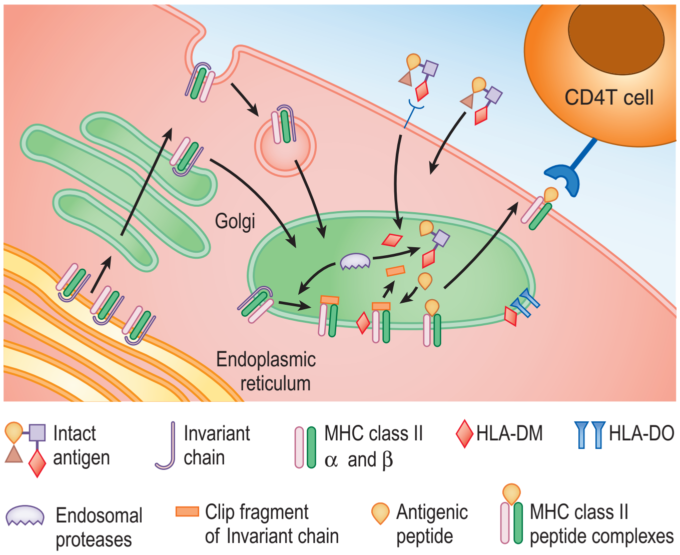


FIG 6.7 Major Histocompatibility Complex (MHC) Class II Presentation of Antigens to CD4 T Cells. Assembly of MHC class II molecules with invariant chain in the endoplasmic reticulum (ER) helps class II molecules fold correctly and blocks acquisition of peptide. After transport through the Golgi body, the class II invariant chain complex can be directly sorted into endosomal compartments or can gain access after a brief cell surface intermediate. Invariant chain is cleaved by endosomal proteases upon arrival into late endosomal compartments. Cleavage leaves the CLIP (class II–associated invariant chain peptide) fragment within the peptide binding pocket of the MHC class II dimer. Antigens can reach endosomal compartments through endocytosis, receptor-mediated uptake, or autophagy (not shown). Endosomal proteases cleave the protein antigens into small proteolytic fragments capable of binding to class II. The replacement of CLIP with an antigenic peptide is facilitated by human leukocyte antigen (HLA)-DM. HLA-DM also promotes binding of peptides with high affinity for the class II molecule. HLA-DO occupies the class II binding site on HLA-DM, preventing its binding to class II. Thus HLA-DO can adjust the active level of HLA-DM within endosomal compartments. After DM-editing of the highest-affinity peptides, a cohort of MHC class II molecules that has stable peptides bound is exported to the cell surface for presentation to CD4 T cells.

Invariant Chain

Invariant chain is a nonpolymorphic, non–MHC-encoded type II membrane glycoprotein protein that associates with class II during biosynthesis. It contributes several discrete functions to the biogenesis and function of MHC class II molecules.[20] Many of the functions of the invariant chain are conveyed by different segments or domains of this protein. Invariant chain also has several different isoforms resulting from alternative splicing and alternative start sites, with the most abundant being approximately 31 kD (mouse) or 33 kD (human).

Early during its biosynthesis in the endoplasmic reticulum (ER), invariant chain forms a trimer that nucleates assembly of three MHC class II dimers, forming a nonamer, which exits the ER. During assembly, a small segment of invariant chain termed *CLIP* (class II–associated invariant chain peptide) facilitates correct class II folding. CLIP occupies the peptide-binding pocket of the class II molecule. This prevents premature occupation of the peptide-binding pocket by self peptides, and it ensures the conformation integrity of the membrane distal, class II peptide binding domains. Successful assembly of invariant chain with MHC class II allows egress of the complex from the ER to later biosynthetic compartments. In this context, invariant chain acts as a molecular chaperone.

As they leave the ER, MHC class II and invariant chain are cotranslationally modified by the addition of complex carbohydrates into the Golgi complex. Either directly from the trans-Golgi network or after a short-lived cell-surface intermediate, they are then selectively sorted into the endosomal compartment of APCs. The amino terminal cytosolic tail of invariant chain plays a key role in the sorting event, which is mediated by the adaptor protein AP-2. Once in the endosomal compartments, invariant chain is degraded by defined sequential proteolytic events mediated by pH-dependent proteases, including aspartyl and cysteine proteases. The key proteases that cleaves invariant at CLIP's amino-terminus are somewhat cell-type specific. They include cathepsin S in peripheral APCs and cathepsin L in the thymus. Because the small CLIP fragment is sequestered within the class II binding pocket, it resists endosomal proteolysis. A rapid exchange reaction must take place between CLIP and the self or antigenic peptide. This process of peptide exchange is promoted by an additional critical protein cofactor termed *DM,* as discussed below.

APCs lacking invariant chain exhibit defects in antigen presentation and class II localization in endosomal compartments. The loss in efficiency of antigen presentation can vary, depending on the allele of class II and the peptide under study. For example, the murine I-A^b molecule expressed by the very commonly used C57BL/6 strain of mouse displays very inefficient egress from the ER, thus blunting all subsequent functions of export, intracellular sorting, and peptide acquisition. Other alleles of class II are able to assemble and exit the ER, although inefficiently. By alternative sorting mechanisms, they can access endosomal compartments. This allows them to access some, although not all, degraded antigens. Invariant chain dependent and independent peptide epitopes are distinct. Epitopes that do not need invariant chain for presentation by class II molecules are thought to represent peptides made available early in the endosomal compartment, perhaps accessing mature invariant chain–free class II from the cell surface via internalization.

KEY CONCEPTS

MHC Class II–Restricted Antigen Presentation

- Major histocompatibility complex (MHC) class II–peptide complexes form in late endosomes.
- Acidic pH in endosomes promotes:
  - Proteolysis of protein antigens
  - Alteration class II conformation to make it more receptive to peptide binding
  - Release of CLIP (class II–associated invariant chain peptide) from class II
  - Class II interactions with human leukocyte antigen (HLA)-DM.
- HLA-DM interaction with MHC class II molecules promotes:
  - Release of the CLIP fragment
  - Binding of antigenic peptide
- DM editing selects for high-affinity peptides to bind to class II and recruit CD4 T cells.

Class II Peptide Loading

The precise intracellular compartment where class II peptide loading occurs has been a subject of much investigation.[21] Endosomal vesicles "mature" over time, becoming more acidic as they progress toward an end-stage lysosome. MHC class II molecules change conformation with acidic pH, becoming more "peptide receptive." Also, as pH decreases, endosomal proteolytic enzymes generally become more active. Some resident endosomal proteases are synthesized as inactive enzymes, termed *zymogens,* which become activated either directly by pH-dependent conformational changes or by cleavage of other endosomal pH dependent proteases. In most APCs, late endosomal compartments are enriched in serine proteases (cathepsins A and G), aspartic proteases (cathepsins D and E), and cysteine cathepsins (cathepsins S and L). The low pH in late endosomes promotes antigen unfolding and access to reductases, such as the IFN-γ–inducible thiolreductase (GILT) that cleaves disulfide bonds. The combination of low pH and reduction of disulfide bonds thus facilitates protein unfolding, proteolytic digestion, and generation of antigenic peptide fragments.

Although MHC class II molecules traffic through many endosomal compartments, they are most highly enriched in late endosomes that have a pH of approximately 4–5. These heterogeneous intracellular compartments are termed *MIICs* (MHC class II compartments) or *MVBs* (multivesicular bodies) and can have either multilamellar or multivesicular organization, consisting of both a limiting membrane and internal membranes. In these compartments, class II molecules localize with the other critical components of the antigen presentation pathway. The precise organization of the class II–containing compartment likely varies with the APC studied (B cells, macrophages, or DCs) and can change upon signaling or antigen receptor engagement.

The exchange reaction of the invariant chain fragment for antigenic peptides represents an intriguing biochemical event that derives from many factors of MHC class II structure and biology. The first is the polymorphic nature of the peptide-binding pocket of the class II heterodimer. Much of the genetic variability in both α and β chains lies within the peptide-binding cleft of class II molecules. This highly localized genetic polymorphism is thought to allow alternate MHC molecules expressed in different individuals to capture distinct subsets of peptides from pathogenic organisms.

All allelic forms of class II thus must be able to accommodate the CLIP segment while they are associated with an intact invariant chain. This ensures conformational integrity of the class II molecule early in biosynthesis. Once invariant chain is proteolytically cleaved, the various allelic forms of class II are differentially able to sustain interactions with CLIP and acquire peptides. Class II molecules with high affinity for CLIP require an efficient mechanism to release CLIP. Class II molecules devoid of peptide have a tendency to aggregate and become degraded, particularly at the low pH of the endosome. Peptide binding to class II molecules, even at low pH where peptide acquisition is facilitated, is quite slow. These combined characteristics of class II structure and biochemistry led to the early speculation that there must be a mechanism to ensure the release of CLIP from all alleles of class II and rapid subsequent binding of antigenic peptide. The protein that facilitates this process is termed *HLA-DM* (in humans) or *H-2M* (in mice).

HLA-DM and Peptide Exchange

HLA-DM (DM) is the critical mediator in the release of CLIP from class II and its replacement with peptide.[22-26] HLA-DM α and β heterodimers are encoded within the HLA gene complex. Instead of associating with invariant chain, HLA-DM gains access to endosomal compartments via a tyrosine-based motif in its cytosolic tail that allows rapid internalization into endosomal compartments after a brief cell-surface intermediate. The release of CLIP from MHC class II molecules (the dissociation reaction) and the acquisition of antigenic peptide (the binding reaction) are greatly enhanced by the presence of DM. Both dissociation and binding are greatly affected by acidic pH. At acidic pH, interactions between DM and class II are initiated along their lateral faces on the side of class II that bears the amino terminal segment of peptide.

DM prefers class II molecules that are open and have unstable interactions with peptide. DM is thought to bind and stabilize this open intermediate of class II–peptide complexes that is promoted by low pH. This accelerates peptide release. This same open conformation is apparently also readily able to capture peptide, promoting a rapid exchange of CLIP for antigenic peptides. When a stable interaction between class II and antigenic peptide forms and the antigenic peptide is fully docked within the MHC class II molecule, DM is released from class II molecules. MHC class II–peptide complexes can then escape to the cell surface for recruitment of CD4 T cells. Accordingly, DM acts as a catalyst of peptide exchange for class II, binding to a transition state until the final product (MHC-class II–peptide) is formed.

Selection of Immunodominant Peptides

Of the many potential peptides that can bind to class II molecules and recruit CD4 T cells when introduced into the host as single peptides, only a subset of these epitopes recruit CD4 T cells when the host encounters an intact antigen or pathogen. For example, in a 50-kDa protein that has more than 80 potential 15–18mer peptides, the CD4 T cells may focus on just 3–5 peptide epitopes. These peptides are termed *immunodominant.* Another subset of peptides is capable of recruiting CD4 T cells in the host when introduced as single peptides. These are termed *cryptic* because they are sequestered from the immune responses to complex protein antigens. Other peptides can neither bind to the host class II nor recruit CD4 T cells.

Early models to explain the selectivity of CD4 T-cell responses suggested that intracellular antigen proteolysis would play a prominent role in selection of epitopes because intact antigens, but not peptides, require internalization and degradation. By this model, the position of the peptide within the three-dimensional structure of antigenic proteins might limit access to proteolytic enzymes and thus availability to bind to class II molecules. Conversely excessive degradation, leading to destruction of the peptide epitopes by endosomal proteases, could lead to diminished yield of some peptides for class II binding and presentation. These differential proteolytic events could thus restrict the potential presented repertoire to a limited number of peptides of the appropriate size (typically 12–25 amino acids). Although it is likely that in some cases such differential antigen processing might impact the yield of available peptide, accumulated data now suggest that selective presentation of potential antigenic peptides in association with MHC class II molecules is primarily as a result of intracellular DM editing.

Studies in both intact APCs and *in vitro* with purified recombinant class II and HLA-DM proteins suggest that DM selectively releases some peptides from class II molecules. Generally, these peptides possess low affinity for the class II molecule. Only those peptides that are stably bound to class II molecules will sustain interactions with MHC class II and survive endosomal DM editing. These MHC class II complexes will be successfully exported to the cell surface, displayed at high density on the APC and recruit CD4 T cells during an immune response. These *DM-selected* peptides are thus the targets of the focused CD4 T-cell response.

The relationship between *DM sensitivity* and the affinity of the peptide–class II complex is not absolute. However, it is generally a good predictive feature of immunodominant peptides. In some cases, the size and diversity of the CD4 T-cell repertoire can also determine the abundance of responding CD4 after encounter with pathogens or vaccines, but the qualitative ability of a peptide to successfully or ineffectively recruit CD4 T cells can be generally attributed to how stably the peptide binds to host HLA-class II molecules and thus DM editing. Importantly, modification of this parameter by increasing the affinity of peptide binding to class II can enhance the immunogenicity of peptides, and thus can be used to enhance immunogenicity of vaccine candidates.

HLA-DO

A final critical component of the MHC class II presentation pathway is HLA-DO (DO) in humans (H-2O in mice). The expression can be regulated independently of invariant chain and HLA-DM. DO is unevenly distributed in MHC class II–positive cells. Factors influencing expression include cell lineage and state of activation. The differential expression of DO has important biological consequences because it inhibits the activity of DM. The expression of DO in APCs can thus diminish CLIP release from class II, peptide loading of MHC class II and DM editing of the final repertoire of class II presented peptides. HLA-DM binds to DO in the same regions that it does the classic antigen-presenting class II molecule, but with a higher affinity and independently of peptide binding. HLA-DO can thus be considered to be a high-affinity *substrate mimic* for HLA-DM. DO initiates its interaction with DM within the ER, and the two protein dimers are transported together into endosomal/lysosomal compartments. *In vivo,* DO essentially titrates the amount of DM that is available to catalyze peptide exchange and editing of the peptide repertoire. The ratio of DM to DO synthesis by an APC will thus determine the amount of DM that is available for CLIP release, peptide binding, and peptide editing.[22,23]

HLA-DO is expressed in resting B cells but downregulated in germinal center B cells. Some subsets of DCs, including Langerhans cells, express DO as do some activated DCs. When B cells have the opportunity to capture antigen and process and present the derived peptides, they appear to compete for the limited number of CD4 Th cells that are needed for isotype switching and Ig affinity maturation in antigen-specific B cells. Because the expression of DO within APCs attenuates the efficiency of antigen presentation for most epitopes, it allows only the highest affinity B cells to successfully recruit CD4 T-cell help. Once this affinity threshold has been met, DO is downregulated. Robust antigen presentation then sustains the CD4 T cell–dependent germinal center response.

Targeting of Antigen into the MHC Class II Processing Pathway

MHC class II molecules use several mechanisms to access self, tumor, or pathogen-derived proteins. In the case of extracellular antigens, proteins can access endosomal compartments through receptor-mediated uptake, fluid phase uptake, or via internalization of cell surface molecules. The major pathways utilized will vary with the type and the activation state of APCs. For DCs and macrophages, calcium-dependent C-type lectins play a major role in accessing exogenous protein antigens or pathogens.[27-29] These lectins recognize carbohydrate recognition domains (CRDs) specific for different sugar side structures on self or foreign proteins. Examples include O-linked or N-linked sugars on glycoproteins or O-linked sugars on glycolipid molecules. Among the C-type lectins are those that recognize mannose-bearing ligands (mannose receptor, SIGNR1, and langerin) and those that recognize complex sugars (dectin 1).

Receptors that have been implicated in antigen presentation include DEC205, expressed on lymphoid-resident DCs and DIR2 (also known as *Clec4*) and its human orthologue DCIR2. Clec9A is a C-type lectin receptor that preferentially induces CD4 T-follicular helper (Tfh) cells, which promote B-cell production of isotype-switched high-affinity antibodies. Combined with ligands for pattern recognition receptors or immunostimulatory agents, such as CD40-specific antibodies, antigens delivered through these cell surface proteins can activate CD4 T cells and induce protective immunity. Use of antibodies to these lectins, which allow precise targeting of specific subsets of APCs with antigen conjugates, has been exploited in vaccine design strategies to elicit the desired immune response to weakly immunogenic antigens, such as tumor antigens and HIV.

In B cells, MHC-restricted antigen presentation is largely limited to those antigens taken up through the specific binding of the somatically variable cell-surface Ig[30-32] (Fig. 6.8). Ig-mediated uptake of antigen is up to 10,000 times more efficient than fluid phase uptake. Antigen capture by B cells within lymph node and spleen follicles can involve direct access of soluble antigens (<70 kDa) delivered via afferent lymphatics to FDCs. For larger membrane-associated antigens, viral aggregates, or immune

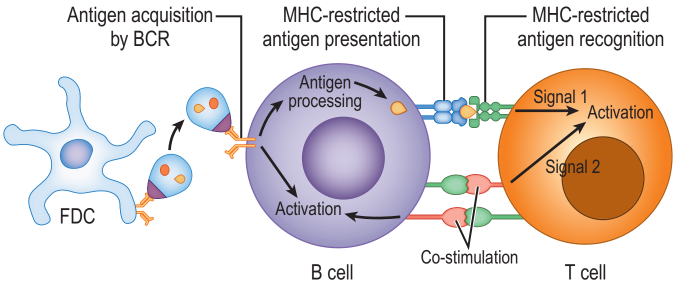

Antigen acquisition by BCR
MHC-restricted antigen presentation
MHC-restricted antigen recognition
Antigen processing
Signal 1
Activation
Signal 2
FDC
Activation
Co-stimulation
B cell
T cell

FIG 6.8 Antigen Presentation by B Cells. Under physiological conditions, B cells acquire antigens through their B-cell receptor (BCR). They present processed peptide epitopes via major histocompatibility complex (MHC) class II molecules to CD4 T cells. CD4 T cell cytokine production and cell-surface interactions between B cells and T cells (*e.g.,* CD40 and CD40L) are nonspecific. In contrast, the interaction between peptide-specific CD4 T cells and antigen-specific B cells facilitates antigen-specific clonal B-cell expansion and immunoglobulin (Ig) production.

complexes, subcapsulary macrophages can bind and capture antigenic material and transfer it across the subcapsulary sinus. For antigen presentation, B cells must *extract* the antigen from the presenting APC surface. This extraction of cell-associated antigen may involve localized protease secretion and/or myosin mediated pulling forces that allow invagination of antigen-containing membranes. Ig-mediated uptake of antigen enhances the ability of B cells to acquire low concentrations of antigen, providing a way to discriminate B cells with high affinity versus low affinity for antigen, and it also directly signals to the B cells. Signal transduction through the BCR is typically initiated by cross-linking of Ig from multivalent antigens (Chapter 4). Rearrangement of endosomal compartments following signal transduction promotes antigen processing and efficient recognition of the antigen-derived peptides by CD4 T cells.

After uptake, the antigen–Ig complex is delivered to late endosomal compartments where the more mild proteolytic compartments prevent terminal degradation of the antigen. MHC class II synthesis and transport to these compartments allows for peptide–class II epitope display. The capture of antigen by the Ig receptor is an essential feature of the ability of the B cell to obtain CD4 T-cell help. Note that to recruit CD4 help, the antigen recognized by the B cell must be physically attached, by either covalent or strong noncovalent interactions, to the antigen that will elicit CD4 T-cell help. The cognate-nature B cell–CD4 T cell help, which requires linkage of CD4 T-cell and B-cell epitopes, is key to vaccine strategies. It underlies the creation of conjugate vaccines used to elicit antibodies that will recognize carbohydrate ligands expressed by bacteria. Carbohydrate moieties coupled to protein carriers contain the recognition structure for the B-cell response (the carbohydrate) and a source of CD4 T-cell help (the protein carrier). The requisite linkage of CD4 T-cell and B-cell epitopes can limit antibody responses to complex viruses, whose protein components can become dissociated from each other during virus replication and ensuing cell death.

Autophagy, or *self eating,* provides a major pathway for viral or self proteins to access MHC class II molecules in APCs.[17,33] It is a catabolic process by which cytosolic, organelle-associated and nuclear materials are delivered to endosomal and lysosomal compartments. Autophagy provides key metabolites that allow cell survival under conditions of stress. It also allows access of self- and pathogen-derived intracellular cellular proteins into the class II–containing compartments. Indeed, 20–30% of peptides presented by class II molecules are derived from cytosolic or nuclear proteins. Autophagy has also been shown to be active in thymic APCs, thus broadening the array of self peptides that can mediate deletion of potentially autoreactive cells during development.

Three routes by which autophagy in APCs can lead to import of cellular proteins into the endosomal pathway of presentation include microautophagy, macroautophagy, and chaperone-mediated autophagy. All of these ultimately lead to degradation of host proteins in the lysosome. Macroautophagy, which is the best understood, involves formation of large, double-membrane vesicles containing cytosolic components or cytosolic organelles into *autophagosomes.* The autophagosomes fuse with endosomes and lysosomes, delivering their content for degradation and cellular reuse. These cytoplasmic proteins can be used as a source of peptides presented by class II molecules. Immunofluorescence studies have documented delivery of autophagosomes to the lysosomal compartments that contains both conventional class II and DM proteins.

Viral infection can activate and increase autophagy, thus increasing the potential the ability of MHC class II molecules to sample cytosolic viral proteins for presentation of their derived peptides. Viral antigens, such as Epstein-Barr virus (EBV)–nuclear antigen 1 and the *Mycobacterium tuberculosis* antigen Ag85B, have been shown to be presented via this pathway, allowing immune detection. Interestingly, a number of viruses have been shown to antagonize specific components involved in autophagy and thus evade immune surveillance.

KEY CONCEPTS

Antigen Presentation by Major Histocompatibility Complex (MHC) Class I Molecules

- Proteolysis for MHC class I–restricted presentation is typically the function of the proteasome, which is found in the cytosol.
- Peptides are imported through the endoplasmic reticulum (ER) membrane by the TAP (for transporter associated with antigen processing) transmembrane channel.
- Peptide binding to class I within the ER is orchestrated by a large multiprotein complex termed the *peptide-loading complex.*
- Tapasin serves as an adaptor between TAP and MHC class I molecules, and it edits the peptide repertoire presented by MHC class I for recruitment of CD8 T cells.

MHC Class I–Restricted Antigen Presentation

The classical pathway of MHC class I–restricted antigen presentation involves antigens derived from endogenously synthesized internal materials. By displaying the peptides made inside the cell, MHC class I proteins allow circulating CD8 T cells to survey host cells for expression of aberrant, mutant self proteins that may be selectively expressed in cancer or for viral protein–derived peptides in infected cells. Infected or cancerous host cells can then be eliminated by the CD8 T cells.

The Proteasome

Proteins made within APC gain access to the host MHC class I following cytosolic proteolysis by the *proteasome.*[16,34,35] The proteasome participates in protein catabolism within all cells. This multimeric protein complex has a 20S catalytic core composed of four stacked heptameric rings, which, together, determine access and the protease specificity of degradation. Proteases include a chymotrypsin specificity that cleaves after hydrophobic residues, a trypsin-like activity that cleaves after basic residues and a caspase-like activity that cleaves after acidic amino acid residues. A 19S regulator forms a lid on the proteasome and promotes recognition and binding of polyubiquitinated proteins, as well as unfolding and translocation of protein substrates.

The composition of the proteasome and its biochemical activity can vary, depending on cell type and conditions. The ubiquitously expressed proteasome is termed the *constitutive* form. In response to IFN-γ in some cell types, there is an exchange of the constitutive β subunits for homologous βi (for induced) subunits to create the *immunoproteasome,* which is expressed by many cells. The specificity of the β_1 proteasomal subunit can lead to diminished cleavage of peptides with acidic residues and enrichment for peptides with hydrophobic P9 pocket residues. These changes fit the preferences of most class I molecules and suggest coevolution of MHC class I and the proteasome. The proteasome and immunoproteasome appear to primarily differ in their preferences for proteolytic cleavage rather than absolute specificity. Recent studies, however, have suggested that some tumor cell recognition epitopes are selectively destroyed by the immunoproteasome. A third form of the proteasome, the *thymoproteasome,* has alternative forms of the β_5 subunit, β_{5t}. The *thymoproteasome* is selectively expressed in the thymus and permits the use of alternative peptides for *positive selection* of CD8 T cells during development (Chapter 8). This leads to a broader CD8 TCR repertoire with diminished reactivity to self peptides expressed in the periphery.

Import of Antigenic Peptides Into the Endoplasmic Reticulum and Final Trimming

After antigenic or self peptides are generated by the proteasome in the cytosol, they are imported into the ER for potential binding to newly synthesized MHC class I molecules. This transport activity, which must cross the ER membrane, is mediated by a protein heterodimer termed *TAP* (for transporter associated with antigen processing). The TAP heterodimer consists of TAP1 and TAP2, each with 6 transmembrane domains, which within the ER form a transmembrane channel. The adenosine triphosphate (ATP) binding domains are on the cytosolic side of the endoplasmic membrane, where peptide binding is initiated. ATP binding and hydrolysis provides the energy needed for conformational changes that drive channel function and the import of the peptides into the lumen of the ER. TAP is selective for peptide length and sequence in its ability to bind and import peptides. The carboxy-terminal amino acid residues are enriched for generally favored residues for MHC class I binding. These are generally hydrophobic for mice and, consistent with the less restricted peptide binding preferences of human class I, both acidic and hydrophobic for humans.

Import of peptides by TAP appears to enrich for peptides of suitable length (8–10mers) for MHC class I binding. Cytosolic peptides that gain access to the ER via TAP can also be trimmed, if needed. Within the ER, an amino peptidase, termed *ERAAP* (ER associated amino peptidase), can trim the peptide length from its amino terminal residues, permitting the peptide to bind firmly within the confines of the peptide-binding pocket.[36] Peptides lacking anchor residues that allow stable binding to class I are terminally degraded by EERAP.

The Peptide Loading Complex

Upon arrival to the lumen of the ER, peptides gain the opportunity to bind newly synthesized class I. As with class II, spontaneous acquisition of peptides by MHC class I molecules is inefficient and requires cofactors that both enhance the local concentration of the peptide and promote class I *peptide receptivity.* The cellular proteins that promote these events for class I are collectively referred to as the *PLC* (for peptide loading complex), which is a highly organized structure within the ER (Fig. 6.9).

Tapasin, an adapter protein, plays a key role in peptide loading. Tapasin helps MHC class I molecules associate with TAP and brings newly imported peptides into close proximity. Tapasin recruits the ERp57, a thiol oxidoreductase, which assists in the folding of class I by mediating disulfide bond formation. It also recruits the chaperone protein calreticulin to the PLC. Finally, tapasin interactions with MHC class I molecules appears to promote peptide acquisition via its maintenance of class I in a peptide-receptive state, serving the same role for class 1 as HLA-DM for class II. Also similar to DM is the ability of tapasin interactions with MHC class I molecules to *edit* or select the repertoire of bound peptides. The PLC thus promotes correct folding and disulfide bond formation for MHC class I molecules, and all of the intermolecular interactions needed for their assembly with peptide.

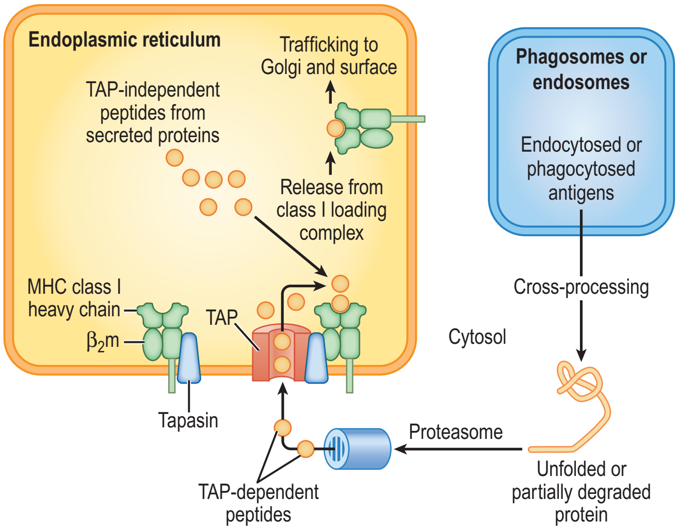

Endoplasmic reticulum
Trafficking to Golgi and surface
TAP-independent peptides from secreted proteins
Release from class I loading complex
MHC class I heavy chain
β_2m
TAP
Tapasin
Phagosomes or endosomes
Endocytosed or phagocytosed antigens
Cross-processing
Cytosol
Proteasome
Unfolded or partially degraded protein
TAP-dependent peptides

FIG 6.9 Major Histocompatibility Complex (MHC) Class I–Restricted Presentation. Internally synthesized proteins, destined for presentation by MHC class I molecules, are degraded by the large proteolytic complex in the cytosol termed the *proteasome.* Peptides that are generally of the appropriate size and sequence for binding to class I molecules are imported from the cytosol by the TAP (for transporter associated with antigen processing) transmembrane channel, catalyzed by adenosine triphosphate (ATP). Once in the endoplasmic reticulum (ER), the peptides can be trimmed to the correct size by an amino peptidase (ERAP). Class I molecule folding is facilitated by chaperones such as calreticulin and ER localized reductases (ERAP) that help class I molecule adopt a transport competent form. Tapasin bridges TAP and class I molecules and helps edit the peptide repertoire bound so that higher-affinity peptides are selected. After peptide binding, the class I–peptide complex is exported through the Golgi body and to the cell surface for recognition by CD8 T cells. There is a second pathway of class I presentation that can be used for externally derived antigens, such as pathogens and tumor cells. This pathway, termed *cross-presentation,* is not shown here but is discussed in the text.

Once assembled with a stable peptide and adopting a stable conformation, the MHC class I–peptide complex is competent to be transported from the ER, through the Golgi complex to the plasma membrane. Overall, the proteasome, TAP, the PLC, and ERAAP cooperate to efficiently load internally synthesized antigens of the correct size and optimal binding affinity onto host class I molecules and to display them at the cell surface for recognition by circulating CD8 T cells.

Cross-Presentation of Antigens for Recognition by CD8 T Cells

T cells are *primed* in lymphoid tissues by specialized APCs that belong to the DC lineage. Under physiological conditions, this process of T-cell priming is restricted to DCs in secondary lymphoid tissues because of their specialized location, access to antigen, and functionality.

Antigen-bearing DCs have several properties that are not shared by most host cells within the lymph node, properties that are essential for activation of antigen-specific T cells. First, subsets of DCs are positioned at sites of host and potential antigen encounter, for example, skin, mucosae, and lungs. Here, they are able to receive molecular signals from pathogens that both activate the DCs and promote internalization and processing of antigens. These pathogen-derived signals initiate DC mobilization, exit from the tissue, and transit to secondary lymphoid organs. Other subsets of DCs are *resident* cells within the lymph node and encounter antigens delivered either via the lymphatics or by other cell types that transit from the site of antigen encounter to the lymphoid tissue.

DCs gain access to secondary lymphoid tissue through expression of chemokine receptors and become localized to regions that allow encounter with recirculating T cells that enter the lymphoid tissue to survey potential antigen-bearing cells. DCs also express key costimulatory molecules, such as CD80 and CD86, and cytokines, such as IL-12, which are essential to prime naïve T cells for expansion and differentiation into the effector cells (Chapter 12).

The intracellular events associated with antigen acquisition and presentation by DCs is easy to envision for CD4 T cells because the antigens that need to access class II molecules can become internalized into the key endosomal compartments. For MHC class I–restricted antigen presentation, particularly for responses to pathogens that do not infect DCs or become systemic, and for responses to tumors that are sequestered in distant sites in the body, this *exogenous presentation* poses a challenge because the site of peptide binding to class I is in the ER, generally not thought to be readily accessible to internalized proteins.

In the last decades, an auxiliary pathway of antigen presentation, termed *cross-presentation,* has been discovered. This antigen presentation event for CD8 T-cell priming to exogenous antigens involves uptake and class I–restricted presentation of cell-associated antigens from tumors or virus-infected cells, and those associated with intracellular bacteria pathogens, such as listeria or fungal pathogens whose proteins are not synthesized within the APC. Cross-presentation also serves to prime CD8 T cells reactive with pathogen-derived proteins and aggregates released by infected cells in distal sites, such as the respiratory tract.[37,38]

Internalization of cell- or pathogen-associated antigens into endosomal/lysosomal compartments through phagocytosis or receptor-mediated uptake initiates cross-presentation and ultimately leads cell-surface presentation of the pathogen-derived peptides in association with class I molecules, leading to CD8 T-cell activation, cell division, and differentiation. These primed CD8 T cells ultimately become programmed to deliver their effector function, most commonly cytolysis, on either pathogen-infected cells or transformed cells, eliminating them from the host.

The intervening, intracellular events involved in cross-presentation are the subject of active research. Early studies pointed to the possibility that this presentation occurs in only a subset of DCs, now known to be distinguished by markers, such as CD8α^+ in mice and BDCA3$^+$ cells in humans. Ablation of DCs *in vivo* eliminates cross-presentation of pathogen-associated antigens, emphasizing the need for this cell type for efficient cross-priming.

There are a number of studies that suggest that cross-presentation is dependent on TAP and the proteasome because in many cases, presentation of the epitope is sensitive to inhibitors of these proteins. Because of the shared cellular cofactors involved in peptide generation and loading, it is envisioned that the same peptides would be presented by the endogenous and the cross-presentation route. This conclusion would seem to be most beneficial to effective surveillance and protection through recognition of the target cell synthesizing the viral or tumor antigen. This cytosolic cross-presenting route requires that internalized or phagocytosed proteins from pathogen-infected or transformed cells gain access to the cytosol. A mechanism that allows transport of an intact or partially unfolded/degraded protein from the lumen of lysosomal compartments into the cytosol is therefore needed. Export has been speculated to involve a transmembrane channel or a direct extraction mechanism. Proteins, such as those involved in ER-associated degradation (ERAD), have been implicated by many studies, as has the Sec61 translocon. Cross-presentation has also been shown to require an endosomal, pH-dependent reductase (GILT), suggesting that protein unfolding may be needed for antigenic protein transfer.

After crossing the lysosomal membrane, the cytosolically localized proteins can now enter the classical or *endogenous* pathway of class I–restricted presentation: proteolysis by the proteasome, import of the derived peptides into the ER via TAP, and peptide binding being facilitated/edited by the components of PLC. The peptide class I complex would then be exported from the ER to the cell surface via the typical default secretory pathway (Fig. 6.9).

The *vacuolar pathway* provides as an alternative means for cross-presentation. In this pathway, the internalized pathogen or tumor-derived proteins remain in the endosomal compartments, where they are degraded by lysosomal proteases. Class I molecules at the cell surface then encounter the antigens during internalization and recycling. Some evidence also suggests that components of the ER-associated PLC may localize to the endosomal compartments and intersect with MHC class I, allowing assembly of the complex within these compartments by a similar mechanism as in the ER. Under what conditions, within what cells *in vivo,* and with what antigens these two potential pathways of exogenous presentation are operational remain to be defined.

Both pathways of cross-presentation require that the internalized antigen be protected from rapid and terminal degradation. Alkalization of the compartment diminishes recruitment of enzymes responsible for acidification of lysosomal compartments, such as nicotinamide adenine dinucleotide phosphate (NADPH) oxidase 2 (NOX-2). Most of the lysosomal enzymes are pH dependent and require acidic pH for optimal activity. Strategies to modify this acidification, either by genetic means or by pharmacological approaches, have been shown to promote cross-presentation.

THE MHC-PRESENTED PEPTIDOME

Isolation and sequencing of host peptides constitutively bound and presented by MHC molecules has yielded significant new insight regarding the diversity, abundance, and source of peptides that are normally displayed.[39,40] The key technical advance was the development and use of tandem mass spectroscopy to sequence very low abundance peptides.[41] In this approach, MHC class I or class II molecules of interest are isolated by antibody affinity chromatography. The bound peptides are then eluted, and the low-molecular-weight material, including peptides, are separated from antibody and class I and then subjected to sequencing.

From these studies, it is now estimated that as many as 10,000 different self peptides are in association with a given class II molecule on the cell surface. For human APCs that can express 6–10 different class II species (HLA-DR, HLA-DQ, and HLA-DP in multiple allelic forms), this diversity of different peptide–class II complexes displayed by APCs was unexpected.

Several key features of MHC biology and biochemistry were revealed by this new technology. First, peptides derived from alternative alleles of MHC proteins yield nonoverlapping sets of peptides. This permits assignment of peptide binding preferences or motifs to different allelic and isotypic forms of MHC molecules (in humans, HLA-A, HLA-B, and HLA-C for class I, and HLA-DR, HLA-DQ, and HLA-DP for class II). This advance has enabled development and refinement of algorithms that predict pathogen-associated epitopes as well as the definition of preferred peptide-binding characteristics for MHC proteins that are associated with susceptibility to autoimmune disease or infection. Sequencing of the peptidome presented by MHC class II molecules also modified our view of class II presented peptides, indicating that the bound peptides extend far beyond the exogenous pathway of antigen access. Many peptide epitopes presented by class II molecules are now known to be derived from cytosolic, nuclear, and membrane-associated proteins synthesized within the APC.

The discovery that antigenic peptides from vaccines, pathogens, or transformed cells represent just a few of the hundreds to thousands of self peptides presented on the surface of APCs supports the current view that the TCR must be capable of acute discriminatory power and behaves more as a signaling protein than a ligand-binding protein. With clarification of the diverse nature of the peptides used for central tolerance induction, tumor-specific antigens and peptides have been identified that can be used for specific immunotherapy. And, methodology to allow sequencing of peptides from just a few million cells from the respiratory tract has led to the identification of respiratory viral peptides presented by the HLA, offering new strategies for vaccine design.

KEY CONCEPTS

Viral Antagonists of Antigen Presentation

- Many viruses, particularly those that lead to chronic infections, encode multiple proteins that antagonize MHC-restricted antigen presentation.
  - Most viral evasion proteins inhibit MHC class I pathways needed for CD8 T-cell recognition.
  - Some human immunodeficiency virus (HIV)–encoded proteins antagonize class II–restricted presentation, as well.
- Viral antagonists typically target the proteins most critical for presentation, including the TAP (for transporter associated with antigen processing) transporter, tapasin, and the MHC molecule itself.

Cryptic Viral or Tumor-Associated Peptides Presented by MHC Class I Molecules

The *transcriptome* consists of the known translated sequence from a viral or endogenous self antigen. An increasing number of tumor- and viral-derived peptides have been identified that are not represented by the transcriptome. These cryptic peptides are derived from many different molecular mechanisms,[40,42] including alternate reading frames, translation of introns, antisense transcript encoded peptides, 5’ untranslated regions, and long noncoding RNAs. In many cases, cytotoxic CD8 T cells have been identified that recognize these peptides. For example, nearly 25% of epitopes from Simian immunodeficiency virus in rhesus macaques appear to be derived from translation of alternative reading frames. Perhaps most surprising is the discovery that the proteasome can ligate distal peptides together through a transpeptidation event, creating new epitopes from noncontiguous segments.

Besides the interest in understanding the precise mechanism responsible for generation of these epitopes is the practical implication of these events. Epitope identification often involves derivation and testing of overlapping peptide libraries from the pathogen or tumor-derived proteins that are based on known protein databases. The cryptic epitopes will not be represented in these libraries, and accordingly, epitopes recognized by many potential effector CD8 T cells will fail to be identified. Thus future vaccine strategies will need to take into account the probability that some of the existing and potentially protective host CD8 T cells may be dedicated to unpredicted peptides from pathogens or transformation.

CLINICAL RELEVANCE

Features of MHC Restricted Presentation that are Important for Clinical and Translational Studies

- Many viral- and tumor-derived peptides for CD8 T-cell recognition are derived from noncoding sequences, alternative reading frames, and antisense coding segments.
- Vaccine strategies must take into account the need for linked recognition of B-cell and CD4 T-cell epitopes and also the selectivity of CD4 T-cell responses that focus only on peptides that bind with very high affinity on the class II molecule.
- Self tolerance within the thymus is selective for peptides that survive tapasin editing for class I presentation and DM editing for class II presentation.
- Autoantigenic peptides that escape self tolerance induction in the thymus often have very low affinity interactions with MHC class II molecules. This can make them difficult to identify or study with MHC–peptide tetramers.
- Chronic viral infection frequently leads to downregulation of the expression of MHC class I–peptide complexes. This leads to escape from immune detection and destruction of infected cells by CD8 T cells.

Pathogen Evasion Strategies

Viral and bacterial pathogens employ extensive evasion strategies that to avoid immune recognition[43-45] (Chapters 24–26). A number of viral proteins, particularly those involved in chronic viral infections, interfere with components of the antigen presentation pathway. Many prevent peptide access to MHC class I molecules, which results in both immune evasion and greatly diminished cell-surface expression of class I molecules.

For example, herpes simplex virus (HSV)– and EBV-encoded proteins both interfere with the binding of peptide to the ER localized TAP transporter. This blocks peptide transport to class I molecules from the site of generation in the cytosol to the ER. The requisite ATP-dependent conformation change of TAP, needed for peptide translocation from the cytosol to the ER, is also a target of viral antagonists. ATP activity is blocked by human cytomegalovirus (hCMV) US6 and EBV-derived BNLF2a. The importance of tapasin in protective CD8 T-cell immunity is evident by the ability of hCMV US3 to bind directly to it and prevent tapasin-dependent loading of antigenic peptides. There are also numerous viral proteins that directly target MHC class I for ER retention, such as adenovirus E3-E19, or degradation directly from the site of synthesis in the ER. Finally, degradation of MHC class I molecules that have successfully emerged to the cell-surface expression is promoted by viral proteins that mediate polyubiquitination and targeting for lysosomal degradation (Kaposi sarcoma–associated herpes viruses Kk3/MIR1 and Kk5/

MIR2 and EBV BILF1). Together, these and other mechanisms prevent detection by CD4 and CD8 T cells and the ensuing protective immune response.

Tumor Escape From Immune Surveillance

Tumor progression is often associated with alterations in the expression and/or function of the MHC class I molecules.[46] Loss of DNA binding factors needed for promoter activation can lead to altered class I transcription. Loss of β2m gene expression can lead to a global defect in class I protein expression at the cell surface. Downregulation of the components of the constitutive and immune-proteasome have been identified in a number of cancers. Low levels of TAP transporter components and the ER aminopeptides (ERAAP) have been seen in others. Loss or diminution of class I presentation of peptides is often associated with disease progression, metastasis, and a detrimental clinical outcome in part as a result of evasion of immune detection and impaired elimination of the transformed cells.

Potential Role of "Peptide Editing" of Presented Peptides on Self Tolerance and Autoimmunity

HLA-DM *edits* the repertoire of peptides presented by HLA-class II molecules at the cell surface. These are the only peptides that are available to interact with circulating CD4 T cells. Tapasin appears to have a similar function for MHC class I peptide binding, enriching for presentation of peptides that bind with high affinity with the MHC class I molecule. This selective peptide display by MHC molecules has important implications for self tolerance and autoimmunity. Tolerance to self arises as a consequence of deletion of developing T cells whose TCR recognizes self peptides presented by host MHC molecules within the thymus (Chapter 8). This is termed *central tolerance.* Accordingly, central tolerance can only be operative if a self peptide is displayed at the cell surface of APCs in the thymus. The bone marrow-derived MHC class II and MHC class I positive APCs responsible for deletion of self-reactive cells express the HLA-DM and tapasin proteins and thus will only express peptides that bind with high affinity to MHC class II molecules. Therefore the host is selectively tolerant only to host peptides that survive this intracellular editing during biosynthesis of the MHC proteins. Self peptides that are removed from MHC molecules during endosomal DM editing or ER-associated tapasin editing are essentially invisible to developing T cells. T cells that have escaped deletion in the thymus will seed the peripheral pool of circulating T cells. The mature CD4 and CD8 T cell repertoire therefore maintains reactivity to some self peptides. Events in the periphery that modify presentation of this subset of lower affinity peptides for MHC molecules will thus allow their display at the cell surface potentially leading to recognition by host CD8 or CD4 T cells and thus initiation or propagation of autoreactivity.

SUMMARY

The intracellular events involved in MHC class I and class II restricted antigen presentation have many commonalities, although they are different in the details. Both can be considered three-chain proteins, where the bound peptide is the third chain. Binding of peptide is needed for conformational integrity of the protein, and accordingly, self or antigenic peptides must be present throughout biosynthesis. Both have a primary site of peptide acquisition: For class I molecules, it is early in biosynthesis within the ER, and for class II, it is within acidic endosomal compartments accessed later during biosynthesis of the MHC protein. Within each of these compartments resides a set of proteins that participate in and direct peptide loading. Both class I and class II molecules acquire peptides inefficiently, and accordingly both require critical protein cofactors that prepare the MHC protein to capture peptide. The presentation events are also similar in that they allow only a subset of the potential self or pathogens peptides to be presented. Peptide editing of the repertoire of peptides presented by MHC proteins is primarily based on the ability to stably interact with the MHC protein.

ON THE HORIZON

- New means to regulate class II antigen presentation could come from a better understanding of the mechanisms by which HLA-DO regulates HLA-DM or from the elucidation of new mechanisms by which viral or bacterial products antagonize the MHC class II presentation pathway.
- New means to regulate class I antigen presentation could come from a better understanding of the mechanisms used to regulate proteasome generated cytosolic peptides and their transport by TAP (for transporter associated with antigen processing) to MHC class I molecules and from a better understanding of cross-presentation.
- A better understanding of the similarities and differences in antigen presentation by the various APC subsets could lead to new vaccination methods.
- A better understanding of the role of cryptic peptides in CD8 T-cell selection events could lead to the manipulation of CD8 T-cell responses to tumors, viruses, and autoantigens.

Because of MHC polymorphism, the peptides selected for protection, self tolerance, and autoimmunity will vary from one individual to the next. Many pathogens have developed specific proteins that degrade or interfere with the essential functions of the protein cofactors that regulate MHC-restricted antigen presentation. These viral evasions strategies allow the infected host cell to avoid immune recognition and destruction by CD8 and CD4 T cells. Interference with antigen presentation also leads to the escape of host cells from tumor-specific surveillance and allow progressive growth and metastases.

We now understand many of the key players in MHC-restricted presentation of antigen to CD4 and CD8 T cells. In the future this understanding will be used for vaccine design strategies, and in personalized medicine aimed at enhancing immunosurveillance and protection mediated by the adaptive immune response.

ACKNOWLEDGEMENT

With the permission of authors J. Rodgers and R. Rich, the section of this chapter that discusses the nature of antigens was reproduced, with modification, from Chapter 6 of the 4th edition of *Clinical Immunology: Principles and Practices.*

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Silverstein AM. Paul Ehrlich's Receptor Immunology: The Magnificent Obsession. New York: Academic Press; 2001.
2. Kawai T, Akira S. Toll-like receptors and their crosstalk with other innate receptors in infection and immunity. Immunity 2011;34:637–50.
3. Desai SN, Clemens JD. An overview of cholera vaccines and their public health implications. Curr Opin Pediatr 2012;24:85–91.

4. Baldauf KJ, Royal JM, Hamorsky KT, et al. Cholera toxin B: one subunit with many pharmaceutical applications. Toxins (Basel) 2015;7:974–96.
5. Pierson TC, Graham BS. Zika Virus: Immunity and Vaccine Development. Cell 2016;167:625–31.
6. Sadanand S, Suscovich TJ, Alter G. Broadly Neutralizing Antibodies Against HIV: New Insights to Inform Vaccine Design. Annu Rev Med 2016;67:185–200.
7. Ren H, Zhou P. Epitope-focused vaccine design against influenza A and B viruses. Curr Opin Immunol 2016;42:83–90.
8. Cicala C, Nawaz F, Jelicic K, et al. HIV-1 gp120: A Target for Therapeutics and Vaccine Design. Curr Drug Targets 2016;17:122–35.
9. Zarei AE, Almehdar HA, Redwan EM. Hib Vaccines: Past, Present, and Future Perspectives. J Immunol Res 2016;2016:7203587.
10. Baker CJ. Prevention of Meningococcal Infection in the United States: Current Recommendations and Future Considerations. J Adolesc Health 2016;59:S29–37.
11. Proft T, Fraser JD. Bacterial superantigens. Clin Exp Immunol 2003;133:299–306.
12. Sundberg EJ, Deng L, Mariuzza RA. TCR recognition of peptide/MHC class II complexes and superantigens. Semin Immunol 2007;19:262–71.
13. Fraser JD, Proft T. The bacterial superantigen and superantigen-like proteins. Immunol Rev 2008;225:226–43.
14. Deshane J, Chaplin DD. Follicular dendritic cell makes environmental sense. Immunity 2010;33:2–4.
15. Heesters BA, van der Poel CE, Das A, et al. Antigen Presentation to B Cells. Trends Immunol 2016;37:844–54.
16. van Kasteren SI, Overkleeft H, Ovaa H, et al. Chemical biology of antigen presentation by MHC molecules. Curr Opin Immunol 2014;26:21–31.
17. Blum JS, Wearsch PA, Cresswell P. Pathways of antigen processing. Annu Rev Immunol 2013;31:443–73.
18. Unanue ER, Turk V, Neefjes J. Variations in MHC Class II Antigen Processing and Presentation in Health and Disease. Annu Rev Immunol 2016;doi:10.1146/annurev-immunol-041015-055420.
19. Roche PA, Furuta K. The ins and outs of MHC class II-mediated antigen processing and presentation. Nat Rev Immunol 2015;15:203–16.
20. Stumptner-Cuvelette P, Benaroch P. Multiple roles of the invariant chain in MHC class II function. Biochim Biophys Acta 2002;1542:1–13.
21. ten Broeke T, Wubbolts R, Stoorvogel W. MHC class II antigen presentation by dendritic cells regulated through endosomal sorting. Cold Spring Harb Perspect Biol 2013;5:a016873.
22. Denzin LK. Inhibition of HLA-DM Mediated MHC Class II Peptide Loading by HLA-DO Promotes Self Tolerance. Front Immunol 2013;4:465.
23. Mellins ED, Stern LJ. HLA-DM and HLA-DO, key regulators of MHC-II processing and presentation. Curr Opin Immunol 2014;26:115–22.
24. Pos W, Sethi DK, Wucherpfennig KW. Mechanisms of peptide repertoire selection by HLA-DM. Trends Immunol 2013;34:495–501.
25. Sant AJ, Chaves FA, Leddon SA, et al. The control of the specificity of CD4 T cell responses: thresholds, breakpoints, and ceilings. Front Immunol 2013;4:340.
26. Schulze MS, Wucherpfennig KW. The mechanism of HLA-DM induced peptide exchange in the MHC class II antigen presentation pathway. Curr Opin Immunol 2012;24:105–11.
27. Caminschi I, Shortman K. Boosting antibody responses by targeting antigens to dendritic cells. Trends Immunol 2012;33:71–7.
28. McGreal EP, Miller JL, Gordon S. Ligand recognition by antigen-presenting cell C-type lectin receptors. Curr Opin Immunol 2005;17:18–24.
29. Lehmann CH, Heger L, Heidkamp GF, et al. Direct delivery of antigens to dendritic cells via antibodies specific for endocytic receptors as a promising strategy for future therapies. Vaccines (Basel) 2016;4.
30. Yuseff MI, Lennon-Dumenil AM. B Cells use Conserved Polarity Cues to Regulate Their Antigen Processing and Presentation Functions. Front Immunol 2015;6:251.
31. Avalos AM, Ploegh HL. Early BCR Events and Antigen Capture, Processing, and Loading on MHC Class II on B Cells. Front Immunol 2014;5:92.
32. Heesters BA, Myers RC, Carroll MC. Follicular dendritic cells: dynamic antigen libraries. Nat Rev Immunol 2014;14:495–504.
33. Duraes FV, Niven J, Dubrot J, et al. Macroautophagy in Endogenous Processing of Self- and Pathogen-Derived Antigens for MHC Class II Presentation. Front Immunol 2015;6:459.
34. Kniepert A, Groettrup M. The unique functions of tissue-specific proteasomes. Trends Biochem Sci 2014;39:17–24.
35. Sijts EJ, Kloetzel PM. The role of the proteasome in the generation of MHC class I ligands and immune responses. Cell Mol Life Sci 2011;68:1491–502.
36. Shastri N, Nagarajan N, Lind KC, et al. Monitoring peptide processing for MHC class I molecules in the endoplasmic reticulum. Curr Opin Immunol 2014;26:123–7.
37. Schuette V, Burgdorf S. The ins-and-outs of endosomal antigens for cross-presentation. Curr Opin Immunol 2014;26:63–8.
38. Segura E, Amigorena S. Cross-Presentation in Mouse and Human Dendritic Cells. Adv Immunol 2015;127:1–31.
39. Stern LJ, Santambrogio L. The melting pot of the MHC II peptidome. Curr Opin Immunol 2016;40:70–7.
40. Granados DP, Laumont CM, Thibault P, et al. The nature of self for T cells-a systems-level perspective. Curr Opin Immunol 2015;34:1–8.
41. Appella E, Padlan EA, Hunt DF. Analysis of the structure of naturally processed peptides bound by class I and class II major histocompatibility complex molecules. EXS 1995;73:105–19.
42. Starck SR, Shastri N. Non-conventional sources of peptides presented by MHC class I. Cell Mol Life Sci 2011;68:1471–9.
43. Hansen TH, Bouvier M. MHC class I antigen presentation: learning from viral evasion strategies. Nat Rev Immunol 2009;9:503–13.
44. van de Weijer ML, Luteijn RD, Wiertz EJ. Viral immune evasion: Lessons in MHC class I antigen presentation. Semin Immunol 2015;27:125–37.
45. Goldberg MF, Saini NK, Porcelli SA. Evasion of Innate and Adaptive Immunity by *Mycobacterium tuberculosis*. Microbiol Spectr 2014;2.
46. Leone P, Shin EC, Perosa F, et al. MHC class I antigen processing and presenting machinery: organization, function, and defects in tumor cells. J Natl Cancer Inst 2013;105:1172–87.

MULTIPLE-CHOICE QUESTIONS

1. How does the expression of human leukocyte antigen (HLA)-DO most commonly modify the peptide repertoire presented by major histocompatibility complex (MHC) class II molecules?
   A. It enhances presentation of most peptide epitopes.
   B. It targets MHC class II presentation to the correct endosomal compartment.
   C. It inhibits peptide editing by HLA-DM.
   D. It promotes release of the class II–associated invariant chain peptide (CLIP) fragment from MHC class II molecules.

2. How do the proteasome and immunoproteasome differ?
   A. They differ primarily in their subcellular localization.
   B. They differ with regard to targeting of ubiquitinated ligands.
   C. They differ in the number of subunits they contain.
   D. They differ in the sequence and structure of the β subunit.

3. How does cross-presentation differ from the classical pathway of class I–restricted antigen presentation?
   A. It involves exogenous antigens gaining access to MHC class I molecules.
   B. It utilizes the TAP (transporter associated with antigen processing) transporter.
   C. It involves degradation by the proteasome.
   D. It uses only one subset of class I molecules for presentation.

4. What is the function of invariant chain in MHC class II biology?
   A. It prevents class II acquisition of peptides made in the endoplasmic reticulum.
   B. It targets class II molecules to endosomal compartments.
   C. It helps degrade antigens internalized through the immunoglobulin receptor.
   D. A and B

7

B-Cell Development and Differentiation

Harry W. Schroeder, Jr., Andreas Radbruch, Claudia Berek

B lymphocytes arise from multipotent hematopoietic stem cells that successively populate the embryonic paraaortic splanchnopleure, the fetal liver, and then bone marrow. Stem-cell daughter cells give rise to lymphoid-primed multipotent progenitors (LMPPs), which, in turn, can give rise to either myeloid or lymphoid cells.[1,2] LMPPs then produce common lymphoid progenitors (CLPs), which can generate T cells, B cells, natural killer (NK) cells, and dendritic cells (DCs). Final B-cell differentiation requires the exposure of CLP daughter cells to specialized microenvironments, such as those found in the fetal liver and bone marrow. These two tissues are the primary B-lymphoid organs. In humans, the shift from fetal liver to bone marrow begins in the middle of fetal life and ends just prior to birth. B cells continue to be produced in bone marrow throughout the life of the individual, although the rate of production decreases with age.

KEY CONCEPTS

B-Cell Development in the Primary Lymphoid Organs

- Commitment to the B-cell lineage reflects differential activation of transcription factors that progressively lock the cell into the B-cell pathway.
- B-cell development is typically viewed as a linear, stepwise process that is focused on the assembly and testing of immunoglobulin function, first in the fetal liver and bone marrow and then in the periphery:
  - Failure to assemble a functional receptor leads to cell death.
  - Expression of a functional receptor subjects the B cell to antigen selection.
  - B cells with inappropriate specificities tend to be eliminated.
  - B cells responding appropriately to external antigen can develop either into immunoglobulin-secreting plasma cells or into memory cells.
- At the clinical level, B-cell development can be monitored by examining the pattern of expression of characteristic surface proteins.

An intact and functional B-cell receptor (BCR) complex, which consists of membrane-bound immunoglobulin (mIg), the Igα and Igβ coreceptors, and ancillary signal transduction components, must be present for the developing B cell to survive (Chapter 4). The composition of the BCR is subject to intense selection. In the primary organs, hazardous self-reactive BCRs, as well as nonfunctional ones, can be culled by changing the light (L) chain (receptor editing), by cell anergy,[3] or by apoptosis of the host cell. Survivors of this initial selection process are released into blood and thence to the spleen, lymph nodes, and other secondary lymphoid tissues and organs, where selection for specificity continues (Chapter 2).

B-cell differentiation (Fig. 7.1) is commonly presented as a linear process defined by the regulated expression of specific sets of transcription factors, immunoglobulin (Ig), and cell-surface molecules. Given the central role of the BCR (Chapter 4), initial developmental steps are classically defined by the status of the rearranging Ig loci. With the development of monoclonal antibody (mAb) technology, analysis of cell-surface markers, such as CD10, CD19, CD20, CD21, CD24, CD34, and CD38 (Fig. 7.2), has facilitated definition of both early and late stages of development, especially in those cases where Ig cannot be used to distinguish between cell types.[4] Of these, CD19, a signal transduction molecule expressed throughout B-cell development up to, but not including, the mature plasma cell stage[5] warrants special mention as the single best clinical marker for B-cell identity.

In practice, B-cell development is a more complex process than the simple, linear pathways depicted in Figs. 7.1 and 7.2. For example, proB cells typically derive from a CLP, but they can also develop from a bipotent B/macrophage precursor. Thus B-lineage subsets identified by one fractionation scheme may consist of mixtures of subsets identified by others. The practitioner would therefore be wise to clarify the fractionation scheme used by the reference laboratory when comparing patient findings to the literature.

Initial commitment to the B-cell lineage requires activation of a series of transcriptional and signal transduction pathways. At the nuclear level, the transcription factors PU.1, Ikaros, ID-1, E2A, EBF, and PAX5 play major roles in committing progenitor cells to the B-cell lineage.[6] After lineage commitment has been established, however, it is generally accepted that the composition of the BCR controls further development.

Each B-cell progenitor has the potential to produce a large number of offspring. Some will develop into mature B cells, and even less into long-lived memory B cells or plasma cells. Others, indeed the majority, will perish.[7] Most of the defined steps in this process of development represent population bottlenecks: developmental checkpoints wherein the developing B cell is tested to make sure that its BCR will be beneficial.[8] In the periphery, exposure to antigen is associated with class switching and hypermutation of the variable domains of the antigen receptor. A few select survivors earn long lives as part of a cadre of memory B cells. These veterans are charged with the responsibility to rapidly engage antigen to which they have been previously exposed, providing experienced protection against repeated assault.

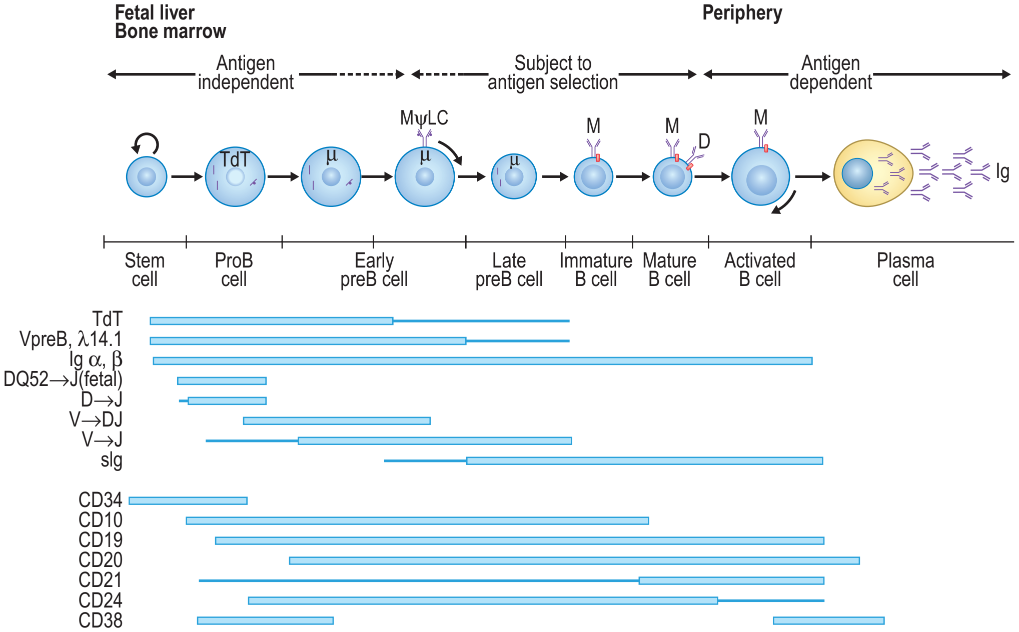
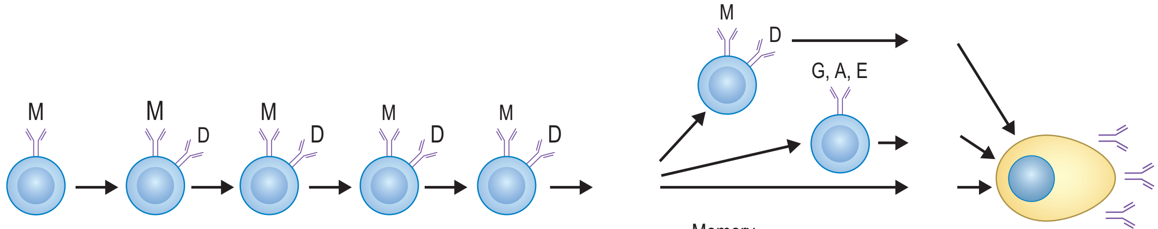

Fetal liver
Bone marrow
Periphery
Antigen independent
Subject to antigen selection
Antigen dependent
TdT
μ
MψLC
μ
μ
M
M
D
M
Ig
Stem cell
ProB cell
Early preB cell
Late preB cell
Immature B cell
Mature B cell
Activated B cell
Plasma cell
TdT
VpreB, λ14.1
Ig α, β
DQ52→J(fetal)
D→J
V→DJ
V→J
sIg
CD34
CD10
CD19
CD20
CD21
CD24
CD38

FIG 7.1 **Model of B-Cell Differentiation.** B-cell development is typically viewed as a linear progression through different stages of differentiation. The various processes associated with the assembly of the B-cell receptor (BCR) complex and the expression pattern of surface molecules whose presence or absence are illustrated through use of bars. The various steps in immunoglobulin (Ig) rearrangement and the pattern of expression of these surface molecules can be used to characterize stages in B-cell development.


M
M
D
M
D
M
D
M
D
M
D
G, A, E

	Immature	Transitional			Naive mature	Memory Unswitched (MZ-like)	Memory switched	Plasmablast
		T1	T2	T3				
IgM	++	+++	++	+	+	+	-	-
IgD	-	+	++	++	++	+	-	-
R123	+	+	+	Int	-	+	+	+
CD10	+	+	+/-	-/+	-	-	-	+/-
CD19	+	+	+	+	+	+	+	lo
CD24	+++	+++	++	+	+	+	-	-
CD27	-	-	-	-	-	+	+	hi
CD38	+++	+++	++	+	+	lo	lo	hi

FIG 7.2 **B-Cell Subsets in Peripheral Blood.** Peripheral blood B-cell subsets can be identified by differential staining for immunoglobulin M (IgM), IgD, CD10, CD19, CD24, CD27, and CD38 and by the use of the rhodamine dye R123, which is extruded by the ABCB1 transporter expressed in naïve, mature B cells and, to a lesser extent, T3 transitional B cells.[36] B cells that are CD20$^+$ CD27$^+$CD43$^+$CD70$^-$ are currently the best candidates for the human B-1-cell counterparts,[4] but this is subject to change. In addition to these subsets found in blood, germinal center B cells found in the spleen, and lymph nodes are characterized as IgD$^-$CD38^{++} CD10$^+$ CD27$^{+/-}$; marginal zone (MZ) B cells are typically CD27$^+$CD21^{++}CD23$^{+/-}$CD1c$^+$IgD$^-$; and long-lived plasma cells found in bone marrow, the spleen, and tonsils are CD138$^+$CD20$^-$CD38^{++}.

Specialized microenvironments also play a role in peripheral B-cell development (Chapter 2), each of which enables the B cell to properly engage different types of antigens or venues of attack. In the marginal zone (MZ), mature splenic MZ B cells await bacterial pathogens. In the lymphoid follicles, B cells reactive with a given antigen collaborate with follicular T helper (Tfh) cells and DCs to maximize the immune response (Chapter 6). In the germinal centers (GCs), B cells use class switching and somatic mutation to modify and optimize the function and affinity of their Igs. And, underneath mucosal surfaces, B cells are primed to express IgA (Chapter 20).

B-CELL DEVELOPMENT BEGINS IN THE PRIMARY LYMPHOID ORGANS

Generation of a Functioning Antigen Receptor Is Key to the Viability of a B Cell

Ig rearrangement is hierarchical. In proB cells, heavy (H) chain $D_H \rightarrow J_H$ joining precedes $V_H \rightarrow DJ_H$ rearrangement (Chapter 4), followed by light (L) $V_L \rightarrow \rightarrow J_L$ joining in late-stage preB cells.

Production of a properly functioning B-cell receptor (BCR) is essential for development beyond the preB cell stage. For example, loss-of-function (LOF) mutations in RAG-1/2 and DNA-dependent protein kinase (DNA-PKcs, Ku 70/80) preclude B-cell as well T-cell development (Chapter 35). Each proB cell faces the probability that only one of three possible splices will place the V_H and J_H in the same reading frame. The opportunity to try rearrangement on the second chromosome gives failing proB cells a second opportunity. Together, this provides the cell with five chances out of nine for initial survival ($\frac{1}{3} + [\frac{1}{3} \times \frac{2}{3}] = \frac{5}{9}$). This calculation reflects the fact that after failure of the first rearrangement for $\frac{2}{3}$ of the cells, these remaining cells will each have a 1 in 3 chance of a functional rearrangement of the second chromosome. In-frame, functional VDJ_H rearrangement allows the proB cell to produce μ H chains, most of which are retained in the endoplasmic reticulum (ER). The appearance of cytoplasmic μ H chains marks initiation of the preB cell stage. These early preB cells tend to be large in size.

VpreB and λ14.1 [λ5], which together form the surrogate light chain (ψLC), and Igα and Igβ (Chapter 4) are constitutively expressed by proB cells. The first H chain quality control checkpoint tests for the ability of the μ H chain to associate with surrogate light chain to form a preB-cell receptor. In addition to checking to see if the scaffolding (frameworks) of the L chain can associate correctly with the scaffolding of the H chain, VpreB encodes a sensing site that can test the H chain antigen-binding site. Thus the surrogate light chain functions as the first, and invariant, antigen to screen for antigen-binding characteristics.

Successful formation of a stable preBCR is followed by the termination of further H chain rearrangement (*allelic exclusion*), which is followed by four to six cycles of cell division, a process associated with a progressive decrease in cell size. Late preB daughter cells reactivate recombinase activating gene 1 (RAG1) and RAG2 and begin to undergo $V_L \rightarrow J_L$ rearrangement. Successful production of a complete κ or λ light chain permits expression of conventional IgM on the cell surface (sIgM), which identifies the *immature* B cell. Immature B cells expressing self-reactive IgM antibodies may undergo repeated rounds of light chain rearrangement to lessen the self specificity of the antibody, a process termed *receptor editing.*

Immature B cells that have successfully produced an acceptable IgM BCR extend transcription of the H chain locus to include the Cδ exons downstream of Cμ. Alternative splicing permits coproduction of IgM and IgD. These now newly *mature* IgM^+IgD^+ B cells enter blood and migrate to the periphery, where they form the majority of the B-cell pool in the spleen and the other secondary lymphoid organs. The IgM and IgD on each of these cells share the same variable domains.

Tyrosine Kinases Play Key Roles in B-Cell Development

Signaling through the BCR is required for continued development. Bruton tyrosine kinase is an important component of the phospholipase Cγ (PCLγ) pathway, which is used in BCR signaling. Deficiency of BTK function results in the arrest of human B-cell development at the preB cell stage[9] and is the genetic basis of X-linked agammaglobulinemia (XLA) (Chapter 34).

BLNK is a SRC homology 2 (SH2) domain–containing signal transduction adaptor. When phosphorylated by SYK, BLNK serves as a scaffold for the assembly of cell activation targets that include GRB2, VAV, NCK, and phospholipase C-[γ] (PLCγ). LOF mutations in BLNK can result in the loss of preB and mature B cells and thus agammaglobulinemia.

FLT3 (FLK2) is a receptor tyrosine kinase belonging to the same family as c-FMS, the receptor for colony stimulating factor-1 (CSF-1). FLT3 ligand, which has homology to CSF-1, is a potent costimulator of early proB cells. In mice, targeted disruption of *flt3* leads to a selective deficiency of primitive B-cell progenitors.

Cell Surface Antigens Associated With B-Cell Development

B-cell development is associated with the expression of a cascade of surface proteins, each of which plays a key role in the fate of the cell (see Fig. 7.1; Table 7.1). The timing of the appearance of each of these proteins can be used to further analyze the process of B-cell development.

CD34 is a highly glycosylated type I transmembrane glycoprotein that binds to CD62L (L-selectin) and CD62E (E-selectin) and thus likely aids in cell trafficking (Chapter 11). It is expressed on a small population (1–4%) of bone marrow cells that includes hematopoietic stem cells (HSCs). Minimal hematopoietic defects have been documented in mice deficient in CD34. However, such observations must be viewed with caution when extrapolated to humans because CD34 is not expressed on HSCs in mice.

CD10, also known as *neprilysin, neutroendopeptidase,* and the *common acute lymphocytic leukemia antigen* (CALLA), is a type II membrane glycoprotein metalloprotease. CD10 has a short *N*-terminal cytoplasmic tail, a signal peptide transmembrane domain, and an extracellular *C*-terminal domain that includes six *N*-linked glycosylation sites. The extracellular domain contains 12 cysteines whose disulfide bonds help stabilize its zinc-binding pentapeptide motif, which is involved in its zinc-dependent metalloprotease catalytic activity. By virtue of its protease activity, it is thought to downregulate cellular responses to peptide hormones and cytokines. Inhibition of CD10 activity on bone marrow stromal cells enhances B-cell maturation. CD10 (CALLA) is used as a marker for preB acute lymphocytic leukemias (ALLs) and for certain lymphomas.

CD19 is a cell-surface glycoprotein of the immunoglobulin superfamily (IgSF) that is exclusively expressed throughout B-cell development from the proB-cell stage up to the plasma cell stage (see Fig. 7.1).[10] CD19 exists in a complex with CD21 (complement receptor 2: CDR2), CD81 (TAPA-1), and Leu 13. With the help of CD21, CD19 can bind the complement C3 cleavage product C3d. The simultaneous binding of sIgM and CD19 to a C3d-antigen complex enables CD19 and the BCR to

TABLE 7.1 Cell-Surface Proteins Active in Early B-Cell Development

Gene	Class or Alternative Name	Associated or Targeted Genes or Molecules	B-Cell Developmental Phenotype in Humans or *Mice* Associated With Disrupted Function of the Indicated Gene
B-Cell Receptor Complex			
μ chain	Immunoglobulin superfamily (IgSF)	κ,λ L chains, ψL chain, CD79 a,b (Igα,β)	AGM1: agammaglobulinemia and no B cells. *Arrest at preB-cell stage*
Immunoglobulin λ-like polypeptide 1; IGLL1 (λ14.1, λ5)	IgSF	VpreB, μ H chain	AGM2: agammaglobulinemia and reduced B-cell numbers. *Arrest at preB-cell stage*
VPREB1 (VpreB)	IgSF	λ14.1, μ H chain	*Arrest at preB-cell stage*
CD79a,b (Igα,β)	IgSF, cytoplasmic immunoreceptor tyrosine-based activation motifs (ITAMs)	H chain, LYN, FYN, BLK, SYK	AGM3 (CD79A); AGM6 (CD79B): Agammaglobulinemia and arrest at proB-cell stage. *Arrest at proB-cell stage*
Other Cell-Surface Proteins			
CD10	Type II metalloproteinase	Hydrolyzes peptide hormones, cytokines	*Not expressed in murine B-cell progenitors*
CD19	IgSF	mIgM, PI-3 kinase, VAV, LYN?, FYN?	CVID3: panhypogammaglobulinemia, normal numbers of $CD20^{+}$ B cells in blood
CD20	Four transmembrane domain surface molecule	B-cell Ca^{2+} channel subunit; indirectly interacts with LYN, FYN, LCK	CVID5: low IgG, normal IgM, variable IgA *20–30% reduction in B-cell numbers*
CD21	Complement control protein	iC3b, C3dg, C3d, CD19, CD81, Leu 13, CD23	CVID7: Low IgG, reduced IgA, low normal IgM. *Diminished T cell–dependent immune responses, decreased germinal center formation, reductions in affinity maturation*
CD24	Glycosyl-phosphatidylinositol (GPI)–linked sialoglycoprotein	Ligand for P-selectin (CD62P)	A57V polymorphism associated with increased risk of multiple sclerosis. *Deletion in mice leads to reductions in late preB-cell and immature B-cell populations*
CD34	Type I transmembrane glycoprotein	Ligand for L-selectin (CD62L) and E-selectin (CD62E)	*Not expressed in murine B-cell progenitors*
CD38	Type II transmembrane glycoprotein adenosine diphosphate (ADP)-ribosyl cyclase, cyclic ADP-ribose hydroxylase	ADP ribosylates proteins	*Diminished T cell–dependent immune responses, augmented responses to T cell–independent type 2 polysaccharide antigens*

interact and thereby provides a link between innate and adaptive immune responses (Chapter 3). CD19-BCR interactions permit the cell to reduce the number of antigen receptors that need to be stimulated to activate the cell. Coactivation also reduces the threshold required for B-cell proliferation in response to a given antigen.

The cytoplasmic domain of CD19 contains nine conserved tyrosine residues that, when phosphorylated, allow CD19 to associate with PI-3 kinase and the tyrosine kinase VAV. Patients deficient in CD19 have normal numbers of $CD20^{+}$ B cells in blood but have panhypogammaglobulinemia and are susceptible to sinopulmonary infections (Chapter 34).

CD20 contains four transmembrane domains and cytoplasmic *C*- and *N*-termini. It is a member of the CD20/FcεRIβ superfamily of leukocyte surface antigens. Differential phosphorylation yields three forms of CD20 (33, 35, and 37 kilodaltons [kDa]). Activated B cells have increased fractions of the 35- and 37-kDa forms of the antigen. CD20 appears to function as a B-cell Ca^{2+} channel subunit that regulates cell cycle progression. It can interact directly with major histocompatibility complex (MHC) class I and II molecules, as well as members of another family of four transmembrane domain proteins, known as the TM4SF (*e.g.*, CD43, CD81, and CD82). It also appears to interact indirectly with LYN, FYN, and LCK. A patient born of consanguineous parents and with a CD20 deficiency present with normal numbers of B cells, reduced circulating memory B cells, low IgG with normal IgM and IgA, and reduced somatic hypermutation (SHM). Rituximab, a common monoclonal biological approved for medical use in 1997, is directed against CD20. It is commonly used to treat certain autoimmune diseases and lymphoid cancers (Chapter 89).

CD21 (complement receptor 2 [CR2]) is a cell-surface protein that contains a small cytoplasmic domain and an extracellular domain consisting of a series of short consensus repeats termed *complement control protein* (CCP) domains. These extracellular domains can bind three different products of complement C3 cleavage, iC3b, C3dg, and C3d. When binding these products, CD21 acts as the ligand-binding subunit for the CD19–CD21–CD81 complex, tying the innate immune system to the adaptive immune response. Mice that lack CD21 exhibit diminished T-dependent B-cell responses. However, serum IgM and IgG are in the normal range. A patient lacking CD21 presents with low IgG, low normal IgM, normal IgA, and normal responses to protein vaccination, but impaired responses to polysaccharide vaccines.

CD24 is a glycosyl-phosphatidylinositol (GPI)-linked sialoprotein that serves as a ligand for P-selectin (CD62P). It is expressed on progenitor, immature, and mature B cells. Its expression decreases in activated B cells and is lost entirely in plasma cells. Monoclonal antibodies (mAbs) against CD24 inhibit human B-cell differentiation into plasma cells. In mice, CD24 is also known as heat-stable antigen (HSA). Mice made deficient in CD24 show a leaky block in B-cell development with a reduction in late preB-cell and immature B-cell populations. However, peripheral B-cell numbers are normal, and no impairment of immune function has been demonstrated.

CD38 is a bifunctional enzyme that can synthesize cyclic adenosine diphosphate–ribose (cADPR) from nicotinamide adenine dinucleotide (NAD$^+$) and also hydrolyze cADPR to ADP-ribose. It is presumed that the enzyme exists to ADP-ribosylate target molecules. CD38 is expressed on preB cells, activated B cells, and early plasma cells, but not on immature or mature B cells or on mature plasma cells. Antibodies to CD38 can inhibit B lymphopoiesis, induce B-cell proliferation, and protect B cells from apoptosis. CD38 knock-out mice exhibit marked deficiencies in antibody responses to T cell–dependent protein antigens and augmented antibody responses to T-cell–independent type 2 polysaccharide antigens.

Transcription Factors Controlling B-Cell Differentiation

Ultimately, B-cell development is a function of differential gene expression. Deficiencies in the function of the transcription factors that regulate lymphoid-specific gene expression can thus be expected to result in abnormal B-cell development (Fig. 7.3; Table 7.2).[11]

PU.1 belongs to the ETS family of loop–helix–loop (winged helix) transcription factors, which bind purine-rich DNA sequences. In B cells, PU.1 regulates a number of critical genes, including CD79a (Igα), J chain, μ chain, κ chain, λ chain, RAG1, and terminal deoxynucleotidyl transferase (TdT), the enzyme responsible for *N* addition (Chapter 4). ETS family members are relatively weak transcriptional activators and typically require the presence of other factors to activate or repress their target genes. PU.1 is no exception. It cooperates with PIP (LSIRF, IRF4), c-JUN, and c-FOS. PU.1-deficient mice demonstrate defective generation of monocyte, granulocyte, and lymphocyte progenitors, indicating a role in the generation of MPPs as well as LMPPs. PIP-deficient mice lack GCs in peripheral lymphoid organs and exhibit defects in B-cell activation.

Ikaros and Aiolos belong to the same zinc finger transcription factor family. Although both are expressed during lymphoid development, Ikaros is expressed in stem cells and mature lymphocytes, whereas Aiolos is only expressed after commitment to the B-cell lineage. Ikaros transcripts are subject to alternate splicing. It can generate several isoforms, each of which differs in its DNA-binding pattern, tendency to dimerize, and nuclear localization. Among the genes bound by Ikaros are TdT, λ14.1 (λ5), VpreB, and LCK. Ikaros-deficient mice lack B cells. Aiolos-deficient mice exhibit elevated levels of IgG and IgE. As they age, these mice tend to develop autoantibodies and B-cell lymphomas.

The *E2A* locus encodes two basic helix–loop–helix transcription factors that represent two alternately spliced products, E12 and E47. Targets for E2A include RAG-1 and TdT. Although the functions of E12 and E47 overlap, E47 appears to play the greater role in driving TdT and RAG-1, whereas E12 is a better activator of EBF and PAX5 and thus helps commit developing cells to the B-cell lineage. In mice, disruptions in the *E2A* gene are marked by an arrest of B-cell development prior to the first transcription of *RAG-1*.

ID-1 has a helix–loop–helix domain, but lacks a DNA-binding domain. Thus it can function as a dominant negative factor, inhibiting the function of helix–loop–helix transcription factors, such as E2A. ID-1 is expressed only in proB cells. *ID-1* transgenic mice have a phenotype similar to *E2A* knock-out mice, suggesting that Id-1 can regulate E2A function.

EBF, or early B-cell factor, is a helix–loop–helix–like transcription factor. It is expressed at all stages of differentiation except plasma cells. In mice, it has been shown to be critical in the progression of B cells past the early proB-cell stage. The developmental block in B-cell differentiation is similar to that seen in E2A mutants, suggesting that these transcription factors act cooperatively and regulate a common set of genes.

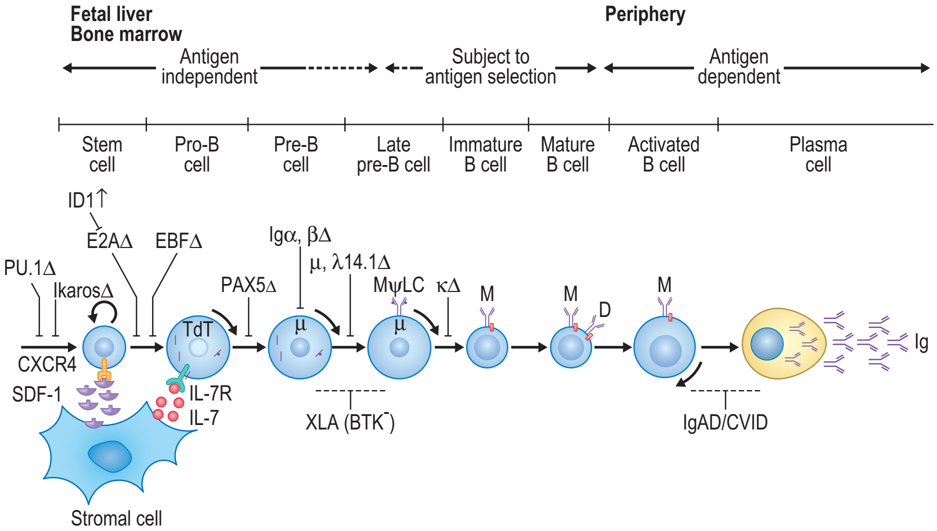


FIG 7.3 Genes Involved in Early B-Cell Development. The stage of development at which abnormal function of selected set of transcription factors, cytokines, chemokines, and signal transduction elements can influence B-cell development is illustrated. A Greek delta (Δ) or a dash (–) indicates a loss of function of the gene in question. An upward arrow (↑) indicates an increase in the function of the gene in question.

TABLE 7.2 Nuclear and Cytoplasmic Factors Active in Early B-Cell Development

Gene	Class or Alternative Name	Associated or Targeted Genes or Molecules	B Cell Developmental Phenotype in Humans or *Mice** Associated With Disrupted Function of the Indicated Gene
Transcription Factors			
PU.1	Loop–helix–loop (winged helix)	CD79a (Igα), μ H chain	Arrest prior to the proB-cell stage*
Ikaros	Zinc finger	RAG1, TdT, IL2R, VpreB, LCK	CVID13: progressive loss of B cells and serum immunoglobulins Arrest prior to the proB-cell stage*
Aiolos	Zinc finger	RAG1, TdT, IL2R	Aging mice develop symptoms of systemic lupus erythematosus*
E2A	Basic helix–loop–helix (BHLH)	RAG1, IgH, Igκ, TdT, EBF, PAX5	AGM8: Agammaglobulinemia, reduced numbers of $CD19^+$ B cells that lacked B-cell receptors (BCRs) Arrest prior to the proB-cell stage*
EBF	EBF/Olf helix–loop–helix (HLH) –like	CD79a (Igα), λ14.1, VpreB, PAX5	Arrest prior to the proB-cell stage*
PAX5	Paired-domain	CD19, λ14.1, VpreB, BLK kinase, J chain, V_H promoters, Vκ promoters	Susceptibility to B-cell acute lymphoblastic leukemia 3 Arrest at proB-cell stage*
The Recombinase Complex			
RAG1, RAG2	Recombinase	Recombination signal sequences of immunoglobulin gene segments	Autosomal recessive severe combined immunodeficiency (SCID) Arrest at proB-cell stage*
TdT	Nontemplated DNA polymerase	Coding ends of rearranging immunoglobulin gene segments	Absence of N nucleotides, diminished production of pathogenic anti-DNA autoantibodies, loss of heterosubtypic immunity against influenza virus*
DNA-PK	DNA repair complex	Multimeric complex consisting of DNA-PKcs, Ku70, Ku80, which repairs double-stranded DNA breaks	SCID Arrest at proB-cell stage, original mouse SCID mutation identified as a loss-of-function mutation in DNA-PKcs*
Protein Tyrosine Kinases			
FLK2/FLT3	Class III receptor tyrosine kinase	GRB2, SHC	Activating mutations contribute to acute myeloid leukemia Selective deficiency of primitive B-cell progenitors*
BLNK	SH2 adaptor protein	SYK, GRB2, VAV, NCK, Phospholipase Cγ (PLCγ)	AGM4: Normal numbers of proB cells, absent preB and B cells. Arrest at proB-cell stage*
BTK	BTK/TEC protein tyrosine kinase	Phospholipase Cγ (PLCγ), SAB	XLA: X-linked agammaglobulinemia - arrest at preB cell stage Xid: Impaired responses to T-cell–independent antigens*

*Indicates B cell development phenotype in mice.

PAX5 is a paired-box, or domain, transcription factor that, among the progeny of HSCs, is expressed exclusively in cells of the B-cell lineage. PAX5 has both a positive and a negative effect on B-cell differentiation. In mice, B-cell precursors require Pax5 to progress beyond the proB-cell stage. The presence of Pax5 also prevents early B-lineage progenitors from transiting into other hematopoietic pathways. Downregulation of PAX5 allows upregulation of BLIMP1 and plasma-cell differentiation. PAX5 downregulation in plasma cells permits expression of genes typically expressed in macrophages and neutrophils.

MicroRNAs and B-Cell Development

MicroRNAs (miRNAs) are a class of small, noncoding RNAs that downregulate target genes at a posttranscriptional level.[12] These RNAs are derived from longer transcripts by the sequential action of RNA polymerase II, the nuclear nuclease Drosha, and the cytosolic nuclease Dicer. Mature miRNAs are incorporated into the multiprotein RNA-induced silencing complex (RISC), which represses target messenger RNAs (mRNAs) by either inducing mRNA cleavage or mRNA degradation or by blocking mRNA translation. Critical miRNAs include miR-150, miR-155, and miR-17-92. Several of these miRNAs play a role in both early and late B-cell development. Abnormal function of these miRNAs can contribute to oncogenesis and immune dysfunction.

Modulation of B-Cell Development by Chemokines, Cytokines, and Hormones

Stromal cells provide the microenvironment for B-cell development and differentiation (see Fig. 7.3). For example, the chemokine CXCL12, also known as preB-cell growth-stimulating factor and as stromal cell-derived factor-1 (PBSF/SCF-1), promotes proB-cell proliferation. Mice with a targeted disruption of this gene exhibit impaired B-lymphopoiesis in fetal liver and bone marrow and fail to undergo bone marrow myelopoiesis.[13] The mechanism by which CXCL12 regulates early B-cell development remains unclear.

Although in mice interleukin-7 (IL-7) plays an essential role in B-lineage differentiation, in humans, it has a minimal proliferative effect on human B-cell progenitors. Nevertheless, IL-7 enhances CD19 expression, which plays an important role in BCR signal transduction (Chapter 4). IL-7 treatment of human proB cells also leads to a reduction in the expression of RAG-1, RAG-2, and TdT. Thus IL-7 can modulate the process of Ig gene segment rearrangement in human.

Interferons-α and -β (IFN-α/β) are potent inhibitors of IL-7-induced growth of B-lineage cells in mice.[14] The inhibition is mediated by cell death (apoptosis). One potential source of IFN-α/β is bone marrow macrophages. Another macrophage-derived cytokine, IL-1, can also act as a dose-dependent positive or negative modulator of B lymphopoiesis.

Systemic hormones also regulate lymphopoiesis.[15] A role for sex steroids is suggested by the reduction in preB cells during pregnancy. Estradiol can also alter later stages of B-cell development, promoting expansion of the marginal zone (MZ) compartment. Prolactin appears to enhance production of both MZ and follicular B cells. Mice with a LOF mutation in the *Pit-1* transcription factor gene do not produce growth hormone, prolactin, or thyroid-stimulating hormone. These dwarf mice exhibit a defect in B-cell development that is correctable by the thyroid hormone thyroxine.[16]

KEY CONCEPTS

B-Cell Development in the Periphery

- T cell–independent activation of naïve B cells results in terminal differentiation into short-lived plasma cells.
- T cell–dependent activation of B cells:
  - Induces germinal center formation, permitting somatic hypermutation and class-switch recombination (CSR)
  - Results in differentiation into high-affinity memory B cells (reactive memory) and plasma cells secreting high-affinity antibodies (protective memory)
  - Generates long-term humoral immune protection
- The longevity of plasma cells is supported by highly specialized survival niches in bone marrow.
- T-follicular helper (Tfh) cells control late B-cell differentiation by cell-bound ligands and secreted cytokines.
- Activated B cells control T-cell development by presentation of antigen and costimulation.

B-CELL DEVELOPMENT IN THE PERIPHERY

The life span of mature B cells expressing surface IgM and IgD appears entirely dependent on antigen selection. After leaving bone marrow, unstimulated cells live for only a few days. Deletion of the transmembrane/intracellular domains of the BCR leads to loss of mature B cells, which indicates that signaling through the BCR is essential for their survival. As originally postulated by Burnet's "clonal selection" theory, B cells are rescued from apoptosis by their response to a cognate antigen. The reaction to antigen leads to activation, which can then be followed by diversification.

The nature of the activation process is critical. T cell–independent stimulation of B cells induces differentiation into short-lived plasma cells with limited class switching. T-dependent stimulation adds additional layers of diversification, including SHM of the variable domains, which permits affinity maturation, class switching to the entire array of classes available (Chapter 4), and differentiation into the long-lived memory B-cell pool or into the long-lived plasma-cell population.

BAFF and APRIL Can Play Key Roles in the Development of Mature B Cells

B cells leave bone marrow while still undergoing initial maturation, demonstrating progressively higher levels of IgD expression with a commensurate lowering of IgM. The splenic environment plays a key role in this maturation process. Immigrant splenic maturing B cells pass through two transitional stages, known as transitional stages 1 (T1) and 2 (T2). Only a minority of these cells successfully make the transition, as this differentiation step is a crucial checkpoint for controlling self-reactivity. Passage through this checkpoint requires the interaction of soluble B cell–activating factor (BAFF) of the tumor necrosis factor (TNF) family, with its receptor, BAFF-R, which is expressed primarily on B cells.[17] Death signals triggered through interaction of the BCR with self antigen can be counterbalanced by stimulation of BAFF-R, which enhances expression of survival factors, such as Bcl-2, and at the same time downregulates proapoptotic factors.

BAFF and a second TNF family member APRIL (a proliferation-inducing ligand) are essential factors for B-cell development and also for their long-term maintenance.[18] With the development of plasma cells, BAFF-R is downregulated while the receptors transmembrane activator and calcium-modulator and cyclophilin ligand (CAML) interactor (TACI) and B-cell maturation antigen (BCMA) are upregulated. In contrast to BAFF-R, these members of the TNF-R family can bind both BAFF and APRIL. APRIL can induce isotype switching in naïve human B cells. More importantly, it is a crucial survival factor supporting the longevity of plasma cells.

B CELLS AND THE RESPONSE TO ANTIGEN

T Cell–Independent Antigens

Unlike T cells, which require presentation of antigen by other cells, B cells can respond directly to an antigen as long as antigen is able to cross-link the BCR. Such antigens, especially those that by nature cannot be recognized by T cells (*e.g.*, DNA or polysaccharides), can induce a B-cell response independent of T-cell help. Depending on the cytokine milieu, B cells may even class switch (Chapter 4; and see below), although the range of available classes appears to be restricted. B cells that are activated by antigen alone do not take part in a germinal center (GC) reaction (see below).

T Cell–Dependent Antigens

Activated B cells express both MHC class I and class II molecules on their cell surface (Chapter 5). They can thus present both intracellular and extracellular antigens to CD4 T-helper (Th) and CD8 T cytotoxic lymphocytes (Chapter 6). Their role as antigen-presenting cells (APCs) is enhanced when they present peptides from the same antigen they have taken up with their antibodies. Cognate recognition of the same antigen by both a B cell and a T cell permits each of these cells to reciprocally activate the other.

T cell–activated B cells express the costimulatory molecules CD80 and CD86. These cell-surface molecules are required for activation of T cells via CD28, and for inactivation by CD152 (cytotoxic T-lymphocyte antigen-4 [CTLA-4]). Since B cells do not express IL-12, they do not induce expression of IFN-γ in the activated T cells but rather favor the differentiation of activated T cells into IL-4, -5, -10, and - 13 expressing Th2 cells and IL-21 secreting T-follicular helper (Tfh) cells.[19] These cytokines can support the CD40-induced expansion of memory B cells (IL-4), CD40-induced class-switch recombination (CSR) to IgG4 or IgE (IL-4), and differentiation of antigen-activated B cells into high-affinity plasma cells (IL-21).

ORGANIZATION OF PERIPHERAL LYMPHOID TISSUES

B lymphocytes enter the secondary lymphoid organs through defined ways. Each organ exhibits a preferred route of entry.[20] For example, most lymphocytes enter the spleen through the bloodstream, whereas lymphocytes enter lymph nodes and the Peyer patches through high endothelial venules. DCs, macrophages,

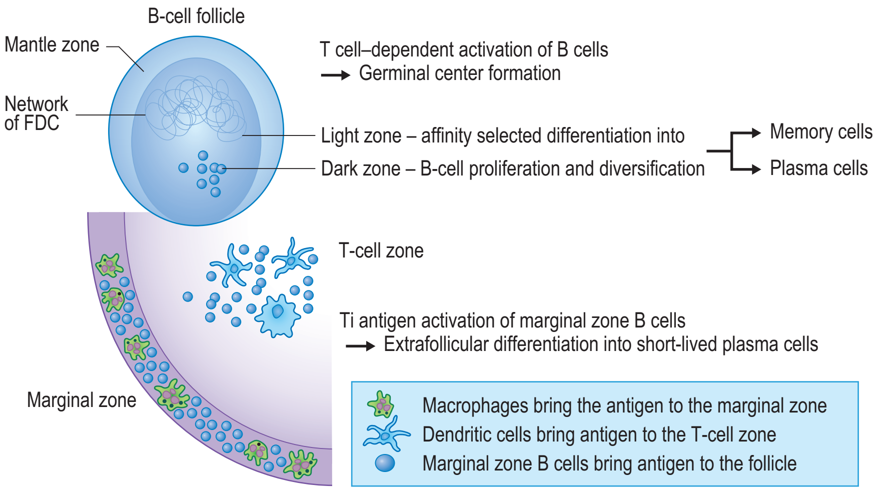

B-cell follicle
Mantle zone
Network of FDC
T cell–dependent activation of B cells
→ Germinal center formation
Light zone – affinity selected differentiation into
Dark zone – B-cell proliferation and diversification
Memory cells
Plasma cells
T-cell zone
Ti antigen activation of marginal zone B cells
→ Extrafollicular differentiation into short-lived plasma cells
Marginal zone
Macrophages bring the antigen to the marginal zone
Dendritic cells bring antigen to the T-cell zone
Marginal zone B cells bring antigen to the follicle

FIG 7.4 **T-Cell and B-Cell Compartments in the Murine Spleen.** T-cell compartments surround the central arterioles (periarticular lymphocyte sheath [PALS]). B cells are found in the adjacent follicles, where they are embedded in a network of follicular dendritic cells (FDCs). Marginal zone (MZ) B cells are located outside of the marginal sinuses, which mark the border of the white pulp and the red pulp.

and other highly specialized cells transport antigens from peripheral sites of entry into the secondary lymphoid organs. Within these organs, circulating lymphocytes survey available antigens (Chapter 6).

During ontogeny, the primary and secondary lymphoid organs are built up in an organized way.[21] This compartmentalization of the immune structures is essential for an efficient and controlled immune response. The establishment of the different immune compartments involves multiple factors. Among these, cytokines of the tumor necrosis factor superfamily (TNFSF), such as TNFα, lymphotoxin α (LTα), LTβ, and their receptors TNFR1 and LTβR, play important roles (Chapter 2). Chemokines also play an essential function in the organized development of the secondary lymphoid organs and the specific localization of immune cells (Chapter 10).[20]

In the secondary lymphoid organs, T cells and B cells are segregated into clearly defined areas, the T-cell zone and the B-cell follicle (Fig. 7.4). B cells are embedded into a network of highly specialized stromal cells, the follicular dendritic cells (FDCs). In contrast to other types of DCs, FDCs do not process antigen. Instead, FDCs have abundant complement receptors and Ig Fc receptors that allow accumulation of antigen in the form of immune complexes within the B follicle. Antigen presentation by FDCs is crucial for B-cell maintenance and for their activation and differentiation (see below).[22]

The Spleen

In the white pulp of the spleen, one can distinguish between a periarticular lymphatic sheath (PALS) of T cells and adjacent B-cell follicles.[23] In the murine spleen, the MZ is separated from the white pulp by the marginal sinus, the site of entry for lymphocytes, macrophages, and DCs into the splenic tissue. A specialized layer of metallophilic macrophages are thought to control the entry of antigen into the white pulp.[22] In the human spleen, one can distinguish an inner and an outer MZ. The latter is surrounded by a perifollicular area, where blood vessels terminate and thus facilitate the entrance of lymphocytes. Macrophages lining the ending of the capillaries help control the entry of antigen into the splenic tissue.

The splenic MZ provides a home for rapid B-cell responses. The ability of MZ B cells to rapidly respond to encapsulated bacteria by differentiating into antigen-specific plasma cells helps keep such infections under control. The MZ takes time to develop and is not present in young infants. It becomes fully populated with MZ B cells only after the age of 2 years. In the physiological absence of these B cells, a poor response to bloodborne infections is commonly observed.[24]

B-1 Cells

In addition to the MZ and conventional (B-2) subsets, differential expression of the cell-surface molecules IgD, CD5, CD11b/CD18, CD23, and CD45 in mice have allowed the identification of two additional peripheral B-cell subsets, B-1a, and B-1b.[25] “Conventional” B cells (B-2 cells) express high levels of IgM and IgD, whereas “CD5” B cells (B-1 cells) express minimal surface IgD. B-1 cells express little CD45 and virtually no CD23. They all express CD5 mRNA, although some display CD5 on their cell surface (B-1a) and some do not (B-1b).

B-1 cells seem to develop from distinct progenitors that represent a majority of B cells in fetal life. Accordingly, in mouse fetal liver, all B cells, and in fetal spleen, 40–60% of B cells are B-1 cells. Later in development, B-1 cells comprise <10% of the splenic IgM$^+$ B cells but are abundant in the peritoneal cavity.

Natural antibodies (NAbs) are IgM antibodies produced by B-1 cells. They are always found in serum and tend to have specificity for bacterial antigens as well as autoantigens. In general, they are polyreactive but of low affinity. Self-reactivity appears to play a major role in tissue homeostasis. NAbs also appear to provide an important and immediate defense against many infectious organisms.

The frequent presence of CD5 on chronic lymphocytic leukemia (CLL) B cells and their tendency to produce poly- and self-reactive antibodies led many to conclude that CLL was a leukemia of the human homologue of B-1 cells. When it became clear that CD5 was not a definitive marker for B-1 cells in humans, considerable effort was expended searching for this elusive subset. NAbs are present in human serum, but definition of the human B-1 subset remains controversial. Currently B cells that are CD2 $0^+CD27^+CD43^+CD70^-$ appear to be the best candidates.[4]

GERMINAL CENTERS

T cell–dependent activation of follicular B cells can induce the formation of a GC, which is the microenvironment where affinity maturation of the humoral immune response takes place. The interplay of hypermutation followed by antigen selection is the basis of affinity maturation. In the germinal center, B cells that express antibodies of high affinity are selected to develop into memory and long-living plasma cells.[26]

GCs develop only after T cell–dependent activation of B cells. Their full function is dependent on the interaction between CD40 expressed on B cells and CD40L (CD154) expressed on activated T cells. Patients with LOF mutations in CD40L have high serum levels of IgM and suffer from recurrent infections (hyper-IgM syndrome, Chapter 34).[27]

In a primary immune response, it takes about a week for the complex GC structure to develop. In the spleen, a few days after activation of antigen-specific B cells and T cells, small clusters of proliferating B cells are observed at the border of the T-cell zone and the primary B-cell follicle. The rapidly expanding B-cell clone seems to push the naïve B cells toward the edge of the primary follicle. The naïve B cells form a mantle zone around the newly developing GCs, and the primary follicle changes into a secondary follicle. Subsequently, the network of FDCs becomes filled with proliferating, antigen-activated B cells. An influx of antigen-activated Tfh cells is also observed. Tfh cells express the chemokine receptor CXCR5, which enable them to enter the B-cell follicle. During the GC reaction, expression of the chemokine CXCL13 by FDCs attracts both antigen-activated B and Tfh cells.[19]

In the second week after immunization, the GC matures into a classic structure that contains a dark zone and a light zone. At this stage of GC development, proliferation is restricted to the dark zone. Amidst the network of FDCs, the B cells differentiate into plasma cells and memory cells. In a fully developed GC, dividing cells are termed *centroblasts,* whereas differentiating cells within the FDC network are termed *centrocytes.*

In the dark zone, proliferating B cells activate a mechanism of SHM (Chapter 4).[28] This is a highly specific process that is targeted toward the gene segments that encode the antigen-binding domain of the antibody molecule. Hypermutation introduces single nucleotide changes into the rearranged variable genes of the Ig molecules. Thus within the dark zone, a clone of variants expressing antigen receptors with various affinities for the antigen is generated from a single B-cell progenitor. By chance, a few of these mutations result in a receptor with higher affinity for antigen. B cells expressing such receptors are favored for activation and proliferation, particularly late in an immune response when availability of antigen is limiting.

FDCs present antigen to B cells, but only those with high-affinity receptors are able to internalize the antigen via their BCR. Processing of the internalized antigen and presentation of peptides to Tfh cells are prerequisites for B-cell differentiation into memory and plasma cells.[19] Thus only the few B cells with high-affinity receptors get adequate help. IL-21 provided by the Tfh cells is crucial in this differentiation phase, and since it controls the fate of the GC B cell, it is essential for affinity maturation of the immune response.

B-CELL FUNCTIONS IN ADDITION TO ANTIBODY PRODUCTION

Mature B cells are not homogeneous. Functionally and developmentally distinct subsets exist. In the spleen, follicular B cells have a key role in the adaptive immune response, whereas MZ B cells are major players at the interface between the immediate innate immune response and the delayed adaptive response.[29]

B cells play an important function in the activation of T cells. Similar to DCs (Chapter 6), B cells can internalize antigen, process it, and the present antigen peptides to the T-cell receptor (TCR). In cancer, B cells can secrete tumor-associated autoantibodies and inflammatory cytokines and alter patterns of antigen presentation to T cells. Thus they can modulate T-cell and innate immune responses to the tumor. Through antigen–antibody complexes, B cells have the potential to influence immune cells that express Fc receptors, which include granulocytes and NK cells. In autoimmune diseases and also in response to inflammation, B cells can have an immune suppressive function. Regulatory B cells (Bregs) appear to exert their activity via the release of suppressive cytokines, such as IL-10, IL-35, and tumor growth factor-β (TGF-β).[30]

MOLECULAR MECHANISM OF SOMATIC HYPERMUTATION AND CLASS-SWITCH RECOMBINATION

Ig SHM and CSR are essential mechanisms for the generation of a high-affinity, adaptive humoral immune response. They allow the generation of effector plasma cells secreting high-affinity IgG, IgA, and IgE antibodies.

Somatic Hypermutation

Hypermutation occurs only during a narrow window in B-cell development. The mechanism is induced during B-cell proliferation within the microenvironment of the GC. With a high rate of about 10^{-3}/base pair/generation, single nucleotide exchanges are introduced in a stepwise manner into the rearranged V-region and its 3’ and 5’ flanking sequences. Mutations are randomly introduced, although there is a preference for transitions (cytidine → thymidine or adenosine → guanine) over transversions. Analysis of the pattern of somatic mutations has revealed that the sequence of the complementarity-determining regions (CDRs; Chapter 4), the loops that form the antigen-binding site, have been selected to form mutation hot spots.

Effective hypermutation requires the V-gene promoter and transcription-enhancer sequences. Indeed, the position of the V-gene promoter defines the start of the hypermutation domain, which spans about 2000 nucleotides. Any heterologous sequence that is introduced into the V gene segment locus will become a target of the hypermutation machinery. Thus SHM can sometimes play a role in lymphomas and leukemias, where oncogenes have been linked to Ig promoters and enhancers.

Class-Switch Recombination

Upon transition from the immature to the mature state and leaving bone marrow, the B cell starts to express IgD as well as IgM. Both IgM and IgD antibodies use the same V_HDJ_H-exon and

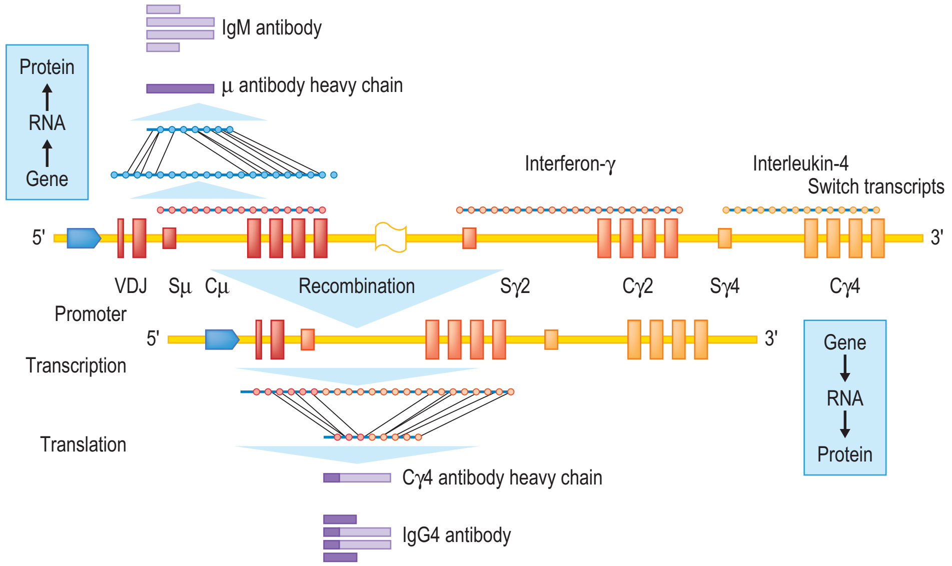


FIG 7.5 Antibody Class-Switch Recombination (CSR). Recombination between switch regions (Sµ and Sε) is preceded by transcription of these switch regions. Transcription is targeted by cytokines to distinct switch regions. IgM, Immunoglobulin M.

promoter (Fig. 7.5). The molecular basis of coexpression of IgM and IgD by the same B cell is attributed to differential termination of transcription and splicing of the primary transcripts. Although sequences have been identified that are required for the control of termination and splicing of the Cµ and Cδ transcripts, none of the proteins involved is known. The role of IgD remains unclear, but there are indications that IgM and IgD form different types of signal transduction structures on the cell surface. In mice, targeted inactivation of IgD has shown that it is not critical for B-cell activation and differentiation. However, IgD$^{-/-}$ B cells show a slightly reduced capability for affinity maturation.

Unlike IgD, the other antibody classes are not stably expressed together with IgM. B cells can switch from expression of their V_HDJ_H-exon with Cµ to expression of the same V_HDJ_H-exon with any of the downstream C_H genes (*e.g.,* $C\alpha_{1,2}$, $C\gamma_{1,2,3,4}$, or Cε) (Chapter 4). CSR, like SHM, is a hallmark of B-cell activation. It can be induced by T cell–independent signals (*e.g.*, lipopolysaccharide [LPS]) or by signals derived from T cells (*e.g.*, CD40L). CD40L-deficient humans (X-linked hyper-IgM syndrome [X-HIM]; Chapter 34) are severely impaired in the expression of Ig classes other than IgM.

TATA-less promoters located in front of the switch regions respond to signals from cytokines and B-cell activation–inducing ligands (*e.g.,* CD40L binding to CD40 on the B cell). These promoters initiate transcription just upstream of a small I-exon located between the promoter and the switch region. Transcription continues through the switch region itself and finishes after including the entire C_H gene sequence. Transcription is essential in targeting switch recombination to the transcribed switch region.

The choice of C_H gene targeted for switch recombination in a particular B cell appears to be dependent on external cytokine signals (Chapter 9). IFN-γ targets CSR to IgG2 in humans and IgG2a in mice, IL-4 to IgG4 and IgE in humans and IgG1 and IgE in mice, and TGF-β to IgA in both humans and mice. Other switch-targeting cytokines have been described, although our knowledge is far from complete. It is evident, however, that the cytokines central for the organization of cellular, humoral, and mucosal immunities, respectively, recruit exactly those classes of antibodies that provide the most useful functions for their respective branches of the immune system.

Both SHM and CSR Require Activation-Induced Cytidine Deaminase

Both CSR and SHM are dependent on an activation-induced cytidine deaminase (AID).[28] Mice deficient in this enzyme express only IgM antibodies without SHM, and patients with homozygous AID LOF mutations present with hyper-IgM syndrome (Chapter 34). Recently it was shown that DNA modifications are not limited to the rearranged V-region gene. As AID is active on single-stranded DNA (ssDNA), it may target all genes that are transcribed during the GC reaction. Subsequently, all DNA modifications, except the ones introduced into the rearranged V-region sequence, undergo repairs. As a result, retained SHMs are concentrated on the V-regions of the BCR alone.

Hypermutation is proposed to occur in two steps.[28] The mechanism is induced by AID-catalyzed deamination of deoxycytidine (C) to deoxyuridine (U). The mispairing of U and deoxyguanosine (G) is than processed by uracil DNA glycosylase and targeted by repair pathways. As a consequence, mutations at C–G pairs are observed. In the second step, mutations at A–T pairs are induced, probably during a mutagenic patch repair of U–G mismatches introduced by AID. A number of proteins, such as MSH2 and MHS6 (homologues 2 and 6 of the *Escherichia coli* MutS), polymerase η, or exonuclease-1, seem to be involved; however, the mechanism is not clear.

In CSR, AID is targeted to the switch regions located upstream (5') of each C_H gene. These switch regions are composed of 1–6 kilobase–long GC-rich repetitive sequence motifs. G–C pairs within these motifs are targeted by AID. Deamination of C and processing by uracil DNA glycosylase creates an abasic site that facilitates the introduction of double-stranded DNA (dsDNA) breaks. Joining and repair requires the presence of DNA-phosphokinases, Ku70, Ku80, and probably other members of the general double-strand repair mechanism (Chapter 4).

Both mechanisms, SHM and CSR, need to be tightly controlled, since the introduction of double-strand breaks into the DNA can not only pose a risk to the longevity of the B cell but also permit translocations involving and activating oncogenes.[31] For example, for Burkitt lymphoma cells and for plasma cell–derived myeloma cells, the translocation and ectopic expression of the *c-MYC* gene is an apparent consequence of abnormal SHM and CSR.

B-CELL MEMORY

One of the key features of the immune system is immunological memory for antigens encountered in the past. In humoral immune responses, there are two layers of memory, long-lived B memory cells and B effector cells (*i.e.*, plasma cells). The generation of these long-lived cells is dependent on antigen activation of B cells in the presence of T helper cells and thus the induction of germinal centers.

Memory B Cells

Although long-term protection of the organism is provided by both memory B and plasma cells, their contribution varies; some individuals are protected mainly by memory B cells, whereas others are primarily protected by plasma cells.[32] This can be of vital importance in special situations, such as transplantation, where activation of the immune system should be avoided. For example, treatment of transplant recipients with rituximab, an mAb specific for CD20, depletes memory B cells but has no effect on long-living plasma cells secreting transplant-specific antibodies.

After a lag-phase of 1–2 days, primary B-cell responses start with secreted low-affinity IgM antibodies. High-affinity antibodies of other Ig classes require the passage of time to be generated. In contrast, a second encounter with antigen induces a rapid development of memory B cells into new plasma cells, secreting antibodies of high quality (Fig. 7.6).

Plasma Cells

Protective humoral memory is provided by long-lived plasma cells.[33,34] These cells are generated in the secondary lymphoid organs and then migrate to bone marrow or to a site affected by inflammation. In bone marrow, plasma cells survive in highly specialized niches provided by the underlying reticular stromal cells. Here they can live on for long periods without further activation and proliferation, but their maintenance is dependent on survival factors, such as APRIL and IL-6. Eosinophils have been shown to be the main providers of these cytokines, and when they are depleted, plasma cells rapidly go into apoptosis.[35]

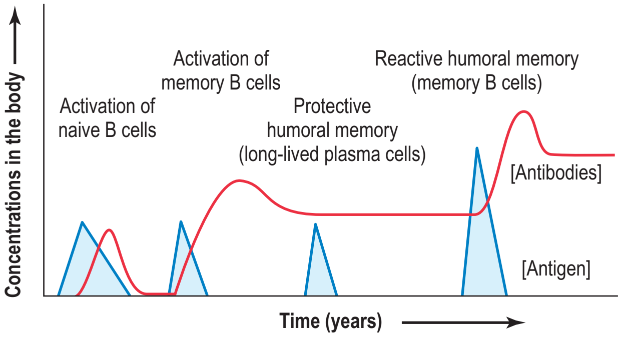


FIG 7.6 Active and Reactive B-Cell Memory. Memory B cells provide reactive memory, whereas long-lived plasma cells provide active protective memory. The relative concentrations of antibody and antigen over time are indicated.

By continuously secreting antibodies, long-lived plasma cells provide the individual with long-term humoral protection.

CLINICAL RELEVANCE

Abnormal B-Cell Development and Diseases of Immune Function

- Failure to generate B cells or a normal repertoire of antibodies leads to humoral immune deficiency, which is commonly marked by recurrent sinopulmonary infections.
- Failure to prevent the formation of antibodies with high avidity or high affinity to self antigens can lead to autoimmune diseases.
- The process of antibody repertoire diversification lends itself to the creation of mutations that can activate and modify oncogenes as well, leading to leukemia or lymphoma. Mechanisms include:
  - Recombinase activating gene (RAG)1/2-catalyzed juxtaposition of an oncogene to an immunoglobulin promoter or enhancer, activating the oncogene.
  - Activation-induced cytidine deaminase (AID)–induced DNA double-strand breaks and chromosome alterations.
  - AID-induced somatic hypermutation (SHM) of oncogene, altering its function.

ECTOPIC LYMPHOID TISSUE AND B-CELL DEVELOPMENT

In autoimmune diseases, in infection, and in tumors (cancer), ectopic lymphoid tissue can develop in the affected tissue or organ. Inflammatory cytokines and the presence of B cells can support the development of additional lymphoid tissue.[32]

The growth of ectopic lymphoid tissue in the rheumatoid synovium (Chapter 52) offers an excellent example of this disease-related phenomenon. In healthy individuals, the synovium is made up by a thin lining layer of synoviocytes. In contrast, in patients with rheumatoid arthritis, the diseased joint is highly infiltrated by varying numbers of T cells, B cells, plasma cells, macrophages, and DCs. In the majority of patients, these mononuclear cells are dispersed loosely throughout the synovium. However, well-organized, large lymphoid structures, which are similar in appearance to the lymphoid follicles seen in the secondary lymphoid organs, can develop in some patients.[32] At the center of these cell clusters, a network of FDCs are found. Antigen presented by FDCs appears to activate B cells, which induces proliferation. The central B cells are surrounded by a layer of T cells, which may support local B-cell differentiation. A central question concerns which antigens drive these immune responses and select B cells to differentiate into memory and plasma cells. The ectopic lymphoid tissue may function as additional lymphoid tissue. Equally likely, it may support a self-specific immune response. These questions remain topics of active investigation.

ON THE HORIZON

- Elucidation of the mechanisms used to control the antibody repertoire and shape B-cell epitope recognition offer the promise of being able to direct immunity toward production of broadly neutralizing or anti-tumorigenic antibodies, and away from pathogenic autoantibodies.
- Elucidation of the mechanisms that prevent the development of self-reactivity during affinity maturation could yield new insights into autoimmunity, as well as vaccination.
- A better understanding of the mechanisms that lead to long-lived plasma cells could lead to single, long-lasting vaccination strategies.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Matthias P, Rolink AG. Transcriptional networks in developing and mature B cells. Nat Rev Immunol 2005;5(6):497–508.
2. Nagasawa T. Microenvironmental niches in the bone marrow required for B-cell development. Nat Rev Immunol 2006;6(2):107–16.
3. Rosenspire AJ, Chen K. Anergic B Cells: Precarious On-Call Warriors at the Nexus of Autoimmunity and False-Flagged Pathogens. Front Immunol 2015;6:580.
4. Griffin DO, Holodick NE, Rothstein TL. Human B1 cells are CD3-: A reply to "A human equivalent of mouse B-1 cells?" and "The nature of circulating CD27+CD43+ B cells". J Exp Med 2011;208(13):2566–9.
5. Mei HE, Wirries I, Frolich D, et al. A unique population of IgG-expressing plasma cells lacking CD19 is enriched in human bone marrow. Blood 2015;125(11):1739–48.
6. Santos PM, Borghesi L. Molecular resolution of the B cell landscape. Curr Opin Immunol 2011;23(2):163–70.
7. Rajewsky K. Clonal selection and learning in the antibody system. Nature 1996;381(6585):751–8.
8. Melchers F. Checkpoints that control B cell development. J Clin Invest 2015;1–8.
9. van der Burg M, van Zelm MC, Driessen GJ, et al. Dissection of B-cell development to unravel defects in patients with a primary antibody deficiency. Adv Exp Med Biol 2011;697:183–96.
10. Haas KM, Tedder TF. Role of the CD19 and CD21/35 receptor complex in innate immunity, host defense and autoimmunity. Adv Exp Med Biol 2005;560:125–39.
11. Somasundaram R, Prasad MA, Ungerback J, et al. Transcription factor networks in B-cell differentiation link development to acute lymphoid leukemia. Blood 2015;126(2):144–52.
12. Aalaei-Andabili SH, Rezaei N. MicroRNAs (MiRs) Precisely Regulate Immune System Development and Function in Immunosenescence Process. Int Rev Immunol 2016;35(1):57–66.
13. Nagasawa T, Hirota S, Tachibana K, et al. Defects of B-cell lymphopoiesis and bone-marrow myelopoiesis in mice lacking the CXC chemokine PBSF/SDF-1. Nature 1996;382(6592):635–8.
14. Wang J, Lin Q, Langston H, et al. Resident bone marrow macrophages produce type 1 interferons that can selectively inhibit interleukin-7-driven growth of B lineage cells. Immunity 1995;3(4):475–84.
15. Grimaldi CM, Hill L, Xu X, et al. Hormonal modulation of B cell development and repertoire selection. Mol Immunol 2005;42(7):811–20.
16. Foster M, Montecino-Rodriguez E, Clark R, et al. Regulation of B and T cell development by anterior pituitary hormones. Cell Mol Life Sci 1998;54(10):1076–82.
17. Metzler G, Kolhatkar NS, Rawlings DJ. BCR and co-receptor crosstalk facilitate the positive selection of self-reactive transitional B cells. Curr Opin Immunol 2015;37:46–53.
18. Mackay F, Schneider P. Cracking the BAFF code. Nat Rev Immunol 2009;9(7):491–502.
19. Vinuesa CG, Linterman MA, Yu D, et al. Follicular Helper T Cells. Annu Rev Immunol 2016;34:335–68.
20. Schulz O, Hammerschmidt SI, Moschovakis GL, et al. Chemokines and Chemokine Receptors in Lymphoid Tissue Dynamics. Annu Rev Immunol 2016;34:203–42.
21. Fu YX, Chaplin DD. Development and maturation of secondary lymphoid tissues. Annu Rev Immunol 1999;17:399–433.
22. Cyster JG. B cell follicles and antigen encounters of the third kind. Nat Immunol 2010;11(11):989–96.
23. Mebius RE, Kraal G. Structure and function of the spleen. Nat Rev Immunol 2005;5(8):606–16.
24. Carsetti R, Rosado MM, Wardmann H. Peripheral development of B cells in mouse and man. Immunol Rev 2004;197:179–91.
25. Baumgarth N. B-1 Cell Heterogeneity and the Regulation of Natural and Antigen-Induced IgM Production. Front Immunol 2016;7:324.
26. Allen CD, Okada T, Cyster JG. Germinal-center organization and cellular dynamics. Immunity 2007;27(2):190–202.
27. Etzioni A, Ochs HD. The hyper IgM syndrome–an evolving story. Pediatr Res 2004;56(4):519–25.
28. Di Noia JM, Neuberger MS. Molecular mechanisms of antibody somatic hypermutation. Annu Rev Biochem 2007;76:1–22.
29. Cerutti A, Cols M, Puga I. Marginal zone B cells: virtues of innate-like antibody-producing lymphocytes. Nat Rev Immunol 2013;13(2):118–32.
30. Rosser EC, Mauri C. Regulatory B cells: origin, phenotype, and function. Immunity 2015;42(4):607–12.
31. Shaffer AL, Rosenwald A, Staudt LM. Lymphoid malignancies: the dark side of B-cell differentiation. Nat Rev Immunol 2002;2(12):920–32.
32. Mamani-Matsuda M, Cosma A, Weller S, et al. The human spleen is a major reservoir for long-lived vaccinia virus-specific memory B cells. Blood 2008;111(9):4653–9.
33. Nutt SL, Hodgkin PD, Tarlinton DM, et al. The generation of antibody-secreting plasma cells. Nat Rev Immunol 2015;15(3):160–71.
34. Manz RA, Hauser AE, Hiepe F, et al. Maintenance of serum antibody levels. Annu Rev Immunol 2005;23:367–86.
35. Chu VT, Berek C. The establishment of the plasma cell survival niche in the bone marrow. Immunol Rev 2013;251(1):177–88.
36. Manjarrez-Orduno N, Quach TD, Sanz I. B cells and immunological tolerance. J Invest Dermatol 2009;129(2):278–88.

MULTIPLE-CHOICE QUESTIONS

1. The first quality control checkpoint for immunoglobulin H chain function occurs at what stage of B-cell development?
   A. ProB cells
   B. PreB cells
   C. Immature B cells
   D. Transitional B cells
   E. B-1 cells

2. Somatic hypermutation (SHM) occurs at what stage of B-cell development?
   A. ProB cells
   B. PreB cells
   C. B-1 cells
   D. Marginal zone (MZ) B cells
   E. Germinal center (GC) B cells

3. The best cell surface marker for B cells is:
   A. CD10
   B. CD19
   C. CD20
   D. CD21
   E. CD24

4. Which one of the following is a transcription factor critical for B-cell development?
   A. TdT
   B. PAX5
   C. LCK
   D. BLNK
   E. BTK

8

T-Cell Development

Laurie E. Harrington

T cells and B cells (Chapter 7) are the two major components of the adaptive immune system. T cells are instrumental for protective immunity to numerous pathogens, yet these cells are also linked to the pathogenesis of multiple autoimmune disorders. One aspect unique to T-cell development as opposed to other hematopoietic cells, including B cells, is the requirement for the thymus (Chapter 2). Defective generation of the thymus manifests as reduction in T-cell numbers and resultant T-cell immune deficiency (Chapter 35). T-cell development begins with hematopoietic stem cells (HSCs) in the fetal liver and later in bone marrow. T-cell progenitors then travel to the thymus, where they undergo a highly intricate and defined series of differentiation steps that ultimately culminate in the mature peripheral T-cell population. The process of T-cell development is dependent on several signaling events and cellular interactions. Ultimately this process provides the host with the diverse repertoire of T cells needed for the recognition of a wide array of ancient and novel antigens (Chapters 4 and 6) and for the subsequent generation of effective adaptive immune responses toward those antigens.

KEY CONCEPTS

Early T-Cell Development

- Peripheral T cells are the progeny of hematopoietic stem cells (HSCs) from bone marrow or fetal liver.
- T-cell development occurs in and requires the thymus.
  - Patients without a thymus as a result of *FOXN1* mutations or DiGeorge syndrome lack circulating T cells.
- Early thymic progenitors (ETPs) are the first cells to seed the thymus.
- T-cell lineage commitment requires:
  - Notch signals
  - Upregulation of lineage-defining transcription factors
  - Progression through the double negative (DN) stage
- T-cell formation requires production of functional αβ or γδ T-cell receptor (TCR)

THYMUS: THE SITE OF T-CELL DEVELOPMENT

T-cell development is dependent on the presence of the thymus. In its absence, the generation of T cells is severely impaired.[1] Whether the absence of a thymus is the result of a germline mutation or of surgical removal at the time of birth, peripheral T cells do not develop in mice that do not have a thymus. For example, *nude* mice lack a thymus and lack T cells. However, when bone marrow from *nude* mice is transplanted into mice with an intact thymus, T-cell development and function are restored. Similarly, the genetic disorder DiGeorge syndrome further exemplifies the central role of the thymus in T-cell development in humans. Patients who develop DiGeorge syndrome (Chapter 35) as a result of deletions in chromosome 22q11 and patients with the rare FOXN1 deficiency are both unable to generate a complete thymus, and both types of patients exhibit the immunodeficiency associated with decreased peripheral T cells. In both human conditions, transplantation of thymic tissue is able to restore T-cell development and function.[2,3] Collectively, these observations demonstrate the fundamental need for the thymus in the generation of mature, peripheral T cells.

The anatomy of the thymus provides the unique microenvironment necessary to support T-cell development. The thymus is home not only to T-cell precursors but also to various stromal cell populations, including endothelial cells, fibroblasts, and importantly thymic epithelial cells (TECs) that are fundamental to the selection of functional T cells. TECs are not hematopoietic in origin and require the transcription factor FOXN1 for development. This helps explain why FOXN1 is central to thymus organogenesis and T-cell development and why transplantation of the thymus, but not bone marrow, will overcome the T-cell defect caused by FOXN1 mutations.

The thymus is divided into outer and inner regions, which are termed *cortex* and *medulla,* respectively (Chapter 2). These distinct areas foster specific aspects of T-cell commitment and differentiation. This can be partially attributed to the TECs resident to those regions: cortical thymic epithelial cells (cTECs) and medullary thymic epithelial cells (mTECs). The cTECs and mTECs can be distinguished by their location within the thymus, the expression of certain proteins, and distinct cytokines and chemokines that facilitate development of a diverse repertoire of mature T cells and deletion of autoreactive T cells.

Development of T cells in the thymus is a precisely orchestrated process involving interactions with specific cell types in specific locations. (The molecular and cellular cues that shape T-cell development and selection will be discussed in detail later in this chapter.) T-cell precursors from the circulation first enter the thymus via vessels at the corticomedullary junction. Guided by cues from stromal cells, such as cTECs, these immature thymocytes traffic into the cortex area, where commitment to the T-cell lineage occurs. As thymocyte differentiation progresses through the double-negative (DN) stage, the cells migrate further into the cortex, to the subcapsular region. Upon rearrangement of the αβ T-cell receptor (TCR), double-positive (DP) thymocytes can be found interspersed in the cortex area, where interactions with cTECs mediate positive selection. The DP thymocytes that survive positive selection begin to downmodulate either the CD4 or CD8 coreceptor. This is accompanied by the movement of these cells into the medulla of the thymus. Once in the medullary region, interactions with mTECs support negative selection of the maturing T cells, ensuring the deletion of autoreactive T cells from the repertoire. Surviving single-positive (SP) thymocytes exit the thymus as mature T cells.

LINEAGE COMMITMENT

T-cell lineage commitment is the result of a stepwise developmental program that starts with HSCs in bone marrow.[4,5] These HSCs are already endowed with the potential to differentiate into T cells when directly transplanted into the thymus. However, these HSCs do not fulfill this fate because they lack the ability to migrate to the thymus. Hence, further differentiation is necessary.

THE COMMON LYMPHOID PROGENITOR

Early in ontogeny, a bifurcation event occurs that gives rise to the common myeloid progenitor (CMP) and the common lymphoid progenitor (CLP). The CLP produces T cells, whereas the CMP represents the first stage at which the potential for the T-cell lineage is extinguished. At the fate bifurcation event, a series of epigenetic modifications are established in the CMP, and they silence the lymphoid differentiation program in the cells and further imprint erythroid, granulocyte, and myeloid potential.[6] In contrast, the CLP program enables its progeny to develop into T cells, B cells, natural killer (NK) cells, or dendritic cells (DCs). Again, however, T-cell commitment will only occur following entry to the thymus, whereas B-cell, NK-cell, or DC differentiation occurs in bone marrow and fetal liver without the need for migration.

Migration into the thymus is not a random event; hematopoietic precursor cells utilize specific homing molecules to facilitate this process. A fraction of the circulating CLP subset expresses the chemokine receptors CCR7 and CCR9, as well as the homing molecule P-selectin glycoprotein ligand-1 (PSGL-1). It is these molecules that enable the recruitment of circulating CLP cells to the thymus.[4]

The Early Thymic Progenitor

The first hematopoietic cell to seed the thymus is termed the early thymic progenitor (ETP) cell. ETP cells do not take up permanent residence in the thymus. Thus they must be continually generated from HSCs in bone marrow. Upon entry into the thymus, the B-cell potential of this cell population is extinguished. However, it still retains some level of multipotency in that it can differentiate into NK cells and some myeloid cell populations as well as T cells.

Double-Negative Thymocytes

In the thymus, T-cell development progresses from the ETP in a highly structured order that is marked by the expression of a particular set of surface molecules (Fig. 8.1).[7] The earliest stages in this process are characterized by the absence of CD3, CD4, and CD8 expression on the surface of maturing thymocytes. CD3 is composed of four distinct chains: CD3γ, CD3δ, and two CD3ε chains. Together, this heterotetramer associates the TCR and the ζ chain (zeta chain) to generate an activation signal in T lymphocytes. The TCR, ζ-chain, and CD3 molecules together constitute the TCR complex. CD4 is a coreceptor that assists TCR with binding to major histocompatibility complex (MHC) class II peptide complexes on the cell surface of antigen-presenting cells (APCs). CD8 is a coreceptor that assists the TCR with binding to MHC class I peptide complexes on the surface of most nucleated cells in the body. CD3$^-$CD4$^-$CD8$^-$ cells are termed *DN cells* in reference to the absence of CD4 and CD8.[5]

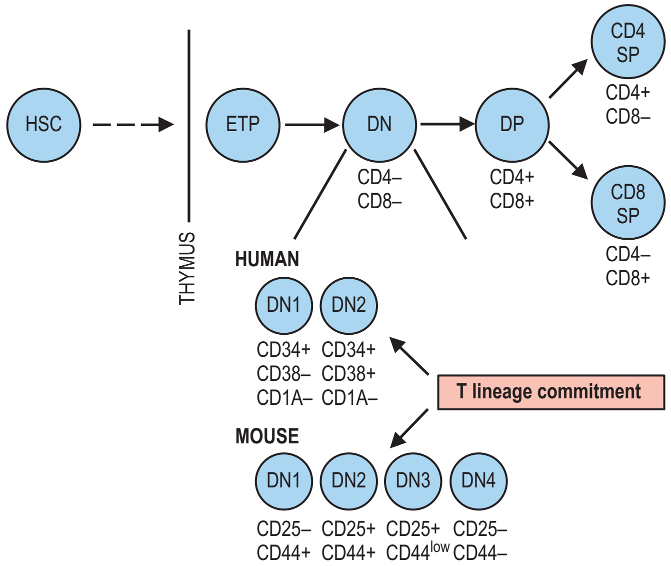


FIG 8.1 **Stages of Thymocyte Development.** T-cell development in the thymus is marked by the progression of cells through distinct stages of differentiation. The stages are defined by the presence of specific cell-surface molecules that are known to correspond with defined epigenetic and transcriptional changes. The double-negative (DN) thymocyte stage is slightly different between humans and mice, and a simplified depiction of this is shown.

KEY CONCEPTS

Double-Negative (DN), Double-Positive (DP), and Single-Positive (SP) Thymocytes

- Progressive stages of thymocyte development are identified by the expression profiles of CD4 and CD8 coreceptors.
- DN (CD4$^-$CD8$^-$) thymocytes give rise to DP (CD4$^+$CD8$^+$) and then SP (CD4$^+$CD8$^-$ or CD4$^-$CD8$^+$) thymocytes.
- CD4 and CD8 T-cell lineage determination is coordinated by the major histocompatibility complex (MHC) recognition of the T-cell receptor (TCR).
- MHC class II restriction leads to CD4 T cells, and MHC class I restriction leads to CD8 thymocytes, which then leave the thymus to become mature CD4 or CD8 T cells.

Cells of the DN phenotype are quite heterogeneous. In mice, differential surface expression of CD44 and CD25 can be used to divide these cells into four distinct stages. CD44 is a broadly distributed cell surface protein thought to mediate cell attachment to extracellular matrix (ECM) components or specific cell-surface ligands, including hyaluronate and chondroitin-4 and -6 sulfates. It thus plays a role in matrix adhesion, lymphocyte activation and lymph node homing. CD25 is the interleukin-2 (IL-2) receptor α (IL-2RA) chain. Homodimeric α chains (IL-2RA) create a low-affinity receptor for IL-2. CD25 can create a high-affinity receptor for IL-2 when associated with CD122, the IL-2 receptor β chain (IL-2RB), and CD132, the IL-2 receptor γ chain (IL-2RG). IL-2RG is also termed the common γ chain (γ_C), and it can contribute to IL-4, IL-7, IL-9, IL-15, and IL-21 receptor complexes in addition to IL-2. Activation of the IL-2 receptor pathway promotes T-cell proliferation.

In mice, CD44 DN1 cells express CD44 but not CD25 ($CD44^+CD25^-$). Thus they are primed for binding to ECM components or specific cell-surface ligands, but not for cell proliferation. DN2 cells express both CD44 and CD25 ($CD44^+CD25^+$) and are thus primed for location and proliferation. DN3 cells express low levels of CD44 and are positive for CD25 ($CD44^{lo}CD25^+$). They are thus more mobile while still being primed for proliferation. Finally, DN4 cells lack both CD44 and CD25 expression ($CD44^-CD25^-$).[8]

In humans, the DN precursor cell population can be segregated on the basis of expression of CD34, CD38, and CD1a into two stages, DN1 and DN2. The DN1 less mature precursors being ($CD34^+CD38^-CD1a^-$), whereas the DN2 more mature DN cells express CD34, CD38, and CD1a ($CD34^+CD38^+CD1a^+$).[8] CD34 is a cell-surface glycoprotein that can mediate the attachment of stem cells to the ECM or directly to stromal cells. CD38 is also known as cyclic adenosine diphosphate (cADP) ribose hydrolase. It is a glycoprotein that also functions in cell adhesion, as well as in signal transduction and calcium signaling. Unlike most MHC class I genes, which are located on the short arm of chromosome 6 within the MHC (Chapter 5), the CD1 genes are a nonpolymorphic cluster of MHC class I–like genes on human chromosome 1. CD1a molecules associate with the β_2 microglobulin and present antigen to T cells through their TCRs (Chapter 4). However, instead of peptides, CD1a molecules bind and present lipid and glycolipid antigens.

Commitment to the T-cell lineage is solidified in thymocytes of the DN phenotype. In mice, lineage commitment occurs during the DN2 to DN3 transition,[5] whereas T-cell lineage commitment is associated with the $CD34^+CD38^+CD1a^+$ population of thymocytes in humans.[8] Rearrangement of the TCR β, γ, and δ loci can be detected at this stage (Chapter 4). The pre-TCR is expressed by αβ T-cell precursors to test for TCRβ structural integrity; and, most importantly, the cells that pass this checkpoint stage no longer retain the ability to differentiate into other hematopoietic lineages.

Fate Commitment

After the cells pass through the DN stage, most of the T lineage–committed thymocytes will follow one of two fates. A minor fraction of these cells will express a functional γδ TCR on the surface and be exported into the periphery. But the bulk of the DN cells will begin to coexpress both CD4 and CD8 on the cell surface.[5,7] The $CD4^+CD8^+$ DP thymocytes constitute the majority of cells in the thymus. These are the precursors to the αβ T-cell lineage. It is at this stage of development that T-cell progenitors finalize the αβ TCR pairing through rearrangement of the TCR α loci and then initiate processes (positive and negative selection) to eliminate cells with nonproductive or autoreactive TCRs. Less than 5% of the DP thymocytes survive the selection processes, and the cells that endure downmodulate one of the two coreceptors from the cell surface, yielding either CD4 or CD8 SP thymocytes. After this stage of development, the SP cells are exported from the thymus and enter the periphery as mature CD4 or CD8 T cells.

TRANSCRIPTIONAL REGULATION OF COMMITMENT TO T-CELL LINEAGE

Notch

T-cell development is dependent on a number of molecular interactions and signaling events mediated by the cells resident to the thymus.[7,9] Among these, the Notch signaling pathway is integral to the generation of T cells.[5,10,11] The highly conserved Notch signaling pathway is known to dictate cell-fate decisions in numerous biological systems. Notch receptors are transmembrane molecules. In mammals, there are four members in the Notch family (Notch1–4). Of these, Notch1 expression plays a crucial role in T-cell progenitors.

Notch Ligand

Of the two types of Notch ligands, experimental evidence indicates that the δ-like (DL) family ligands function primarily during T-cell development. Upon interaction with its ligand, Notch receptors undergo a proteolytic cleavage event that releases the intracellular portion of the molecule. This portion then translocates into the nucleus to mediate transcription of target genes. Using elegant *in vitro* systems to dissect the factors that drive T-cell development, it has been determined that Notch signaling is critical for commitment to the T-cell lineage.[12] Ligation of Notch1 on hematopoietic progenitors by OP-9 cells expressing DL1, as well as by overexpression of active Notch1 in the progenitor cells, promotes T-cell development and suppresses B-cell development. Conversely, deletion of Notch1 expression in hematopoietic precursor cells prevents the development of T cells in the thymus and instead allows the emergence of immature B cells at this anatomical site.[13] Notch signaling alone does not commit a cell to the T-cell lineage, but it works with other transcription factors to imprint the T-cell fate. Importantly, some of the transcription factors are targets of active Notch in the cells.

T-Cell Factor 7

Notch target genes that encode molecules critical for T-cell commitment include the transcription factors T-cell factor-7 (TCF-7) and the TCF-7 splice variant TCF-1, which are both encoded by the *Tcf7* gene.[14,15] TCF1 is highly expressed in the ETP population in the thymus and deletion of *Tcf7* results in a reduction in these cells. Overexpression of TCF-1 in progenitor cells upregulates the expression of genes linked to the T-cell lineage and restores T-cell development even in the absence of Notch signaling. Hence, TCF-1 is necessary for initiating T-cell lineage development.

Enhancer Binding Protein GATA-3

Another transcription factor regulated by Notch signaling and necessary for T-cell development is GATA-binding protein 3 (GATA-3). GATA-3 is a zinc-finger transcription factor that is required during multiple stages of T-cell development, as well as in T-cell function. GATA-3 is expressed as early as the ETP stage and is critical for the development of this cell population. Loss of GATA-3 does not perturb the early stages of hematopoiesis, including the CLP stage of development, but does yield few, if any, ETPs, indicating a nonredundant function for GATA-3 in promoting T-cell development. As cells progress from the ETP stage to the DP stage, GATA-3 expression is maintained, and this transcription factor is known to have distinct functions in the various steps of T-cell development, including CD4 T-cell delineation.[16]

B-Cell Chronic Lymphocytic Lymphoma/Lymphoma 11B (Bcl11b)

Although Notch, TCF-1, and GATA-3 are all essential in promoting the T-cell lineage, an additional molecule, the transcriptional repressor Bcl11b, is important in suppressing alternative fates. Bcl11b is only expressed in T-cell progenitors, and it is not expressed in other hematopoietic cells. This molecule is upregulated as early as the DN2 stage of differentiation. It is necessary for T-lineage commitment, but it does not control the expression of many T lineage–specific genes.[17,18] Bcl11b is necessary to suppress the myeloid potential in the developing T-cell progenitors, and interestingly, in mice, deletion of Bcl11b from T-cell progenitors results in the cells adopting an alternative NK cell phenotype.[19]

T-CELL RECEPTOR REARRANGEMENT AND β SELECTION

One key step in T-cell lineage commitment is rearrangement of a functional TCR, be it an αβ or γδ TCR.[7] Rearrangement of the TCR gene loci within the developing thymocyte is a highly ordered sequence of recombination events mediated by the recombinase activating gene (RAG) enzymes (Chapter 4). RAG proteins are induced early in DN thymocytes and initiate recombination of the TCR γ, δ, and β genes. VDJ (or VJ) recombination of these loci must yield successful, in-frame rearrangements or the cell will stall at this stage and die.

γδ T cells

For cells of the γδ T-cell lineage, signaling via the newly rearranged γδ TCR complex directs the maturation and export of the γδ T cell into the periphery. This developmental checkpoint is critical for the selection of the γδ T-cell repertoire and, if unsuccessful, the thymocytes may either die or be redirected to the αβ lineage.

αβ T cells

For cells that will enter the αβ T-cell compartment, the developmental checkpoints are distinct from that of γδ T cells. Upon productive rearrangement of the TCR β gene, this protein will pair with preT α chain (surrogate α chain) to create the pre-TCR complex. The pre-TCR complex associates with the CD3 molecules and induces constitutive signaling within the cells.

β Selection and the Appearance of Double-Positive Thymocytes

Signaling by the pre-TCR complex in DN thymocyte facilitates the process of β selection, which instructs the cell to undergo a burst of proliferation and suppresses further β–chain rearrangement. After β selection, the thymocytes will begin to coexpress both CD4 and CD8 on the cell surface. These cells are termed DP cells. They subsequently initiate recombination of the TCR α gene loci. As a result of the organization of the TCR α loci, DP thymocytes have multiple opportunities to generate a successful TCR α rearrangement. Consequently most DP thymocytes express an αβ TCR. Only a fraction of these cells will progress into the periphery as CD4 or CD8 SP T cells.

KEY CONCEPTS

Positive Selection and Negative Selection

- Positive selection promotes the survival of double-positive (DP) thymocytes whose αβ T-cell receptor (TCR) can interact with MHC–self peptide complexes expressed by the thymic cortical epithelial cells.
- Negative selection deletes self-reactive T cells from the repertoire by inducing apoptosis in thymocytes expressing an αβ TCR with high affinity for self peptides, promoting central tolerance.

POSITIVE AND NEGATIVE SELECTION

Following the productive rearrangement and pairing of an αβ TCR, the DP T-cell precursor must pass through multiple checkpoints in the thymus before the cell can be released into the periphery. These processes, termed *positive selection* and *negative selection,* aim to eliminate T-cell clones expressing self-reactive, as well as unnecessary, nonfunctional αβ TCRs from the host.[20] This is critical, as rearrangement of both the α and β TCR genes occurs independent of MHC and antigen (Chapter 5). Therefore the antigen-specificity and the MHC restriction of any given αβ TCR pair are not predetermined. During both positive selection and negative selection, the ability of the αβ TCR pair to interact with peptide–MHC complexes is tested (Chapter 6), and the strength of the reaction determines the fate of the cell.

Positive Selection

Positive selection refers to the ability of the αβ TCR on the DP thymocyte to interact with self peptide–MHC complexes expressed by the cortical epithelial cells of the thymus.[20] If the αβ TCR is unable to bind the peptide–MHC complex, the DP thymocytes will not receive the appropriate signal and thus die by neglect in a matter of days. This is the fate of the majority of DP thymocytes for two main reasons: MHC alleles are highly diverse, and rearrangement of both α and β TCR genes is random. The combination of these two factors does not favor the generation of an αβ TCR pairing capable of binding the peptide–MHC complex with appropriate affinity.[21] If, however, the interaction is successful, the cell is rescued from programmed cell death. Positive selection has been demonstrated experimentally in mice that are genetically engineered to express a single αβ TCR with a known antigen-specificity and MHC restriction.[22] In mice that express the αβ TCR transgene and the correct MHC allele in the thymus, positive selection will proceed, and mature SP T cells will be exported into the periphery. In contrast, if the mice do not express the appropriate MHC allele, the αβ TCR transgenic DP thymocytes will fail positive selection and die by neglect in the thymus.

Importantly, the positive selecting MHC molecule, be it MHC class I or MHC class II, dictates expression of the corresponding CD8 or CD4 coreceptor that will be retained by the DP thymocyte as it matures (Fig. 8.2). In mice that lack MHC class I complexes, selection of CD4 SP cells proceeds normally, but CD8 SP cells are not generated.[23,24] Likewise, in mice deficient in MHC class II expression, CD8 SP cells develop and mature, but CD4 SP cells do not.[25,26] Notably, positive selection of DP thymocytes cells can be rescued in MHC-deficient hosts by restricting MHC expression to the thymic cortical epithelial cells, and this demonstrates the crucial role of MHC and the cortical epithelial cells of the thymus in regulating T-cell development.

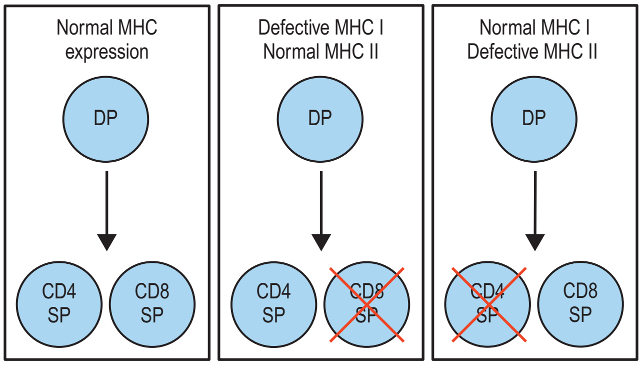


FIG 8.2 **Positive Selecting Major Histocompatibility Complex (MHC) Dictates CD4/CD8 Coreceptor Expression.** Positive selection in the thymus is determined by the ability of the αβ T-cell receptor (TCR) to bind to MHC. Upon selection, the immature T cell must coordinate expression of the CD4 or CD8 coreceptor with that MHC. If the thymocyte is positively selected by a class I MHC–peptide complex, the T cell will utilize the CD8 coreceptor. Conversely, if the double-positive (DP) thymocyte is positively selected by a class II MHC–peptide complex, the developing T cell will maintain expression of the CD4 coreceptor. SP, Single-positive.

Negative Selection

Positive selection ensures that the αβ TCR expressed by a DP thymocyte is capable of interacting with MHC, whereas negative selection is the means by which an immature T-cell clone encoding a self-reactive αβ TCR is eliminated from the host.[20] In a seminal experiment, negative selection was validated by using transgenic mice engineered to express an αβ TCR specific for the male antigen encoded by the Y chromosome (Fig. 8.3).[27] In female mice expressing the male-specific TCR, T-cell development proceeded normally. However, in male mice with the same TCR transgene, T-cell development was aborted at the DP stage, and no SP T cells emerged from the thymus. The different outcomes are the direct result of the presence or absence of self-antigen recognition, in this case the male antigen, by the αβ TCR in the thymus. The deletion of self-reactive T-cell clones mediated by negative selection is key to establishing *central tolerance.*

During negative selection, self antigens are displayed by the MHC on the surface of thymic epithelial cells and in the cortex and the medulla, or thymic-resident, bone marrow–derived DCs and macrophages. The self antigens presented in the thymus encompass ubiquitously expressed proteins, as well as tissue-restricted proteins, such as insulin. The expression of some tissue restricted antigens in the thymus is mediated by the autoimmune regulator (AIRE).[28] Humans who possess mutations in *AIRE* develop the autoimmune disease autoimmune polyendocrinopathy with candidiasis and ectodermal dysplasia (APOCED) (Chapter 50).[29] In mice, *Aire* deficiency also results in tissue-specific autoimmune disease.[30,31] Hence, the breadth of self antigens presented in the thymus during negative selection directly impacts the deletion of autoreactive αβ T-cell clones, thereby influencing autoimmunity.

Role of the Major Histocompatibility Complex in Negative Selection and Positive Selection

The processes of positive and negative selection are highly sensitive to the degree by which the TCR interacts with self peptide–MHC complexes. Weak interactions with peptide–MHC complexes are necessary to protect the developing T cells from death by neglect and promote the positive selection of the cells. However, an interaction with the peptide–MHC complex that is too strong induces negative selection, causing the cells to die via apoptosis. Nevertheless cells do escape this process, and self-reactive cells can thereby enter the periphery. Interestingly, the current concept is that regulatory T cells (Tregs) have a TCR that can bind with higher affinity to the selecting peptide–MHC complexes, inferring the self-reactivity of this particular subset of cells.

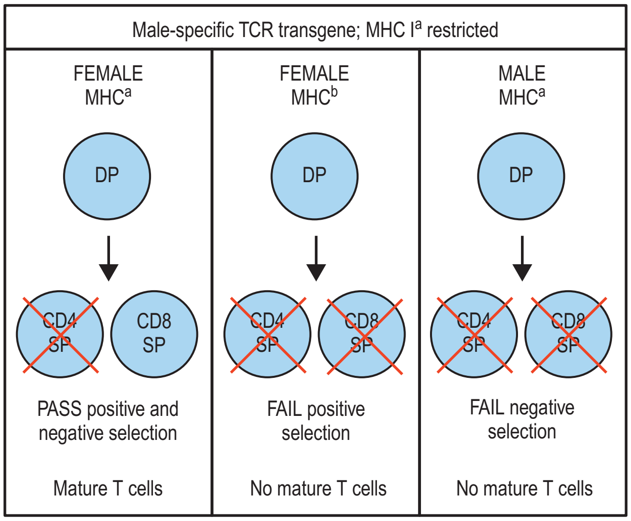


FIG 8.3 **Positive Selection and Negative Selection of Double-Positive (DP) Thymocytes Is Driven by Major Histocompatibility Complex (MHC) Restriction and Self Antigen Recognition.** The study of T-cell receptor (TCR) transgenic mice has yielded considerable insight into the processes of positive selection and negative selection. Positive selection will only occur if an MHC allele capable of binding the appropriate peptide and the transgenic TCR is expressed in the thymus. If the mice do not express an MHC molecule that will bind both the peptide and the TCR, the cells will die via apoptosis (middle panel). Negative selection will delete cells that possess a TCR with too high an affinity for self-antigen in the context of MHC (right panel). SP, Single-positive.

CD4 AND CD8 T-CELL DETERMINATION

Transition From Double-Positive to Single-Positive CD4 or CD8 Thymocytes

The final stage in T-cell development prior to exit from the thymus is transition of the DP thymocyte into the immature, SP CD4 or CD8 T-cell pool. The developing αβ T cell must coordinate expression of the CD4 or CD8 coreceptor with the corresponding MHC molecule, hence the CD4 versus CD8 decision is made during positive selection. Currently a kinetic model of TCR signaling is proposed to govern the CD4/CD8 determination. This model purports that CD4 T cells require a longer period of TCR signaling compared with CD8 T cells. When the DP thymocyte interacts with the MHC on thymic cortical epithelial cells, the CD8 molecule is transiently downmodulated. If the TCR–MHC interaction is sustained, this selects CD4 SP cells. However, if the TCR–MHC interaction is limited, the thymocyte will reexpress CD8 on the cell surface and repress CD4 expression.

Factors That Dictate CD4 Versus CD8 Commitment

Extensive research is being performed to understand the mechanisms and factors that dictate the CD4/CD8 decision. There are known transcription factors that facilitate this process and are necessary for the differentiation of functional CD4 and CD8 T cells.[7] Two transcription factors identified as being key to enforcing the CD4 and CD8 T-cell lineages are T-helper–inducing POZ-Kruppel factor (ThPOK) and Runt-related transcription factor 3 (Runx3), respectively. These molecules function to counterregulate each other. ThPOK promotes CD4 T-cell commitment and silences CD8 T-cell development, in part by repressing expression of Runx3.[16,32-34] Conversely, Runx3 is essential for CD8 T-cell generation and silences CD4 expression in these cells.[35] Importantly, these two transcription factors are necessary, but not sufficient, to induce CD4 and CD8 SP cells, and other factors, including GATA-3, are critical in this decision.[16]

Migration of Thymocytes Into the Periphery

Upon maturation, specific proteins facilitate the migration of CD4 and CD8 SP T cells from the thymus. One molecule responsible for certain aspects of the thymic egress process is the transcription factor Kruppel-like factor 2 (Klf2).[36] Klf2 is increased after positive and negative selection in SP cells and drives the expression of sphingosine 1 phosphate receptor 1 (S1P1) and L-selectin. Upregulation of S1P1 on the maturing T cells is necessary for migration from the thymus, as mice deficient in S1P1 have significantly reduced numbers of circulating peripheral T cells because the cells are unable to exit the thymus.[37] L-selectin is involved in the recruitment of the newly emerged T cells to lymph nodes. L-selectin interacts with various ligands, including GlyCAM-1 and CD34, expressed by the high endothelial venules in the lymph nodes to coordinate movement of the cells into the secondary lymphoid organs.[38] Together, these molecules, as well as others, deliver mature CD4 and CD8 T cells to the lymph nodes where these cells can function in adaptive immune responses.

ON THE HORIZON

- Deducing how the T-cell receptor (TCR) senses major histocompatibility complex (MHC) class I versus MHC class II and coordinates the expression of the appropriate CD8 or CD4 coreceptor, respectively
- Understanding additional molecular events that facilitate negative selection of tissue-specific antigens in the thymus
- Better tools and markers to study αβ T-cell development and the acquisition of innate-like functions prior to thymic export
- Developing strategies to combat thymic involution with age and increase T-cell development, especially following bone marrow transplantation

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Miller JF. Immunological function of the thymus. Lancet 1961;2:748–9.
2. Markert ML, Boeck A, Hale LP, et al. Transplantation of thymus tissue in complete DiGeorge syndrome. N Engl J Med 1999;341:1180–9.
3. Markert ML, Marques JG, Neven B, et al. First use of thymus transplantation therapy for FOXN1 deficiency (nude/SCID): a report of 2 cases. Blood 2011;117:688–96.
4. Yang Q, Jeremiah Bell J, Bhandoola A. T-cell lineage determination. Immunol Rev 2010;238:12–22.
5. Ciofani M, Zuniga-Pflucker JC. Determining γδ versus αβT cell development. Nat Rev Immunol 2010;10:657–63.
6. Cedar H, Bergman Y. Epigenetics of haematopoietic cell development. Nat Rev Immunol 2011;11:478–88.
7. Shah DK, Zuniga-Pflucker JC. An overview of the intrathymic intricacies of T cell development. J Immunol 2014;192:4017–23.
8. Taghon T, Rothenberg EV. Molecular mechanisms that control mouse and human TCR-αβ and TCR-γδ T cell development. Semin Immunopathol 2008;30:383–98.
9. Yui MA, Rothenberg EV. Developmental gene networks: a triathlon on the course to T cell identity. Nat Rev Immunol 2014;14:529–45.
10. Pui JC, Allman D, Xu L, et al. Notch1 expression in early lymphopoiesis influences B versus T lineage determination. Immunity 1999;11:299–308.
11. Yashiro-Ohtani Y, Ohtani T, Pear WS. Notch regulation of early thymocyte development. Semin Immunol 2010;22:261–9.
12. Holmes R, Zuniga-Pflucker JC. The OP9-DL1 system: generation of T-lymphocytes from embryonic or hematopoietic stem cells in vitro. Cold Spring Harb Protoc 2009;2009:pdb prot5156.
13. Radtke F, Wilson A, Stark G, et al. Deficient T cell fate specification in mice with an induced inactivation of Notch1. Immunity 1999;10:547–58.
14. Weber BN, Chi AW, Chavez A, et al. A critical role for TCF-1 in T-lineage specification and differentiation. Nature 2011;476:63–8.
15. Germar K, Dose M, Konstantinou T, et al. T-cell factor 1 is a gatekeeper for T-cell specification in response to Notch signaling. Proc Natl Acad Sci USA 2011;108:20060–5.
16. Wang L, Wildt KF, Zhu J, et al. Distinct functions for the transcription factors GATA-3 and ThPOK during intrathymic differentiation of CD4(+) T cells. Nat Immunol 2008;9:1122–30.
17. Wakabayashi Y, Watanabe H, Inoue J, et al. Bcl11b is required for differentiation and survival of αβ T lymphocytes. Nat Immunol 2003;4:533–9.
18. Li L, Leid M, Rothenberg EV. An early T cell lineage commitment checkpoint dependent on the transcription factor Bcl11b. Science 2010;329:89–93.
19. Li P, Burke S, Wang J, et al. Reprogramming of T cells to natural killer-like cells upon Bcl11b deletion. Science 2010;329:85–9.
20. Klein L, Kyewski B, Allen PM, et al. Positive and negative selection of the T cell repertoire: what thymocytes see (and don't see). Nat Rev Immunol 2014;14:377–91.
21. Merkenschlager M, Graf D, Lovatt M, et al. How many thymocytes audition for selection? J Exp Med 1997;186:1149–58.
22. Kisielow P, Teh HS, Bluthmann H, et al. Positive selection of antigen-specific T cells in thymus by restricting MHC molecules. Nature 1988;335:730–3.
23. Koller BH, Marrack P, Kappler JW, et al. Normal development of mice deficient in β_2M, MHC class I proteins, and CD8+ T cells. Science 1990;248:1227–30.
24. Muller D, Koller BH, Whitton JL, et al. LCMV-specific, class II-restricted cytotoxic T cells in β_2-microglobulin-deficient mice. Science 1992;255:1576–8.
25. Cosgrove D, Gray D, Dierich A, et al. Mice lacking MHC class II molecules. Cell 1991;66:1051–66.
26. Grusby MJ, Johnson RS, Papaioannou VE, et al. Depletion of CD4+ T cells in major histocompatibility complex class II-deficient mice. Science 1991;253:1417–20.
27. Kisielow P, Bluthmann H, Staerz UD, et al. Tolerance in T-cell-receptor transgenic mice involves deletion of nonmature $CD4^+8^+$ thymocytes. Nature 1988;333:742–6.
28. Mathis D, Benoist C. A decade of AIRE. Nat Rev Immunol 2007;7:645–50.
29. Villasenor J, Benoist C, Mathis D. AIRE and APECED: molecular insights into an autoimmune disease. Immunol Rev 2005;204:156–64.
30. Anderson MS, Venanzi ES, Klein L, et al. Projection of an immunological self shadow within the thymus by the AIRE protein. Science 2002;298:1395–401.

31. Anderson MS, Venanzi ES, Chen Z, et al. The cellular mechanism of Aire control of T cell tolerance. Immunity 2005;23:227–39.
32. He X, He X, Dave VP, et al. The zinc finger transcription factor Th-POK regulates CD4 versus CD8 T-cell lineage commitment. Nature 2005;433:826–33.
33. Egawa T, Littman DR. ThPOK acts late in specification of the helper T cell lineage and suppresses Runx-mediated commitment to the cytotoxic T cell lineage. Nat Immunol 2008;9:1131–9.
34. Muroi S, Naoe Y, Miyamoto C, et al. Cascading suppression of transcriptional silencers by ThPOK seals helper T cell fate. Nat Immunol 2008;9:1113–21.
35. Setoguchi R, Tachibana M, Naoe Y, et al. Repression of the transcription factor Th-POK by Runx complexes in cytotoxic T cell development. Science 2008;319:822–5.
36. Carlson CM, Endrizzi BT, Wu J, et al. Kruppel-like factor 2 regulates thymocyte and T-cell migration. Nature 2006;442:299–302.
37. Matloubian M, Lo CG, Cinamon G, et al. Lymphocyte egress from thymus and peripheral lymphoid organs is dependent on S1P receptor 1. Nature 2004;427:355–60.
38. Rosen SD. Ligands for L-selectin: homing, inflammation, and beyond. Annu Rev Immunol 2004;22:129–56.

MULTIPLE-CHOICE QUESTIONS

1. The process of positive selection ensures that the developing T cell can interact with the corresponding major histocompatibility complex (MHC) molecules with appropriate affinity. How does the process of negative selection impact the maturing T cell?
   - **A.** Negative selection refers to the inability of the rearranged T-cell receptor (TCR) to interact with the MHC.
   - **B.** Negative selection induces the deletion of maturing T cells that express the inappropriate CD4 or CD8 coreceptor.
   - **C.** Negative selection induces apoptosis in the maturing T cells that recognize self antigen with high affinity.
   - **D.** Negative selection induces proliferation in maturing T cells that recognize self antigen with high affinity.
   - **E.** Negative selection promotes expression of the CD4 or CD8 coreceptor in the developing T cells.

2. DiGeorge syndrome is a rare genetic disorder, which results in immunodeficiency in some patients. Why does this group of patients have defects in T-cell development?
   - **A.** These patients have a mutation in the gene *FOXN1* and do not develop a thymus.
   - **B.** These patients have a deletion on chromosome 22 and do not develop a thymus.
   - **C.** These patients have a mutation in the gene *AIRE* that prevents T-cell export from the thymus.
   - **D.** These patients have a mutation in the gene that encodes the sphingosine 1 phosphate receptor 1 (S1P1), which results in failure to export T cells from the thymus.
   - **E.** These patients have a mutation in the *RAG* gene that prevents rearrangement of a functional TCR.

3. During T-cell development in the thymus, early thymic progenitors progress through a sequential series of events that culminates in a mature T cell. How do these processes differ between humans and mice?
   - **A.** In humans, the double-negative (DN) subset of thymocytes comprises the DN1 and DN2 stages, whereas in mice the DN subset consists of the DN1, DN2, DN3, and DN4 stages.
   - **B.** In humans, T-cell lineage commitment occurs at the double-positive (DP) stage of development, whereas in mice lineage commitment occurs in the DN stage.
   - **C.** In mice, surface expression of CD1A is used to discriminate the DN stages, whereas in humans surface expression of CD44 and CD25 distinguishes cells in the DN stages.
   - **D.** In humans, the DN subsets of thymocytes comprises the DN1, DN2, DN3, and DN4 stages, whereas in mice the DN population consists of the DN1 and DN2 stages.
   - **E.** There are no known differences in T-cell development between humans and mice.

9

Cytokines and Cytokine Receptors

John J. O'Shea, Massimo Gadina, Richard M. Siegel

Cytokines play pivotal roles in controlling the development and function of a variety of immune and nonimmune cells. These include immune regulation, disease pathogenesis, and, increasingly, modulation and treatment of immune-mediated diseases. The term *cytokine* encompasses factors that are structurally or functionally unrelated. Included among cytokines are a number of different factors produced by lymphoid and nonlymphoid cells that mediate intercellular communication. The term *lymphokines* was originally used to denote products of lymphocytes,[1] whereas the term *interleukin* was introduced to emphasize the importance of these factors in communication between leukocytes.[2] Although the designation *interleukin* has remained in use, it is inaccurate in that some cases individual interleukins are made by cells other than leukocytes.

Because of the way they were discovered, the complex nomenclature and classification that have developed can be a barrier to understanding cytokines. Many were first identified by researchers in different disciplines, and they were named on the basis of their original observed functions, which may not necessarily reflect the full spectrum of an individual cytokine's actual biological functions.

Cohen et al.[3] coined the word *cytokine* to emphasize the point that these factors need not be made by one specific cell source. This was an important insight because many immunologically relevant cytokines are made by nonlymphoid cells. Cytokines are thus defined operationally as polypeptides secreted by leukocytes and other cells that act principally on hematopoietic cells, the effects of which include modulation of immune and inflammatory responses. However, there are clear exceptions to even this broad definition. Some definitions distinguish cytokines from hormones and growth factors, which act on nonhematopoietic cells.

Cytokines are typically characterized as factors made by more than one cell type that act locally, whereas hormones are secreted by specialized cells and act at a distance on a restricted set of target cells. Although many cytokines act locally in an autocrine or paracrine fashion, some do enter the bloodstream and can act in a typical endocrine fashion. Consequently, the boundary between cytokines and hormones is rather indistinct. In fact, classic hormones, such as growth hormone (GH), prolactin (PRL), and erythropoietin (EPO), and a more recently identified hormone, leptin, are all clearly cytokines, as evidenced by the structure of their receptors and their modes of signaling. Perhaps it is just simplest to accept that cell–cell communication and host defense went hand-in-hand during evolution, and so functional and structural similarities exist among families of molecules that act on the immune, hematopoietic, endocrine, and nervous systems.

KEY CONCEPTS

Cytokine Characteristics

- Cytokines have pleiotropic effects: they may have more than one receptor.
- Cytokines can be redundant—their receptors often share subunits.
- Cytokines can have specific and unique functions—their receptors typically have ligand-specific subunits as well.

CYTOKINE CLASSIFICATION

A major challenge in discussing cytokines is how best to classify them. One can legitimately group cytokines in different ways, but here, we classify cytokines on the basis of the type of receptor that they bind. Our scheme emphasizes the evolutionary relatedness of cytokines, growth factors, and hormones and highlights the similarities in signal transduction. The classification used is adapted from Vilcek[4] and includes the following receptors: the so-called type I (hematopoietin family) and type II (interferon family) cytokine receptors, tumor necrosis factor (TNF) family receptors, interleukin (IL)-1 receptor and the related Toll-like receptors (TLRs), IL-17 receptors, receptor tyrosine kinases, and the transforming growth factor-β (TGF-β) family receptor serine kinases (Table 9.1; Fig. 9.1). A sixth group, known as chemokines, form a separate family and bind seven transmembrane domain receptors (Chapter 10). This chapter reviews in detail only a selected set of cytokines with important immunological functions (Fig. 9.2).

TYPE I AND II CYTOKINE RECEPTORS (HEMATOPOIETIN FAMILY AND INTERFERON RECEPTORS)

Ligand and Receptor Structure

Cytokines (see Table 9.1) that bind the class of receptors, termed *type I* or *hematopoietic cytokine receptor superfamily,* include hormones, such as EPO, thrombopoietin (TPO), PRL, GH, and leptin; colony-stimulating factors (CSFs), such as granulocyte–colony-stimulating factors (G-CSFs), granulocyte macrophage–colony-stimulating factors (GM-CSFs); and IL-2–IL-7, IL-9, IL-11–IL-13, IL-15, IL-21, IL-23, IL-27, IL-31, and IL-35. Also included in this family are ciliary neurotrophic factor (CNTF), leukemia inhibitory factor (LIF), oncostatin M (OSM), and cardiotropin 1 (CT-1). Closely related are the interferons (IFN-α, -β, -τ, -ω, limitin) and IL-10-related cytokines,

TABLE 9.1 Cytokines Classified by Receptor Families

Receptor Family	Cytokine	Signaling	Source	Target	Action	Knock-out Phenotype
Type 1 (hemato-poietin)	GH	Janus kinase (JAK)2, signal transducer and activator of transcription 5b (STAT5b)	Two growth hormone (GH) genes, pituitary, placental	Diverse tissues	Growth, adipocyte differentiation	Dwarfism
	Prl	JAK2, STATa	Two *Prl* genes pituitary, uterus	Mammary epithelium	Growth, differentiation	Infertility, lactation defects
	Erythropoietin (EPO)	JAK2, STAT5	Kidney, liver	Erythroid precursors	Erythroid differentiation	Embryonic lethal, severe anemia
	Thrombopoietin (TPO)	JAK2, STAT5	Liver, kidney	Committed stem cells and megakaryocytes	Platelet	Severe thrombocytopenia
	Leptin	JAK2/STAT3	Adipocytes	Hypothalamus, thyroid	Satiety, controls metabolic rate	Obesity
	Granulocyte–colony-stimulating factor (G-CSF)	JAK2, STAT3	Many tissues, macrophages, endothelium, fibroblasts	Committed progenitors	Differentiation, activates mature granulocytes	Neutropenia
	Interleukin (IL)-6	JAK1, STAT3	Macrophages, fibroblasts, endothelium, epithelium, T cells, other	Liver, B cells, T cells, thymocytes myeloid cells, osteoclasts	Acute-phase reactants proliferation, differentiation, costimulation	Reduced immunoglobulin (Ig), especially IgA; T lymphopenia; impaired acute-phase response; and T-helper 17 (Th17) cells
	IL-11	JAK1, STAT3	Stromal cells, synoviocytes, osteoblasts	Hematopoietic stem cells, hepatocytes, macrophages, neurons	Proliferation	Female infertility
	IL-27	JAK1, STAT1, STAT3, STAT4, STAT5	Activated dendritic cells (DCs), macrophages, epithelial cells	T cells and natural killer (NK) cells, other cells	Enhancement of Th1 responses, and IL-10; inhibition of Th1, Th2, and Th17 responses	Fatal inflammatory disease with infection
	IL-31	JAK1, STAT3, STAT5	Th2 cells CD8 T cells	Monocytes, epithelial cells, keratinocytes, eosinophils basophils	Induces chemokines, PMN recruitment	
	Ciliary neurotropic factor (CNTF)[a]	JAK1, STAT3	Schwann cells	Neuronal	Survival	Progressive atrophy and loss of motor neurons
	Leukemia inhibitory factor (LIF)[a]	JAK1, STAT3	Uterus, macrophages, fibroblasts, endothelium, epithelium, T cells	Embryonic stem cells, neurons hematopoietic cells	Survival	Decreased hematopoietic progenitors, defective blastocyst implantation
	Oncostatin M (OSM)	JAK1, STAT3	Macrophages, fibroblasts, endothelium, epithelium	T cells, myeloid cells, liver, embryonic stem cells	Differentiation, acute-phase induction	
	Cardiotropin-1 (CT-1)	JAK1, STAT3	T cells, others, myocardium	Myocardium	Growth	
	Granulocyte macrophage–colony-stimulating factor (GM-CSF)	JAK1, STAT3	T cells, macrophages, endothelium, fibroblasts	Immature and committed myelomonocytic progenitors macrophages and granulocytes, DCs	Growth, differentiation survival, activation	Pulmonary alveolar proteinosis
	IL-3	JAK2, STAT5	T cells, macrophages, mast cells, natural killer T cells (NKT cells), eosinophils	Immature hematopoietic progenitors of multiple lineages	Growth, differentiation survival	No defects in basal hematopoiesis

TABLE 9.1 Cytokines Classified by Receptor Families—cont'd

Receptor Family	Cytokine	Signaling	Source	Target	Action	Knock-out Phenotype
	IL-5	JAK2, STAT5	Th2 T cells, activated eosinophils, NK cells, NKT cells	Eosinophil, B cells, basophils, mast cells	Proliferation, activation	Decreased eosinophilia, defective CD5, B1-cell development
	IL-2	JAK1, JAK3, STAT5	T cells, NK cells, NKT cells	T cells, B cells, NK cells, macrophages	Proliferation, cytotoxicity interferon-γ (IFN-γ) secretion, antibody production	Lymphoproliferation[a]
	IL-4[b]	JAK1, JAK3, STAT6	Th2 cells, mast cells, NKT cells, γ/δ T cells	T cells, B cells, macrophages	Proliferation, Th2 differentiation, IgG1 and IgE production, inhibition of cell-mediated immunity	Defective Th2 differentiation and IgE production, decreased allergic responses
	IL-7	JAK1, JAK3, STAT5	Bone marrow, thymic stromal cells, spleen DCs, keratinocytes, monocytes, macrophages	Thymocytes, T cells, B cells	Growth, differentiation, survival	Severe combined immunodeficiency (SCID)[a]
	IL-9	JAK1, JAK3, STAT5	Th2 and Th9 T cells, mast cells, eosinophils	T cells, B cells, mast cell precursors	Proliferation, Th1 inhibition	Not essential for Th2 pathology
	IL-15[b]	JAK1, JAK3, STAT5	Many cells	T cells, especially memory cells, NK and NKT cells	Proliferation, survival and activation	Absence of NK and memory cells
	IL-21	JAK1, JAK3, STAT3	T cells, Th17 cells, Tfh cells	T cells, B cells, and NK cells, DCs, macrophages, keratinocytes	Isotype switching, plasma cell differentiation, enhances CD8 and NK-cell responses, promotes Th17 cell differentiation	Acts in concert with IL-4 Decreased Th17 cells
	IL-13	JAK1, TYK2, STAT6	Activated T cells, NKT cells, mast cells, basophils	B cells, mast cells, macrophages, epithelial cells, smooth muscle cells	Costimulator of proliferation, IgE increased CD23 and class II, inhibits cytokine secretion and cell-mediated immunity	Defective Th2 responses and IgE production, decreased allergic responses
	IL-12	JAK2, TYK2, STAT4	Macrophages, DCs, B cells	T cells, NK cells	Th1 differentiation, proliferation, cytotoxicity	Defective Th1 differentiation, susceptibility to bacterial infections*
	IL-23	JAK2, TYK2, STAT3, STAT5	Macrophages, DCs	T cells, macrophages	IL-17 production	Reduced arthritis, inflammation
	IL-35	?	Tregs	T cells	Treg proliferation Suppresses proliferation and functions of Th17	Reduced Treg activity
	Thymic stromal lymphopoietin (TSLP)	JAK1, JAK2, STAT1 STAT3, STAT5	Epithelial cells, keratinocytes	DCs (human) B cells (mouse)	Th2 differentiation (human)	Shared receptor usage with IL-7R
Type II (interferon)	IFN-α/β	JAK1, TYK2, STAT1, STAT2	Plasmacytoid DCs, macrophages, fibroblasts, other	All, NK cells	Antiviral, antiproliferative increased major histocompatibility complex (MHC) class I activation	Susceptibility to viral infections[a]
	IFN-γ	JAK1, JAK2, STAT1	Th1 cells, NK cells	Macrophages, endothelium, NK cells	Activation, increased MHC class II expression, increased antigen presentation	Susceptibility to bacterial infections[a]
	IL-10	JAK1, TYK2, STAT3	Th2 cells, other cells	Macrophages	Decreased MHC class II expression, decreased antigen presentation	Exaggerated inflammatory response and autoimmune disease

Continued

TABLE 9.1 Cytokines Classified by Receptor Families—cont'd

Receptor Family	Cytokine	Signaling	Source	Target	Action	Knock-out Phenotype
	IL-19, -20, -22, -24, -26	STAT1, STAT3	T cells, monocytes, melanocytes, NKT cells	T cells, keratinocytes, epithelial cells	Induces production of inflammatory cytokines, Th2 responses, activation of epithelial cells	
	IL-28, -29, -30	STAT1, STAT2, STAT3, STAT4, STAT5	DCs, many cells	Many cells	Antiviral	
IL-1/TLR	IL-1α/β	IRAK (IL-1 receptor-associated kinase), MyD88, TRAF6 (TNF receptor–associated factor 6), nuclear factor (NF)-κB	Many cells, especially macrophages	Central nervous system, endothelial cells, liver, thymocytes, macrophages, T cells	Fever, anorexia, activation acute-phase reactants costimulation, activation, cytokine secretion, differentiation of Th17 cells	Reduced inflammation, cooperates with tumor necrosis factor (TNF) in host defense
	IL-18	IRAK, MyD88, TRAF6, NF-κB	Many cells, especially macrophages, keratinocytes, osteoblasts	T cells, NK cells, macrophages, epithelial cells		Increased susceptibility to infection, reduced arthritis
	IL-33			T cells, nuocytes (ILC2)	Enhanced Th2 responses	
	IL-36		Skin			
	IL-37					
	IL-38			Macrophages		
IL-17	IL-17A		Th17 cells, CD8 T cells, γ/δ T cells	Endothelium, many cells	Inflammation	Susceptibility to extracellular bacteria
	IL-17B, -C, -D		Many cells	Monocytes, epithelial cells	Inflammation, chondrogenesis	
	IL-17E (IL-25)	TRAF2	Mast cells, Th2 cells	Th2 cells	Enhanced Th2 responses	Increased susceptibility to helminths
	IL-17 F		Th17 cells, CD8 T cells, γ/δ T cells	Endothelium, many cells	Inflammation	
Transforming growth factor (TGF)-β receptor serine kinase family	TGF-β_1, -β_2, -β_3		T cells, macrophages, other	T cells, macrophages, other	Inhibits growth and activation, promotes Th17	
Receptor tyrosine kinases	Stem cell factor	Ras/Raf/ mitogen-activated protein kinase (MAPK), stromal cells	Bone marrow	Pluripotent stem cells	Activation, growth	Defective hematopoietic stem cell proliferation, melanocyte production and development
	CSF-1 (macrophage (M)-CSF)	Ras/Raf/MAPK	Macrophages, endothelium, fibroblast, other	Committed myelomonocytic progenitors	Differentiation, proliferation, survival	Monocytopenia, osteopetrosis, female infertility
	FMS-like tyrosine kinase 3 ligand (FLT-3) ligand	Ras/Raf/MAPK	Diverse tissues	Myeloid cells, especially DCs	Proliferation, differentiation	Reduced repopulating hematopoietic stem cells; reduced B-cell precursors
	IL-32	NF-κB, p38 MAPK	T cells, NK cells, monocytes, epithelia	Monocytes	Induces TNF, IL-1, IL-6, IL-8	
	IL-16		T and B cells, mast cells, eosinophils	CD4 T cells		
	IL-3	Extracellular signal–regulated kinase (ERK)	Many cells	Monocytes	Proliferation binds CSF-1 receptors	

In cases where STAT5a or STAT5b are designated, the cytokines appear to use either interchangeably.

[a] LIFR is shared by these cytokines.

[b] Note that two forms of the IL-4 and perhaps IL-15 receptor exist.

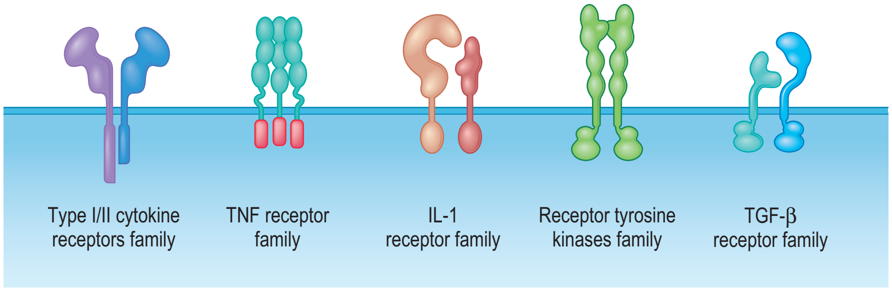

Type I/II cytokine receptors family
TNF receptor family
IL-1 receptor family
Receptor tyrosine kinases family
TGF-β receptor family

FIG 9.1 Schematic representation of prototypical receptors from five of the major cytokine receptor superfamilies.

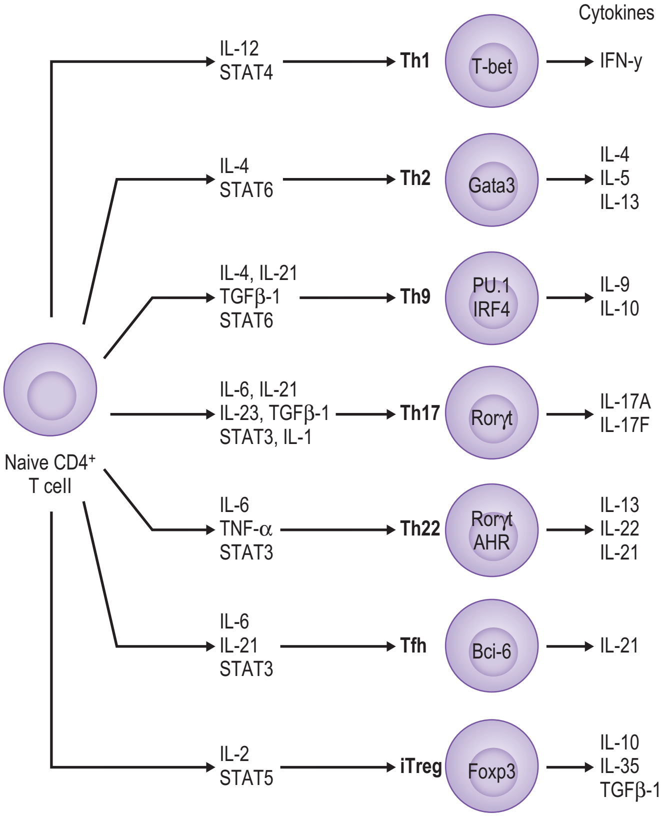


FIG 9.2 Differentiation of T-helper cell subtypes.

IL-19, IL-20, IL-22, IL-24, IL-26, and the IFN-related cytokines IL-28A (IFN-λ2), IL-28B (IFN-λ3), IL-29 (IFN-λ1), which bind type II receptors. The ligands and receptors in this superfamily are structurally similar and utilize related molecules for signal transduction.[5,6]

A central feature of type I cytokines is a similarity in their basic structure. Each contains four antiparallel α helices with two long and one short loop connections arranged in an up–up/down–down configuration. Because of this structure, these cytokines have also been referred to as the *α-helical bundle cytokine family.*

Structurally, the receptors in the type I family have conserved cysteine residues, a conserved Trp–Ser–X–Trp–Ser motif (where X indicates any amino acid), and fibronectin-like repeats in their extracellular domains. These receptors have a single transmembrane domain and divergent cytoplasmic domains. Within the cytoplasmic portion of these receptors, two segments of homology can be discerned, termed *Box 1* and *Box 2 motifs.* The membrane proximal domain binds Janus kinases (JAKs; see below). Some of the cytokine receptors are homodimers, such as the receptors for EPO, TPO, PRL, and possibly leptin, whereas other receptors for type I cytokines are heterodimers, containing two distinct receptor subunits. On the basis of this characteristic, the type I family of receptors can be divided into subfamilies. Each member of the subfamily uses a shared receptor subunit in conjunction with a ligand-specific subunit. For example, the receptors for IL-2, IL-4, IL-7, IL-9, IL-15, and IL-21 all use a common cytokine γ chain, γc (see Table 9.1), whereas a common β chain, βc, is shared by IL-3, IL-5, and GM-CSF. Similarly, gp130 is a shared subunit for IL-6 family cytokines (IL-6, IL-11, IL-27, IL-35, CNTF, LIF, OSM, and CT-1). IL-12 and IL-23 also share a receptor subunit, as do members of the IL-10 family.

Other levels of shared receptor usage also exist. For example, the receptors for LIF, CNTF, OSM, and CT-1 all share the LIF receptor subunit, IL-31 and OSM also share one receptor chain, whereas IL-2 and IL-15 utilize the same β and γ c chains. Conversely, IL-4 can bind two different receptor complexes. The classic IL-4 receptor is composed of the IL-4Rα chain and the γc chain. Additionally, IL-4 can also bind the IL-13 receptor, which comprises a heterodimer of the IL-4Rα chain and the IL-13Rα chain. IL-13 only utilizes the IL-13 receptor complex for signaling.

The utilization of common receptor subunits explains the phenomenon of shared biological activities (cytokine redundancy) between cytokines that belong to the same subfamily. Within a subfamily, actions distinct for each cytokine can be attributed, at least in part, to the ligand-specific subunits. The pleiotropic effects of a single cytokine can be accounted for by the existence of more than one receptor for that cytokine.

Family Members and Their Actions

Homodimeric Receptors

Many of the cytokines that use homodimeric receptors are classic hormones. These include EPO GH, PRL, and leptin. EPO is required for erythrocyte growth and development and is widely used to treat anemia. Similarly, TPO is required for megakaryocyte development and may have a use in the treatment of thrombocytopenia. G-CSF not only regulates the production of neutrophils through its action on committed progenitor cells but also supports the survival of mature neutrophils, enhancing their functional capacity. G-CSF is widely used clinically to treat patients with granulocytopenia. As one would predict, G-CSF–deficient mice have marked neutropenia, and mutations of the G-CSF receptor (G-CSFR) result in severe congenital neutropenia in humans.

Cytokine Receptors Utilizing gp130

gp130 is a receptor component for IL-6, IL-11, IL-27, and IL-35 as well as LIF, OSM, CNTF, and CT-1.[7] Targeted disruption of the gp130 gene is lethal in early embryogenesis, causing defects in myocardial, hematological, and placental development. LIF binds to gp130 in association with the LIF receptor (LIFR), as do the cytokines OSM, CNTF, and CT-1. Deletion of the *LIFR* gene is also embryonically lethal, creating defects in placental architecture and developmental abnormalities in neural tissue and bone. Targeted disruptions of LIF lead to failure of blastocyst implantation. Another critical role of LIF is the maintenance of stem cell pluripotency in culture.

Interleukin-6. The IL-6 receptor (IL-6R) consists of a soluble IL-6 binding protein (α chain) (CD126) and membrane bound gp130. IL-6 has a wide array of biological actions on both lymphoid and nonlymphoid cells with the consequences of signaling by the membrane bound and soluble receptors being distinct.[8] IL-6 is important in host defense, and IL-6–deficient mice are susceptible to infection with *Candida* and *Listeria.* IL-6 is a growth and differentiation factor for B cells, inducing the production of immunoglobulin (Ig), including IgE. IL6$^{-/-}$ mice have normal numbers of B cells with reduced Ig response to immunization and reduced IgA production. IL-6 also promotes T-cell growth and differentiation. Consequently, IL6$^{-/-}$ mice have reduced numbers of thymocytes and peripheral T cells. IL-6 is important for T-helper 17 (Th17) cell differentiation and the cytotoxic T-cell response to viruses. IL-6 functions synergistically with IL-3 in hematopoiesis, and IL-6–deficient mice have reduced numbers of progenitor cells.

IL-6 is a major inducer of fever, inflammation, and the synthesis of acute-phase proteins (*e.g.,* fibrinogen, serum amyloid A, haptoglobin, C-reactive protein [CRP], etc.) in the liver. The elevation of the erythrocyte sedimentation rate (ESR) in inflammatory disease largely reflects the accelerated synthesis of these proteins, and IL-6–deficient mice are defective in this response. IL-6 reduces synthesis of albumin and transferrin in the liver and initiates hepatocyte regeneration. IL-6 induces adrenocorticotrophic hormone and anterior pituitary hormones, such PRL, GH, and luteinizing hormone (LH). IL-6 also plays a role in osteoporosis by affecting osteoclast function. IL-6–deficient mice are protected from bone loss following estrogen depletion.

Levels of IL-6 in serum are low in the absence of inflammation but rapidly increase in response to bacterial and viral infections, inflammation, or trauma. Patients with rheumatoid arthritis (RA), cardiac myxoma, Castleman disease, and other autoimmune diseases have high serum levels of IL-6. This cytokine may also contribute to malignancies, such as multiple myeloma.

IL-6 is produced by many cells, but its expression in mononuclear phagocytes has been well documented. Stimulation of monocytes with IL-1, TNF, or lipopolysaccharide (LPS) stimulates the expression of IL-6, whereas IL-4 and IL-13 inhibit its production. The *IL6* gene contains binding sites for nuclear factor-κB (NF-κB), nuclear factor for IL-6 (NF-IL-6, or CCAAT element-binding protein), activator protein-1 (AP-1), cyclic adenosine monophosphate (cAMP) response element-binding protein (CREB), and the glucocorticoid receptor.

Not surprisingly, much effort has been made to develop biological IL-6/IL-6R antagonists. Tocilizumab, an anti–IL-6 receptor antibody approved for treatment of RA, juvenile idiopathic arthritis, and other diseases and another anti–IL-6R monoclonal antibody (mAb), sarilumab, is in development. Anti–IL-6 mAbs (olokizumab, siltuximab, and clazakizumab) are also being developed for similar indications.

Interleukin-11. IL-11 and its receptor are widely expressed. IL-11 stimulates stem cells, megakaryocytes, myeloid precursors, and erythroid precursors, as well as promoting B-cell differentiation. It also acts on nonhematopoietic cells, including bone and liver cells. IL-11 is induced by proinflammatory cytokines (IL-1, TNF) and by TGF-β.

Interleukin-27. IL-27 is composed of two subunits designated EBI3 and p28 and signals through gp130 and WSX-1/TCCR (T-cell cytokine receptor). The receptor is expressed on naïve CD4 T cells. IL-27 promotes Th1 differentiation but also has essential antiinflammatory properties, inhibiting Th17 differentiation and enhancing IL-10 production.[9]

Cytokine Receptors Utilizing the β_c Chain

IL-3, IL-5, and GM-CSF bind to a ligand-specific α subunit associated with the common βc receptor subunit (common β subunit). Mice, but not humans, have a second β chain, βIL3. This species-specific redundancy may explain why gene targeting of βc in the mouse did not result in loss of IL-3 responses, although βc-null mice did have reduced GM-CSF and IL-5 responses.

Interleukin-3. IL-3 synergizes with other cytokines to stimulate the growth of immature progenitor cells of all lineages and is termed *multilineage colony-stimulating factor*. It promotes survival of macrophages, mast cells, and megakaryocytes. IL-3 is produced mainly by lymphoid cells, but also by mast cells and eosinophils. IL-3–deficient mice have no obvious defect in hematopoiesis, suggesting that the major role of IL-3 *in vivo* may be in the response to stress.

Interleukin-5. IL-5 is unusual in that it is a disulfide-linked homodimer, with each component containing three α-helical bundles. It promotes the growth, differentiation, and activation of eosinophils and so is very important in pathogenesis of allergic disease. IL5$^{-/-}$ mice fail to develop eosinophilia in response to parasitic or aeroallergen challenge and exhibit minimal signs of inflammation and damage to lungs. IL-5 deficiency does not affect the worm burden of infected mice, indicating that eosinophilia may not play an essential role in the host defense against helminths *per se.* Both IL-5 and IL-5R knock-out mice have decreased numbers of CD5^{+} B cells (B-1 cells) and concomitant low serum IgM and IgG3 levels. IL-5 is produced by activated helper T cells of the Th2 phenotype (see below), mast cells, and eosinophils in an autocrine manner. Mepolizumab and reslizumab are anti–IL-5 mAbs that have been approved for the treatment of severe eosinophilic asthma disease.

Granulocyte macrophage–colony-stimulating factor. GM-CSF acts on hematopoietic precursors to support myelomonocytic differentiation. It activates mature neutrophils and macrophages, increasing their microbicidal activity and inducing the production of proinflammatory cytokines. Along with IL-4 and IL-13, GM-CSF is a major stimulatory cytokine for the *in vitro* production of dendritic cells (DCs). GM-CSF induces proliferation and activation of eosinophils and upregulates adhesion molecules on fibroblasts and endothelial cells. Deletion of the GM-CSF gene (*Csf2*) in mice, however, does not affect steady-state hematopoiesis. Instead, these animals develop alveolar proteinosis and lymphoid hyperplasia. βc$^{-/-}$ mice also develop alveolar proteinosis, characterized by the accumulation of surfactant in the lungs. A similar defect may be responsible for disease in a subset of humans with this abnormality.

The production of GM-CSF can be induced by proinflammatory cytokines and LPS. GM-CSF is made by activated lymphocytes and other stimulated cells and is an important driver of immune pathology in murine models of autoimmunity.[10,11] GM-CSF is not ordinarily detectable in blood except under pathological conditions, such as asthma. GM-CSF has been used clinically to treat chemotherapy-induced neutropenia, especially in the context of certain infections (*e.g.,* fungal), and has been tested in myelodysplastic syndrome and aplastic anemia. Mavrilimumab is an anti–GM-CSF antibody being studied in rheumatoid arthritis.[12]

Cytokine Receptors Utilizing the γc Chain

The cytokines IL-2, IL-4, IL-7, IL-9, IL-15, and IL-21 all bind to receptors that share a common γc receptor subunit. The γc subunit and the ligand-specific subunits are expressed predominantly on lymphocytes, although they can be found on other hematopoietic cells as well. Mutation of the γc gene is responsible for X-linked severe combined immunodeficiency (SCID), characterized by a lack of T cells and natural killer (NK) cells, and poorly functioning B cells (T^-B^+ SCID) (Chapter 35).[13] The lack of γc abrogates signaling by all cytokines that utilize this subunit (IL-2, IL-4, IL-7, IL-9, IL-15, and IL-21). The lack of IL-7 and IL-15 signaling is largely responsible for the lack of T and NK cells, respectively.

Interleukin-2. The IL-2 receptor consists of three subunits, α, β, and γc. The latter two are members of the type I cytokine receptor family. NK cells constitutively express these latter two subunits and respond to high doses of IL-2, whereas in T cells the IL-2Rα subunit is induced upon activation, creating a high-affinity receptor for IL-2. IL-2Rα is also inducible in activated monocytes and B cells. IL-2Rα, however, is not a member of the type I cytokine receptor family. Rather, it resembles members of the complement family and IL-15R (see below).

IL-2, one of the first cytokines to be intensively studied, is produced principally by activated T cells. It is a prototypical autocrine T-cell growth factor and is required for *in vitro* T-cell proliferation. It is an important factor in determining the magnitude of T-cell and NK-cell responses *in vivo,* although other factors also contribute. It is thus redundant to some degree for this function. It augments the cytolytic activity of T and NK cells and enhances IFN-γ secretion. IL-2 is also important in programming CD8 memory T cells, which undergo secondary expansion in viral infections.[9] IL-2 is a growth factor for B cells and induces class switching. It also activates macrophages.

IL-2–deficient mice develop massive enlargement of peripheral lymphoid organs, hemolytic anemia, inflammatory bowel disease (IBD), and infiltrative granulopoiesis. In addition, the mice have high levels of IgG1 and IgE. They succumb to widespread autoimmune and lymphoproliferative disease, which points to the nonredundant roles of IL-2 in promoting tolerance and preventing autoimmunity.[14] A similar phenotype is seen in humans with IL-2Rα mutations.

Regulatory T cells (Tregs) in both mice and humans express high levels of IL-2Rα. Maintenance of peripheral Treg numbers is dependent on IL-2.[15] In addition, IL-2 inhibits Th17 and T-follicular helper (Tfh) cell differentiation.[16,17]

IL-2 is produced by activated T cells but also myeloid cells, including DCs. The *IL2* gene has been extensively characterized and contains binding sites for nuclear factor of activated T-cells (NFATs), AP-1, and NF-κB. IL-2 production is also regulated by stabilization of its messenger RNA (mRNA).

IL-2 is approved by the US Food and Drug Administration (FDA) for the treatment of renal cancer; however, its clinical utility is limited by its toxicity, two important manifestations being hepatic dysfunction and the so-called capillary or vascular leak syndrome. However, the anti–IL-2Rα mAbs daclizumab and basiliximab are used to prevent rejection of allotransplants and, more recently, have been found to be effective in the treatment of multiple sclerosis (MS).[18] Polymorphisms of IL-2Rα are also associated with MS.

Interleukin-4. Two classes of IL-4R appear to exist. The first consists of the IL-4Rα subunit in conjunction with γc. It is expressed on hematopoietic cells.[19] The second, which is more widely expressed, is a “type II” receptor. It consists of the IL-4Rα subunit in associated with IL-13R α chain. Two IL-13R subunits have been cloned, so the exact composition of the type II IL-4R remains uncertain. The existence of two receptors helps explain why IL-4 has diverse actions on both hematopoietic and nonhematopoietic cells. The loss of IL-4Rα would be predicted to block the actions of both IL-4 and IL-13, which could explain why gene targeting of IL-4Rα leads to a more severe phenotype than that observed in IL-4–deficient mice. Polymorphisms of IL-4Rα have been reported to be associated with a propensity to atopy.[20]

In general, IL-4 promotes allergic responses and inhibits cell-mediated immune responses. Among the most important roles of IL-4 is its ability to promote differentiation of naïve CD4 T cells into a subset that produces IL-4 and IL-5,[21] denoted as Th2 subset, as opposed to Th1 cells, which produce IFN-γ (Chapter 16). In conjunction with CD40 activation, IL-4 also promotes B-cell proliferation and class switching, particularly to IgG1 and IgE in mice and to IgG4 and IgE in humans. Mice deficient in IL-4 have normal B lymphopoiesis but marked reductions in IgG1 and IgE production in response to parasites. These mice have residual Th2 responses because IL-13, which also binds IL-4Rα, can partially compensate for the defect.

IL-4 upregulates the expression of surface IgM, major histocompatibility complex (MHC) class II, and CD23 on B cells. In conjunction with GM-CSF, it is a growth factor for mast cells and basophils, as well as a potent inducer of DC differentiation. IL-4 inhibits macrophage activation and the production of proinflammatory cytokines. It antagonizes the effects of IFN-γ, blocks cytokine-induced proliferation of synoviocytes, downregulates the expression of adhesion molecules, and antagonizes the induction of some acute-phase reactants in hepatocytes by IL-6.

IL-4 is made by the Th2 subset of CD4 T cells, NK1.1$^+$ CD4$^+$ T cells, basophils, and mast cells. A number of transcription factors appear to be important in regulating IL-4 production, including NFAT, NF-IL6, C/EBP, c-MAF, and GATA-3. The *IL4* gene has multiple signal transducer and activator of transcription 6 (STAT6) binding sites, consistent with the fact that IL-4 regulates its own expression. Epigenetic control and chromatin remodeling are also important aspects.[22]

Clinically, IL-4 has been tested in the treatment of malignancies and some autoimmune disorders. The ability of IL-4 to generate DCs is being exploited in the use of tumor vaccines. Conceivably, soluble IL-4R might also be useful in the treatment of allergic disease. Dupilumab is an anti-IL-4/13R mAb being studied in severe atopic dermatitis.[23]

Interleukin-7. The IL-7 receptor consists of the IL-7Rα chain (CD127) in association with γc. It is expressed on both immature and mature thymocytes. Humans with loss-of-function (LOF) mutations of IL-7Rα have T^-B^+ SCID but, unlike in individuals with γc mutations, display normal NK-cell development (Chapter 35). Gain-of-function (GOF) mutations of the IL-7Rα chain result in constitutive JAK1 signaling and cell transformation and give rise to T-cell acute lymphoblastic leukemia (Chapter 78).

IL-7Rα expression is tightly regulated during thymocyte development (Chapter 8). IL-7 plays an important role in both developing thymocytes and mature T cells. Its receptor is expressed in double-negative thymocytes, downregulated in double-positive cells, and then reexpressed in single-positive thymocytes and mature peripheral T cells. This may be a reflection of its anti-apoptotic effects, which are attributable to the induction of Bcl-2 family members. IL-7 promotes the growth of thymocytes, as well as the expression and rearrangement of T-cell receptor (TCR) genes and the expression of *RAG1* and *RAG2* (Chapter 4). IL-7Rα is expressed on cutaneous T-cell lymphomas, which also produce this cytokine; thus the autocrine response to IL-7 may contribute to the growth of these tumors.

IL-7– and IL-7R–deficient mice exhibit impairments in both T- and B-cell development. Postnatal B-cell development in $Il7^{-/-}$ mice is blocked at the transition to pre–B cells and is arrested even earlier in $Il7ra^{-/-}$ mice. Why these abnormalities do not occur in humans with IL-7Rα mutations is not clear.

IL-7 is produced by a wide variety of cells, including marrow and perhaps thymic stromal cells, as well as in the kidney, spleen, and epithelial cells and keratinocytes. This is consistent with its role in the maintenance of function in both immature and mature lymphocytes.

Clinically, IL-7 may be useful to restore immune function in some congenital immunodeficiencies, after bone marrow transplantation, and in human immunodeficiency virus (HIV) infection. Polymorphisms of the *IL7R* gene are associated with MS.

Interleukin-9. IL-9 has some of the same properties as IL-4. It synergizes with stem cell factor to promote the growth and differentiation of mast cells and regulate mast cell function. IL-9 potentiates IgE production induced by IL-4 in B cells. Although first identified as a T-cell growth factor, a physiological role in T-cell development has not been established. IL-9 is produced by activated Th2 cells, mast cells, and eosinophils. Recently, a new subset of Th cells, termed *Th9,* has been proposed. These cells have been shown to not only secrete IL-9 but also to enhance inflammatory responses. Interestingly, IL-9 inhibits Th1 cytokine production. Some lymphoid tumors also produce IL-9, where it may serve as an autocrine growth factor. A subset of innate lymphoid cells, ILC2 (Chapter 3), also produces IL-9, which is thought to contribute to the pathogenesis of allergy and asthma.

Interleukin-15. The IL-15 receptor consists of the IL-2Rβ and γc subunits in association with a unique ligand-specific subunit, IL-15Rα, which is homologous to IL-2Rα. These receptor proteins contain protein-binding motifs termed "sushi domains." In both humans and mice, these receptors and their cognate ligands are physically linked in the genome. Given their shared receptor usage, there are many similarities in the actions of IL-2 and IL-15, particularly in terms of the effects on lymphoid cells. Like IL-2, IL-15 induces proliferation and cytokine production in T and NK cells. However, despite the similarities between these two ligands/receptors, there are some important differences. In T cells, IL-15 is less efficient than IL-2 in inducing effector memory T-cell differentiation or sensitivity to apoptosis.[24] IL-15Rα is more widely expressed than IL-2Rα, IL-2Rβ, and γc. IL-15Rα is expressed by lymphoid cells, DCs, fibroblasts, and epithelial, liver, intestine, and other cells and is thought to present IL-15 in *trans* to cells expressing IL-15β and γ chains. IL-15– and IL-15Rα–knock-out mice are defective in NK-cell production and in the generation of memory T cells, explaining the absence of NK development in patients with γc mutations.

IL-15 mRNA is expressed broadly in hematopoietic and nonhematopoietic cells but is not typically produced by T cells. (human T-lymphotropic virus 1 [HTLV-1]–transformed T cells are an exception in that they produce abundant IL-15.) Following the pattern seen in IL-7 and IL-9, there are multiple triplet repeats of adenosine, uradine, guanosine (AUGs) in the 5′ untranslated portion of the IL-15 message that interfere with translation. Thus IL-15 is controlled by translational regulation. IL-15 protein is also controlled at the level of secretion of the protein, but this is incompletely understood. High levels of IL-15 protein have been reported in synovial fluids from patients with RA, alveolar macrophages from patients with sarcoidosis, and peripheral blood mononuclear cells from patients with ulcerative colitis.

Interleukin-21. IL-21 is a T-cell–derived cytokine that works in concert with other γc cytokines. It synergizes with IL-7 and IL-15 to expand and activate CD8 T cells. IL-21 also augments the activity of NK cells. IL-21, along with IL-6, drives differentiation of Tfh. Tfh cells are found preferentially in B-cell follicles, where, under the control of the transcription factor BCL6, they regulate B-cell development, activation, and class switching. Tfh cells are also a source of IL-21.[25] IL-21 appears to have some anticancer properties, and it has been tested in the treatment of melanoma.[26,27]

Other Heterodimeric Receptors

Interleukin-12. IL-12 is a heterodimer composed of two disulfide-linked polypeptide chains, p35 and p40, derived from two distinct genes.[28,29] IL-12 p35 shares homology with other cytokines, such as IL-6, whereas p40 resembles the IL-6 receptor. Thus IL-12 can be viewed as being synthesized as a ligand–receptor complex. IL-12R consists of two chains, and because the ligand already comprises the α subunit, the two chains of the IL-12R are denoted IL-12Rβ_1 and β_2. Expression of high-affinity IL-12R is very restricted, being found predominantly on T and NK cells. IL-12Rβ_1 and β_2 are highly inducible upon T-cell activation, and IL-4 inhibits IL-12Rβ_2 expression. This is important because IL-12Rβ_2 is required for IL-12 signaling and the activation of the STAT4. NK cells constitutively express IL-12Rβ_1 and IL-12Rβ_2.

IL-12 plays a pivotal role in promoting cell-mediated immune responses. Humans with *IL12R* mutations, as well as mice with IL-12 and IL-12R deficiency, have very blunted immune responses and are highly susceptible to infections by intracellular pathogens. An important function of IL-12 is that it promotes the differentiation of uncommitted Th cells to the Th1 subset (*i.e.,* T cells that produce LT-α and INF-γ). Th1 differentiation is markedly impaired in IL-12– and IL-12R–deficient mice. A major action of IL-12 is its ability to induce the production of IFN-γ, doing so synergistically in combination with IL-2 or IL-18. Consequently, many of the actions of IL-12 are blocked in IFN-γ or IFN-γR knock-out mice. IL-12 also induces proliferation and cytolytic activity of T and NK cells.

DCs and macrophages are the major producers of IL-12 in response to various pathogens, occupancy of TLRs, and CD40. The *IL12p40* gene is complex and contains NF-κB sites, IFN response elements (IREs), and ETS-binding sites. As with other cytokine genes, nucleosome remodeling is important in the regulation.[30]

Because of its profound effects on cell-mediated immunity, IL-12 has been used in the treatment of malignancies and infectious diseases. However, its utility has been limited because of significant toxicity. IL-12 may also have use in vaccines as an adjuvant. Conversely, antagonizing the actions of IL-12 has been found to be useful in Th1-mediated diseases, including IBD (Chapter 75). Ustekinumab inhibits both IL-12 and IL-23 and is approved for psoriasis and psoriatic arthritis.

Interleukin-23. IL-23 is another heterodimeric type I cytokine. It is composed of two disulfide-linked polypeptide chains, p19 and IL-12 p40. The IL-23 receptor also shares the IL-12Rβ_1 chain paired to the IL-23R. The IL-23R complex is expressed on T cells and ILCs and regulates production of IL-17 (Th17, see below). As such, IL-23 is thought to be important in host defense against extracellular bacteria and the pathogenesis of autoimmune and autoinflammatory disorders. IL-23 is produced primarily by DCs in response to TLR agonists. IL23R polymorphisms are associated with IBD, ankylosing spondylitis, and other autoimmune diseases. Tildrakizumab and guselkumab inhibit IL-23 but not IL-12.

Interleukin-35. IL-35 is a dimer consisting of IL-12 p35 and EB13. It is preferentially produced by Tregs. Tregs are also the main cellular target of IL-35, where it induces proliferation and production of IL-10. A synthetic form of IL-35, obtained by covalently linking EBI3 to IL-12p35, can reduce the incidence of arthritis in mouse models.[26]

Interleukin-13. IL-13 has many of the same effects as IL-4 and shares a receptor subunit(s) with IL-4. IL-13–deficient mice have reduced levels of IL-4, IL-5, and IL-10, with lower IgE levels and eosinophils. In mice deficient for both IL-4 and IL-13, these Th2 responses are abolished, and the ability to clear parasites is severely impaired. These double-knock-out mice default to Th1 responses, with concomitant production of INF-γ, IgG2a, and IgG2b. It thus appears that IL-4 and IL-13 cooperate in promoting Th2 responses, having both overlapping and additive roles. Anti–IL-13 mAbs in clinical trials include lebrikizumab and tralokinumab for severe asthma.[31,32]

Interleukin-31. IL-31 signals through the heterodimeric receptor IL-31Rα and oncostatin M receptor (OSMR). It is produced by activated Th2 cells. Overexpression of IL-31 results in atopic dermatitis, but surprisingly, IL-31Rα–deficient mice showed an increased Th2 response.[33,34]

Thymic stromal lymphopoietin. Thymic stromal lymphopoietin (TSLP) is an IL-7–like cytokine expressed by epithelial cells and keratinocytes. Its receptor comprises TSLP receptor (TSLPR) and IL-7Rα, which is expressed primarily on monocytes and myeloid-derived DCs, as well as on B cells. TSLP-treated human DCs promote Th2 differentiation.[35] A major means by which TSLP exerts its effect is through promotion of basophil hematopoiesis.[36] Elevated TSLP levels have been found in humans and animal models of airway inflammatory disease and atopic dermatitis. In the mouse, TSLP contributes to prenatal B-cell development.

Interferons

Type I Interferons

Interferon-α/β. The type I IFNs include IFN-α, IFN-β, and IFN-ω. IFN-β and IFN-ω are encoded by single genes, whereas IFN-α includes at least 14 separate genes, each encoding structurally distinct forms. These intronless genes are all clustered on the short arm of chromosome 9 and appear to have diverged from a common ancestor more than 100 million years ago. Each of these molecules binds to the same IFN-α/β receptor, and their actions are similar. The receptor is a heterodimer composed of two subunits termed *IFNAR1* and *IFNAR2*. These subunits have limited similarity to type I cytokine receptors, although they lack the WSXWS motif.

A major effect of type I IFNs is their antiviral action.[37] Discovered in 1957, they act on all cells to inhibit viral replication as well as cellular proliferation. It is unclear why there are so many type I genes. Given that their relative potencies differ, it is possible that these genes evolved in response to various viral pathogens. Alternatively, IFN gene duplication may affect the magnitude of antiviral responses. A major mechanism is the inhibition of protein translation. Type I IFNs also upregulate MHC class I and can block the ability of IFN-γ to upregulate MHC class II expression.[38] IFN-α/β increase the cytolytic activity of NK cells. IFNARI knock-out mice are extremely susceptible to infections, even though lymphoid development is normal.

IFNs are produced ubiquitously. Recognition of extracellular and intracellular foreign DNA, produced in viral infection, is a major inducer of their transcriptional regulation. Type I IFN is also induced by intracellular bacterial pathogens and LPS. Immunoregulatory effects of IFN-α/β are being increasingly recognized, and it is notable that a subset of DCs produces very high levels.[39,40] IFN genes are bound by multiple transcription factors, including NF-κB, interferon regulatory factor 3 (IRF-3), IRF-7, and STAT1.

Type I IFN is used clinically in the treatment of certain infections (*e.g.,* viral hepatitis). Because of its antiproliferative action, it is also used in the treatment of certain malignancies, particularly hairy cell leukemia. IFN-β is used in the treatment of MS.

Newer IFN-like cytokines, including IL-28A, IL-28B, and IL-29 (also designated IFN-λ1, -λ2, and -λ3), have been identified. They bind to a receptor designated IL-28R. The exact *in vivo* functions of these IFN-like cytokines are poorly understood, although they probably contribute to antiviral responses.

Interferon-γ. IFN-γ is a major activator of macrophages, enhancing their ability to kill microorganisms by augmenting their cytolytic machinery. IFN-γ exerts this effect by causing the cell to increase its production of reactive oxygen intermediates, including hydrogen peroxide, nitric oxide, and indoleamine dioxygenase. It also upregulates MHC class II expression. IFN-γ acts on CD4 T cells to promote Th1 differentiation while inhibiting the generation of Th2 cells. It promotes the maturation of CD8 T cells to cytotoxic cells. IFN-γ augments NK-cell cytolytic activity and regulates B-cell class switching. Endothelial cells and neutrophils are also activated by IFN-γ. Like IFN-α/β, IFN-γ also contributes to antiviral defenses.

The IFN-γ receptor is a heterodimer composed of two subunits, IFN-γRα and IFN-γRβ. When one IFN-γ homodimer binds, a complex of two α and two β receptors is created.[41] Mice with a disrupted IFN-γR develop normally and have normal lymphoid development but are highly susceptible to viral and bacterial infections, especially those by intracellular microbes. They have diminished macrophage MHC class II expression, decreased NK function, and reduced serum IgG2a concentrations. Humans with mutations of IFNGR subunits (OMIM #209950) are also susceptible to mycobacterial and *Salmonella* infections.

IFN-γ is produced by Th1 and NK cells. Transcription factors, including STAT4, T-BET, and EOMES, play an important roles in IFN-γ gene regulation.[22] IFN-γ has been used to treat patients with immunodeficiencies (*e.g.,* chronic granulomatous disease [CGD]) and in certain patients with disseminated mycobacterial infections.[22] A monoclonal anti-IFN-γ antibody, fontolizumab, is being studied in the treatment of autoimmune diseases.

Interleukin-10 and related cytokines. The major function of IL-10 is to serve as an antiinflammatory and immunosuppressive cytokine. Unlike other cytokines in this family, it is a disulfide-linked dimer. A single IL-10R has been cloned, but the receptor may have additional components. IL-10R is expressed on macrophages, mast cells, and most other hematopoietic cells. It is also inducible in nonhematopoietic cells by stimuli, such as LPS.

IL-10 strongly inhibits the production of proinflammatory cytokines. It inhibits macrophage antigen presentation and decreases expression of MHC class II, adhesion molecules, and the costimulatory molecules CD80 (B7.1) and CD86 (B7.2). The importance of IL-10 as an endogenous inhibitor of cell-mediated immunity is emphasized by the finding that IL-10–deficient mice develop autoimmune disease, which manifests with severe IBD and exaggerated inflammatory responses. Humans with mutations of the *IL10* and *IL10R* genes suffer from IBD.

IL-10 is made by T cells, B cells, macrophages, keratinocytes, bronchial epithelial cells, and other cells. LPS and TNF are inducers of IL-10. IL-10 is readily detected in the blood of patients with septic shock and other inflammatory and immune disorders. Because of its antiinflammatory properties, IL-10 has been used experimentally in the treatment of some Th1-mediated autoimmune diseases. IL-10 is elevated in patients with systemic lupus erythematosus (SLE), and there is a correlation between levels of IL-10 and autoantibody production.

There are viral homologues of IL-10 that may blunt the immune response to these pathogens. IL-10 also contributes to the immunosuppression seen in lepromatous leprosy or parasitic infestations. Other IL-10–related cytokines include IL-19, IL-20, IL-22, IL-24, and IL-26, but their biological actions are incompletely understood.[42]

Signaling

Neither type I nor type II receptors exhibit intrinsic enzymatic activity. However, the conserved membrane proximal segment of each of these receptors serves as the site at which these receptors bind JAKs (see Table 9.1). These JAKs play a pivotal role in signaling via this family of cytokine receptors.[5]

Janus Kinases

Four mammalian JAKs, JAK1, JAK2, JAK3, and TYK2, have been identified. JAKs are structurally unique, consisting of a C-terminal catalytically active kinase domain that is preceded by a segment termed the *pseudokinase domain.* The latter gives JAKs their name and has regulatory functions. A key feature of the JAKs is their association with cytokine receptors, which appears to be mediated by the N-terminus.

Ligand binding to type I and II receptors induces the aggregation of receptor subunits, which brings JAKs in close proximity and allows them to phosphorylate and activate each other. After activation, the JAKs phosphorylate receptor subunits on tyrosine residues, which allow the recruitment of proteins with SRC homology-2 (SH2) or phosphotyrosine-binding (PTB) domains. These proteins can also be phosphorylated by JAKs. Phosphorylation results in the activation of a number of biochemical pathways. Importantly, phosphorylation of cytokine receptors generates docking sites for a class of SH2-containing transcription factors termed *STATs* (see below) (Fig. 9.3).

The pivotal function of the JAKs is vividly illustrated by mice or humans that are deficient in these kinases. Consistent with the association of JAK3 with the common gamma chain, γc, mutations of JAK3 cause autosomal recessive T^-B^+ SCID (Chapter 35).[43] *Jak3*$^{-/-}$ also exhibit SCID. Indeed, mutation of either γc or JAK3 leads to the same functional defects. These findings led to the notion that JAK inhibitors might represent a new class of immunomodulatory drugs. Multiple JAK inhibitors have been approved or are in late phase clinical development (Chapter 87).

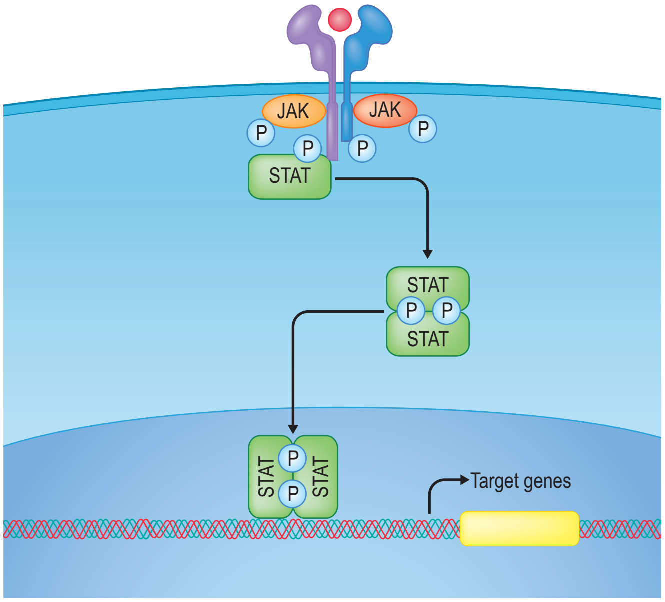


FIG 9.3 The roles of Janus kinases (JAKs) and signal transducer and activator of transcription factors (STATs) in signal transduction by type I and II cytokine receptors.

Gene targeting of *Jak1* and *Jak2* results in diverse abnormalities. *Jak1*$^{-/-}$ mice not only die perinatally as a result of neurological defects but also have a SCID phenotype similar to *Jak3*$^{-/-}$ mice. This is explained by the fact that γc-containing cytokine receptors utilize JAK1 in association with their ligand-specific receptor subunit. Other cytokines that are dependent on JAK1 include those that use gp130 cytokine receptors (IL-6, LIF, OSM, CNTF, and IL-11) and type II receptors (IL-10, IFN-γ, and IFN-α/β). Gene-targeting of *Jak2* was embryonically lethal, principally because JAK2 is essential for erythropoietin function, and the mice fail to form blood. In addition, JAK2 is necessary for signaling of other cytokines, including IL-3.

The importance of the JAKs is also substantiated by mutations in many leukemias and lymphomas. Mutation in the pseudokinase domain of *JAK* underlies most cases of polycythemia vera. Loss of function mutations of *TYK2* are associated with immunodeficiency.

Signal Transducer and Activator of Transcription (STAT)

Members of the STAT family of DNA-binding proteins serve a key role in transducing signals from cytokine receptors on the cell surface to the nucleus, where they regulate gene transcription. STATs are latent, cytosolic transcription factors that have SH2 domains (phosphotyrosine-binding modules) that allow them to be recruited to phosphorylated cytokine receptors (see Fig. 9.3). Different STATs bind to specific cytokine receptors (see Table 9.1). STATs are themselves tyrosine phosphorylated by JAKs, and this promotes their dimerization. STATs translocate to the nucleus, bind DNA, and regulate transcription.

There are seven mammalian STATs: STAT1, STAT2, STAT3, STAT4, STAT5A, STAT5B, and STAT6. *Stat* knock-out mice document the essential and specific functions of these transcription factors in transmitting cytokine signals. *Stat1*$^{-/-}$ mice develop normally but have extreme susceptibility to viral and some bacterial infections, consistent with the defects seen in *Ifng*$^{-/-}$ and *Ifngr*$^{-/-}$ mice. Humans with loss of function mutations of *STAT1* have increased susceptibility to *Salmonella* and mycobacterial infections, whereas GOF mutations of *STAT1* cause chronic mucocutaneous candidiasis (CMC). Gene targeting of *Stat3* leads to early embryonic lethality, the lethality being related, in part, to interference with Lif function. Conditional knock-outs of *Stat3* in myeloid cells display exaggerated inflammatory responses as a result of failure of IL-10 signaling. STAT3 is also essential for Th17 cells, and mutations of *STAT3* underlie the primary immunodeficiency Hyper-IgE syndrome (HIES, or Job's syndrome) associated with impaired IL-17 production. Polymorphisms of *STAT3* are associated with IBD (Chapter 75). STAT4 is activated by IL-12. *Stat4*$^{-/-}$ mice develop normally but have defective Th1 differentiation and IFNγ production combined with augmented Th2 development.

STAT6 is activated by IL-4. *Stat6*$^{-/-}$ mice have defective Th2 development with defective IgE responses following infection with parasites. Lack of STAT6 dramatically attenuates allergic and asthmatic disease in animal models of these diseases. IL-13 also activates STAT6, and its responses are abrogated in *Stat6*$^{-/-}$ mice.

STAT5A and STAT5B are highly homologous, but nonetheless have different functions. *Stat5a*$^{-/-}$ mice have impaired mammary gland development and failure of lactation, whereas *Stat5b*$^{-/-}$ mice have defective sexually dimorphic growth and growth hormone–dependent regulation of liver gene expression. *Stat5a/5b* doubly deficient mice manifest increased perinatal lethality, decreased size, female infertility, and impaired lymphocyte development.

Stat5$^{-/-}$ mice develop lymphoproliferative disease reminiscent of IL-2- and IL-2R–deficient mice, related to loss of Tregs with expansion of Tfh and Th17 cells.

Patients with *STAT5B* mutations have short stature and immune dysregulation.

CLINICAL RELEVANCE

Type I and II Cytokine Receptors

- Mutations of genes encoding interleukin (IL)-7R, γc, and Janus kinase 3 (JAK3) cause severe combined immunodeficiency (SCID).
- TYK2 and signal transducer and activator of transcription 3 (STAT3) mutations cause hyper-IgE syndrome (HIES).
- STAT1 mutations cause autosomal dominant chronic mucocutaneous candidiasis (CMC) with increased susceptibility to mycobacterial and viral infections.
- Mutations of genes encoding IL-12, IL-12R, and interferon-γ receptor (IFN-γR) are associated with susceptibility to intracellular infections.
- Polymorphisms of IL-2R and IL-7R are associated with multiple sclerosis (MS).
- Polymorphisms of IL-23R are associated with inflammatory bowel disease (IBD).
- Polymorphisms of STAT4 are associated with rheumatoid arthritis (RA) and systemic lupus erythematosus (SLE).
- Erythropoietin (EPO), granulocyte–colony-stimulating factor (G-CSF), and thrombopoietin (TPO) are used to treat cytopenias.
- Anticytokine and/or cytokine receptor monoclonal antibodies (mAbs) are used to prevent transplant rejection and treat several autoimmune and inflammatory diseases.

Attenuation of Type-I and Type-II Cytokine Signaling

Perhaps as important as the triggers that initiate signal transduction are the mechanisms for extinguishing the response.[44] There are several families of proteins involved in downregulating cytokine signaling. Among these are phosphatases, cytokine-inducible inhibitor molecules, and transcriptional repressors. The phosphatase SHP-1 interacts with cytokine receptors and downregulates signaling. Mice with the naturally occurring "moth-eaten" mutation in *SHP-1* lack a functional SHP-1 protein and die at an early age as a result of autoimmune disease.

Another family of proteins that attenuate cytokine signaling is the suppressors of cytokine signaling (SOCS), which are alternatively termed JAB, SSI, and CSIS. These are SH2-containing proteins that bind to either cytokine receptors or to JAKs and inhibit signaling. There are at least eight members of this family. Largely as a result of systemic hyperresponsiveness to IFN-γ, *Socs-1*$^{-/-}$ mice die within a few weeks of birth. SOCS-2 has been recently shown to regulate Th2 differentiation and allergic responses. Another family member, SOCS-3, is important in controlling Th17 differentiation.[45-47]

KEY CONCEPTS

Properties of the Tumor Necrosis Factor (TNF) Receptor Superfamily

- Activation of a TNF receptor can lead to a wide range of effects, from proliferation to apoptosis.
- Transduction of signals through TNF receptor–associated factors (TRAFs) leads to the enhancement of survival.
- Signaling through death domains leads to the induction of apoptosis.

THE TNF CYTOKINE AND RECEPTOR SUPERFAMILY

This large family of structurally related ligands, receptors, and inhibitory decoy receptors has various roles both within and outside the immune system. The first two members of this family to be discovered were TNF and lymphotoxin-α (LTα; formerly named TNF-β). These molecules are secreted principally by activated myeloid and T cells. They have similar proinflammatory functions and belong to a large family of related molecules that includes CD30 ligand, CD40 ligand, FAS ligand, and TNF-related apoptosis inducing ligand (TRAIL). Indeed, the TNF cytokine family now contains 18 ligands and 29 receptors, each of which exhibits marked differences in tissue expression, ligand specificity, receptor binding, and biological function (Fig. 9.4) This section describes general aspects of TNF and TNF receptor (TNFR) biology, with examples from three of the best-studied TNF-family members. More complete listings of TNF-family cytokines and their receptors can be found in Table 9.2 and Table 9.3.

Ligand and Receptor Structure

Much of our understanding of the structural and functional characteristics of the TNF ligand and receptor superfamilies has come from the analysis of TNF (TNFSF2), LTα (TNFSF1), FASL (TNFSF6), and their receptors. Among these, TNF and LTα are closely related homotrimeric proteins (32% identity). Human TNF is synthesized as a 233-amino acid glycoprotein. It contains a long (76 residue) amino-terminal sequence that anchors it to the cell membrane as a 25-kilodalton (kDa) type II membrane protein. A secreted 17-kDa form of TNF is generated through

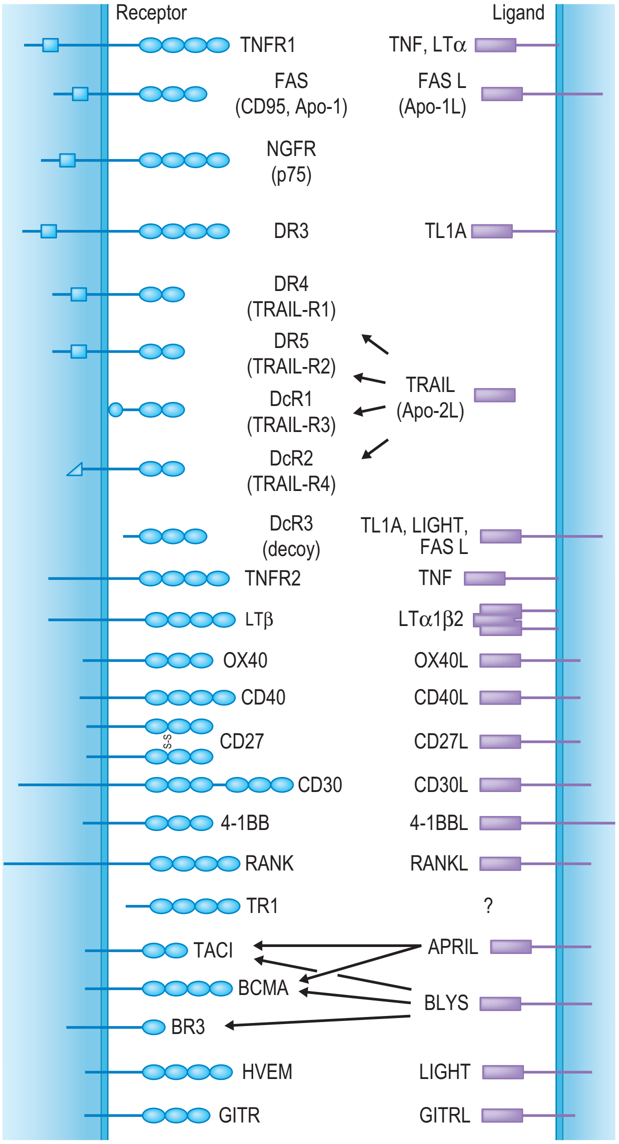


FIG 9.4 Schematic representation of members of the tumor necrosis factor (TNF) ligand and receptor superfamily.

the enzymatic cleavage of membrane-bound TNF by a metalloproteinase termed TNF-α–converting enzyme (TACE). Both soluble and membrane forms of TNF are thought to be homotrimers held together by noncovalent interactions through a trimerization domain. Both membrane-bound and soluble forms of TNF are biologically active. They have different affinities for the two TNF receptors and thus can exhibit different biological properties (see below).

LTα differs from TNF in that it is synthesized as a secreted glycoprotein. Human LTα is synthesized as a 205-amino acid glycoprotein that in native form exists as a 25-kDa homotrimer. It can bind both TNF receptors with affinities comparable to those of TNF and has similar biological effects. A membrane-bound form of LT has been identified that consists of a heterotrimeric complex containing one LTα subunit noncovalently linked to two molecules of an LTα-related type II membrane protein termed LTβ. The LT$\alpha_1\beta_2$ heterotrimer, also known as *mLT,* is not cleaved by TACE and is thought to exist exclusively as a membrane-bound complex. mLT does not bind either of the two TNF receptors, but rather exerts its effects on another member of the TNF receptor superfamily, the lymphotoxin β receptor (LTβR). TNF and the two LT subunits are encoded by closely linked single-copy genes situated in the MHC class III locus, at chromosome 6p21.3 in humans (Chapter 5).

The two receptors for TNF (and LTα) are type-I transmembrane glycoproteins. They are designated TNFR1 (TNFRSF1A) and TNFR2 (TNFRSF1B). (TNFR1 is also known as p60 in humans and p55 in mice, and TNFR2 is known as p80 in humans and p75 in mice.) These receptors are characterized by cysteine-rich repeats of about 40 amino acids in their aminoterminal extracellular domains. Each extracellular domain consists of three or four cysteine-rich regions containing 4–6 cysteines involved in intrachain disulfide bonds. The cytoplasmic domains of these receptors have no obvious similarity to any known kinase and are thought to lack intrinsic enzymatic activity. Signal transduction is, therefore, achieved by the recruitment and activation of adaptor proteins that recognize specific sequences in the cytoplasmic domains of these receptors. Recruitment of adaptor molecules activates a number of characteristic signaling pathways that can lead to a remarkably diverse set of cellular responses. These include differentiation, activation, release of inflammatory mediators, and apoptosis.

Family Members and Their Actions

Tumor Necrosis Factor, Lymphotoxin-α, and Receptors

TNF is a major physiological mediator of inflammation. It is one of the first cytokines made in response to TLR4 stimulation by bacterial (LPS) and other TLR ligands.[48] IFN-γ also induces TNF and augments its effects. TNF has been shown to induce fever, activate the coagulation system, induce hypoglycemia, depress cardiac contractility, reduce vascular resistance, induce cachexia, and activate the acute-phase response in the liver. Thus TNF is the major mediator of septic shock. TNF also upregulates MHC class I and class II expression, activates phagocytes, and induces mononuclear phagocytes to produce cytokines, such as IL-1, IL-6, chemokines, and TNF itself. Activation by TNF causes increased adhesion of cells to the endothelium and can be cytotoxic, particularly to tumor cells. TNF-deficient mice are resistant to septic shock induced by high doses of LPS but have increased susceptibility to bacterial infection. The dual role of TNF in controlling bacterial replication and in septic shock emphasizes the point that although the goal of an immune response is to eliminate invading microorganisms, the response itself can be injurious to normal host tissues. Septic shock is an extreme example of this. Although the primary source of TNF is the mononuclear phagocyte, it is also produced by T cells, NK cells, and mast cells. LTα shares many of the same biological effects as TNF, mainly because of its ability to bind the same receptors. However, LTβR plays a unique role in the development of secondary lymph nodes.

Fas Ligand and Its Receptor Fas/APO-1/CD95

Fas (Apo-1/CD95/TNFRSF6) is a type I integral membrane protein structurally related to TNFR1. Fas can trimerize and transduce proapoptotic signals upon binding of its ligand, FasL. Similar to TNF, the physiological ligand for Fas (CD95L or FasL)

Text continued on p. 143

TABLE 9.2 TNF Superfamily Cytokines

Symbol	Common Name	Aliases	Binds to Receptor(s)	OMIM ID	Key Functions	Phenotype Associated With Over Expression	Phenotype Associated With Deficiency	Human Genetic Disease Associations
TNFSF1	Lymphotoxin alpha (Ltα)	LT, TNFB, TNFSF1	TNFR2 (1B), TNFR1 (1A), HVEM (14)	153440	Lymphoid organ formation		Absence of LN and PP, defective GC formation	
TNFSF2	Tumor necrosis factor (TNF)	TNFα, TNFSF2, CACHECTIN	TNFR2 (1B), TNFR1 (1A)	191160	Inflammation	Wasting syndrome, arthritis	Defective GC formation, resistance to endotoxic shock and experimental arthritis	TNF2 (G-308A) promotor polymorphism associated with increased susceptibility to septic shock. Asthma and RA severity
TNFSF3	Lymphotoxin beta (LTβ)	p33, TNFC, TNFSF3	As a $\beta_2\alpha_1$ heterotrimer with LTα binds to LTβ receptor (3)	600978	Lymphoid organ formation	Ectopic lymphoid organ formation		
TNFSF4	OX40 Ligand	GP34, OX4OL, TXGP1, CD134L, OX-40 L	OX40 (4)	603594	CD4 T-cell expansion, survival, and Th2 development	Increased Th2 responses	Th2 deficiency, blockade improves EAE	Associated with SLE in GWAS
TNFSF5	CD40 Ligand	IGM, IMD3, TRAP, gp39, CD154, CD40L, HIGM1, T-BAM,	CD40 (5)	300386	Costimulation and differentiation of B cells and APC	Constitutive expression in B cells or keratinocytes leads to SLE-like syndrome	Immunodeficiency due to defective Ig class switching and germinal center formation	X-linked hyper-IGM syndrome associated with CD40L loss of function mutations
TNFSF6	Fas Ligand	FASL, CD178, CD95L, APT1LG1	Fas (6), DcR3 (6B)	134638	Mediator of CD4(+) T-cell apoptosis as a result of restimulation and apoptosis in other cell types		Lymphadenopathy and systemic autoimmunity	Autoimmune lymphoproliferative syndrome (ALPS) type Ib
TNFSF7	CD27 Ligand	CD70, CD27L, CD27LG	CD27 (7)	602840	T-cell costimulation	T-cell hyperactivation eventually leading to immunodeficiency (HIV-like)		
TNFSF8	CD30 Ligand	CD153, CD30L, CD30LG	CD30 (8)	603875				

Continued

TABLE 9.2 TNF Superfamily Cytokines—cont'd

Symbol	Common Name	Aliases	Binds to Receptor(s)	OMIM ID	Key Functions	Phenotype Associated With Over Expression	Phenotype Associated With Deficiency	Human Genetic Disease Associations
TNFSF9	4-1-BB Ligand	4-1BB-L	4-1BB (9)	606182	T-cell costimulation			
TNFSF10	TRAIL (TNF-like apoptosis inducing ligand)	TL2, APO2L, TRAIL, Apo-2 L	DR4 (10A), DR5 (10B), DcR1 (10 C), DcR2 (10D)	603598	Dendritic cell apoptosis, NK-cell mediated tumor cell killing		Defective NK-mediated tumor eradication	
TNFSF11	RANK-L	ODF, OPGL, sOdf, RANKL, TRANCE, hRANKL2	RANK (11A)	602642	Mediates osteoclast formation and bone remodeling. Stimulation of APC			
TNFSF12	TWEAK	APO3L	TWEAK-R (12A)	602695	Potential role in inflammation and lymphocyte function			
TNFSF13	APRIL	APRIL, TALL2, TWE-PRIL	TACI (13B), BCMA (17)	604472	Promotes T-independent type-2 responses through interactions with TACI	Overexpression in T cells produces prolonged T-cell survival and enhanced TI-2 responses		
TNFSF13B	BlyS, BAFF	BAFF, BLYS, TALL1, THANK, ZTNF4	TACI (3B), BAFF-R (13 C), BCMA (17)	603969	Promotes B-cell maturation, plasmablast survival	SLE-like systemic autoimmunity and arthritis		
TNFSF14	LIGHT	LTg, TR2, HVEML, LIGHT	HVEM (14), LT-βR (3), DcR3 (6B)	604520	CD8 T-cell and APC costimulation	inflammation, T-cell hyperactivation, Th1 bias	Defective CD8 T-cell costimulation	
TNFSF15	TL1A	TL1, VEGI	DR3 (25)	604052	Ligand for DR3 (TNFRSF25) on lymphocytes	T-cell activation, IL-13 dependent small intestinal hyperplasia and inflammation	Reduced immunopathology in T-cell-dependent autoimmune diseases	Common variant associated with inflammatory bowel disease through GWAS
TNFSF18	GITR Ligand	TL6, AITRL, GITRL, hGITRL	GITR (18)	603898	T-cell costimulation, $CD25^+$regulatory T cells			
ED1	Ectodermal dysplasia 1, anhidrotic (EDA1)	EDA, HED, EDA1, XHED, XLHED	EDAR	305100	Tooth, hair, and sweat gland formation			X-linked ectodermal dysplasia

TABLE 9.3 TNF Superfamily Receptors (Receptors in Italics Have a C-Terminal Death Domain)

Symbol	Common Name(s)	Aliases	Binds to Ligand(s)	OMIM ID	Key Functions	Phenotype Associated With Deficiency	Human Genetic Diseases
TNFRSF1A	Tumor necrosis factor receptor 1 (TNF-R1)	FPF, p55, p60, TBP1, TNF-R, TNFAR, TNFR1, p55-R, CD120a, TNFR55, TNFR60, TNF-R-I, TNF-R55, MGC19588	TNF-α (2), Ltα (1)	191190	Mediates TNF-induced inflammation (and apoptosis in some cells)	Resistance to TNF-induced arthritis models, resistant to endotoxic shock, increased sensitivity to bacterial pathogens	Periodic fever syndrome (TRAPS) associated with heterozygous extracellular mutations.
TNFRSF1B	Tumor necrosis factor receptor 2 (TNF-R2)	p75, TBPII, TNFBR, TNFR2, CD120b, TNFR80, TNF-R75, p75TNFR, TNF-R-II	TNF-α (2), Ltα (1)	191191	May enhance proapoptotic effect of TNFR1	Still susceptible to TNF-induced arthritis models, defective CD8T-cell apoptosis after restimulation, increased sensitivity to bacterial pathogens	
TNFRSF3	Lymphotoxin β receptor	LTBR, CD18, TNFCR, TNFR-RP, TNFRSF3, TNFR2-RP, LT-BETA-R, TNF-R-III	LIGHT (14), LTβ (3)	600979	Lymphoid organ formation	No LN, PP, defective GC formation	
TNFRSF4	OX40	OX40, ACT35, CD134, TXGP1L	OX40L (4)	600315	T-cell costimulation	Defective CD4 T-cell responses	
TNFRSF5	CD40	p50, Bp50, CD40, CDW40, MGC9013	CD40L (5)	109535	Costimulation and differentiation of B cells and antigen-presenting cells (APCs)	Defective immunoglobulin (Ig) class switching and germinal center formation	Autosomal hyper-IgM syndrome associated with loss-of function mutations Locus associated with rheumatoid arthritis (RA) through GWAS
TNFRSF6	*Fas, CD95*	FAS, APT1, CD95, APO-1	FasL (6)	134637	Apoptosis of restimulated CD4 T cells, B cells, ?others	Defective apoptosis of restimulated CD4 T cells	Autoimmune lymphoproliferative syndrome (ALPS) associated with heterozygous dominant-interfering mutations
TNFRSF6B	Decoy receptor 3	M68, TR6, DCR3, DJ583P15.1.1	FasL (6), TL1A (15), LIGHT (14)	603361	Soluble decoy receptor for FasL, LIGHT, and TL1A May have a role in tumor immune evasion		Associated with inflammatory bowel disease (IBD) in GWAS
TNFRSF7	CD27	T14, CD27, S152, Tp55, MGC20393	CD27L (7)	186711	T-cell costimulation	Defective T-cell responses	
TNFRSF8	CD30	CD30, KI-1, D1S166E	CD30L (8)	153243	?Inhibition of CD8 T-cell effector function		
TNFRSF9	4-1BB, CD137	ILA, 4-1BB, CD137, CDw137, MGC2172	4-1BBL (9)	602250	T-cell costimulation	Defective CD8 T-cell responses	Locus associated with ulcerative colitis in GWAS
TNFRSF10A	*Death Receptor 4 (DR4)*	DR4, APO2, MGC9365, TRAILR1, TRAILR-1	TRAIL (10)	603611	Mediator of dendritic cell and tumor cell apoptosis		
TNFRFRSF10B	*Death Receptor 5 (DR5)*	DR5, KILLER, TRICK2, TRICKB, ZTNFR9, TRAILR2, TRICK2A, TRICK2B, TRAIL-R2, KILLER/DR5	TRAIL (10)	603612	Mediator of dendritic cell (DC) and tumor cell apoptosis		
TNFRSF10C	Decoy receptor 1	LIT, DCR1, TRID, TRAILR3	TRAIL (10)	603613	GPI-linked decoy receptor, interferes with TRAIL function		

Continued

TABLE 9.3 TNF Superfamily Receptors (Receptors in Italics Have a C-Terminal Death Domain—cont'd

Symbol	Common Name(s)	Aliases	Binds to Ligand(s)	OMIM ID	Key Functions	Phenotype Associated With Deficiency	Human Genetic Diseases
TNFRSF10D	Decoy receptor 2	DCR2, TRUNDD, TRAILR4	TRAIL (10)	603614	Transmembrane decoy receptor, interferes with TRAIL function		
TNFRSF11A	(Receptor activator of NF-κ B) RANK	OFE, ODFR, PDB2, RANK, TRANCER	RANKL (11)	603499	Mediates DC costimulation and osteoclast maturation and activation	Osteopetrosis caused by osteoclast deficiency, no lymph nodes, impaired B-cell development	
TNFRSF11B	Osteoprotegerin (OPG)	OPG, TR1, OCIF, MGC29565	TRAIL (10), RANKL (11)	602643	Soluble decoy receptor for RANK	Osteoporosis, arterial calcification	
TNFRSF12A	TWEAK-receptor	FN14, TWEAKR	TWEAK (12)	605914			
TNFRSF13B	TACI	TACI	APRIL (13), BAFF (13B)	604907	May inhibit some of the pro-survival effects of BAFF-R	Decreased TI-2 B-cell responses, but B-cell hyperplasia and autoimmunity	Loss-of-function mutations associated with familial common variable immune deficiency (CVID)
TNFRSF13C	BAFF receptor (BAFF-R)	BAFF/BLyS receptor 3	BAFF (13B)	606269		Impaired survival of immature transitional B cells	
TNFRSF14	Herpes virus entry mediator (HVEM)	TR2, ATAR, HVEA, HVEM, LIGHTR	LIGHT (14), herpes viruses	602746			Locus associated with RA and ulcerative colitis in GWAS
TNFRSF16	NGF-R	TNFRSF16, p75(NTR)	NGF (not a TNF family member)	162010	Nerve growth factor (NGF) receptor - evolutionary outlier as NGF not classic TNF family molecule	Defective sensory neuron innervation; impaired heat sensitivity	
TNFRSF17	B-cell maturation antigen (BCMA)	BCM	APRIL (13), BAFF (13B)	109545		Apparently no B-cell phenotype	
TNFRSF18	Glucocorticoid-induced TNF receptor (GITR)	AITR, GITR, GITR-D	GITRL (18)	603905	T-cell costimulation, marker for CD4$^+$CD25$^+$ regulatory T cells (Tregs), modulates Treg function	T-cell hyperactivation	
TNFRSF19	Toxicity and JNK inducer (TAJ)	TROY, TRADE, TAJ-alpha		606122	Similar to ectodysplasin A receptor (EDAR), expressed in skin and brain		
TNFRSF19L	RELT	FLJ14993			Possible T-cell costimulator		
TNFRSF21	*Death Receptor 6 (DR6)*	BM-018		605732	Negative regulator B- and T-cell responses	Enhanced T- and B-cell activation	
TNFRSF25	*Death Receptor 3 (DR3)*	*DR3, TR3, DDR3, LARD, APO-3, TRAMP, WSL-1, WSL-LR*	*TL1A (15)*	*603366*		Impaired thymic negative selection. Reduced immunopathology and T-cell accumulation at site of autoimmune disease	
EDAR	*Ectodysplasin 1, anhidrotic receptor*	DL, ED3, ED5, ED1R, EDA3, EDA-A1R	E1	604095	Tooth, hair, sweat gland formation	Abnormal tooth, hair, and sweat gland formation	Ectodermal dysplasia
XEDAR	XEDAR: ectodysplasin A2 isoform receptor	EDAA2R, EDA-A2R	EDA-A2	300276			

LN, lymph node; PP, Peyer patch; GC, germinal center.
Locuslink ID: Gene "homepage" curated by NCBI. Go To and type the locuslink ID in the search window.OMIM: ID in the Online Mendelian Inheritance in Man Database. Go to and type in the OMIM ID in the search window.

is synthesized as a type II membrane protein and is expressed on activated B cells, T cells, and NK cells. Fas-induced apoptosis can play an essential role in the termination of T-cell responses, particularly in the peripheral immune system. Fas can also play a key role in the induction of cell death by cytotoxic T cells (CTLs) and NK cells (Chapter 17), where it functions in conjunction with perforin. Nonapoptotic functions of FasL include lymphocyte costimulation and T-cell differentiation into short-lived effector memory cells.[49,50]

CD40 Ligand and CD40

CD40 is expressed by a variety of cell types, including B cells, DCs, monocytes, macrophages, and endothelial cells. It plays a major costimulatory role in B-cell differentiation and recombination and promotes survival through the induction of BCL-2 family members. Studies of both CD40-deficient mice and patients with hyper-IgM syndrome (Chapter 34) reveal that its function extends beyond the humoral immune response, with CD40 signaling also playing a role in cell-mediated immunity. CD40 ligand (CD154) is a 39-kDa protein expressed by activated CD4 T cells that can bind to and activate CD40 by cell–cell contact.

CD40L on T cells triggers antigen-presenting cell (APC) activation, including upregulation of the CD28 ligands B7-1 and B7-2. This indirectly boosts costimulation of the T-cell response. Because of its critical role in mediating T-cell help for B-cell class switching and autoantibody formation, blocking CD40L/CD40 interactions has been a therapeutic goal in autoimmune diseases. Clinical trials of a blocking anti-CD40L antibody in SLE showed promising results but were halted because of thrombotic events, likely resulting from the off-target effects of these antibodies on CD40L expressed on platelets.[51] More recently, antagonistic antibodies against CD40 with possible applications to treat transplantation and autoimmune diseases have been developed.[52]

Other TNF-Family Cytokines

Other members of the TNFR family play various roles in the development and function of the immune system. OX-40, CD27, CD30, and 4-1BB can mediate costimulation of T-cell activation. The TNF family ligand BAFF (BlyS/TALL1/TNFSF13B) promotes B-cell maturation and antibody secretion and can bind three distinct receptors, TACI (TNFRSF13B), BADD-R (TNFRSF13C), and BCMA (TNFRSF17).

Signaling

The TNF receptor superfamily can be divided into three subfamilies on the basis of the types of intracellular signaling molecules recruited (*e.g.,* FAS-associated death domain [FADD], TNF receptor-associated death domain [TRADD], or TNF receptor–associated factor [TRAF]) (Fig. 9.5).[53] The cytoplasmic domains of several receptors, including death receptor 3 (DR3), DR4, DR5, TNFR1, and FAS contain a conserved ≈80-amino acid motif termed the death domain (DD). This element is required for recruitment of DD-containing adaptor molecules that are involved in the initiation of apoptotic cell death (Chapter 13). For this reason, these receptors have been termed "death receptors." The function of a number of death receptors can be regulated by decoy receptors. These are cell surface molecules that bind ligand but lack functional intracellular domains. Other TNF receptor superfamily receptors that lack death domains (*e.g.,* CD27, CD30, CD40, HVEM, TNFR2, LT-βR, OX-40, and 4-1BB) associate with different types of adapter molecules, most notably members of the TRAF family.

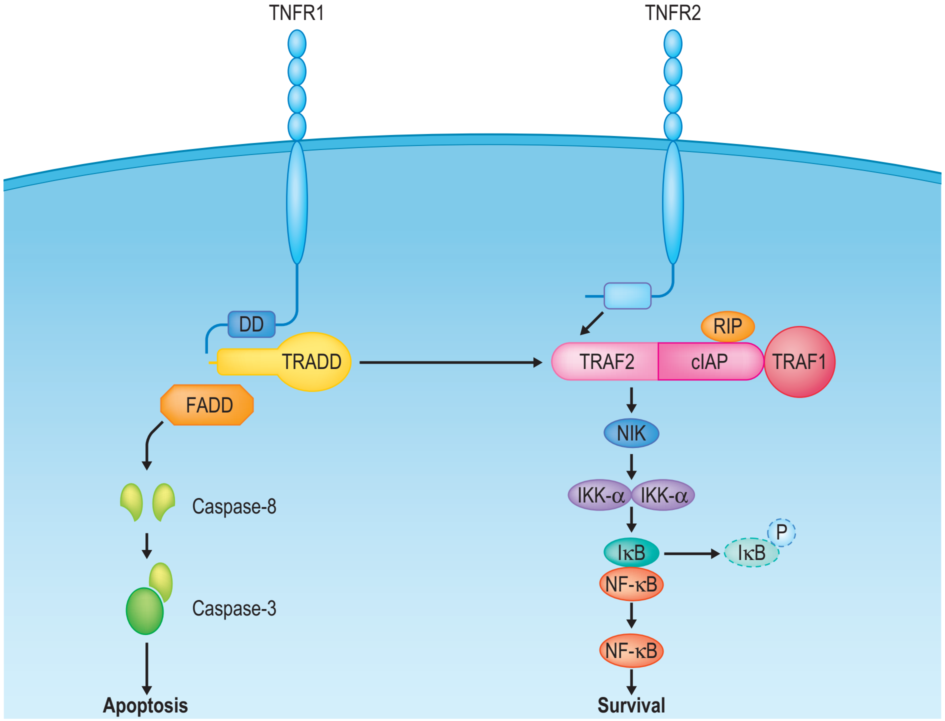


FIG 9.5 The role of the death domain (DD)– and death effector domain (DED)–containing molecules (e.g., Caspase-8) in signaling by tumor necrosis factor receptor 1 (TNFR1) and TNFR2.

Death Domains: TNF Receptor-Associated Death Domain and FAS-Associated Death Domain

The primary molecule transducing signals for TNFR1 is TRADD, which is directly recruited to TNFR1 after activation by TNF. This DD motif is found in both adaptor molecules, such as TRADD, and the cytoplasmic domains of the receptor itself (Fig. 9.6).

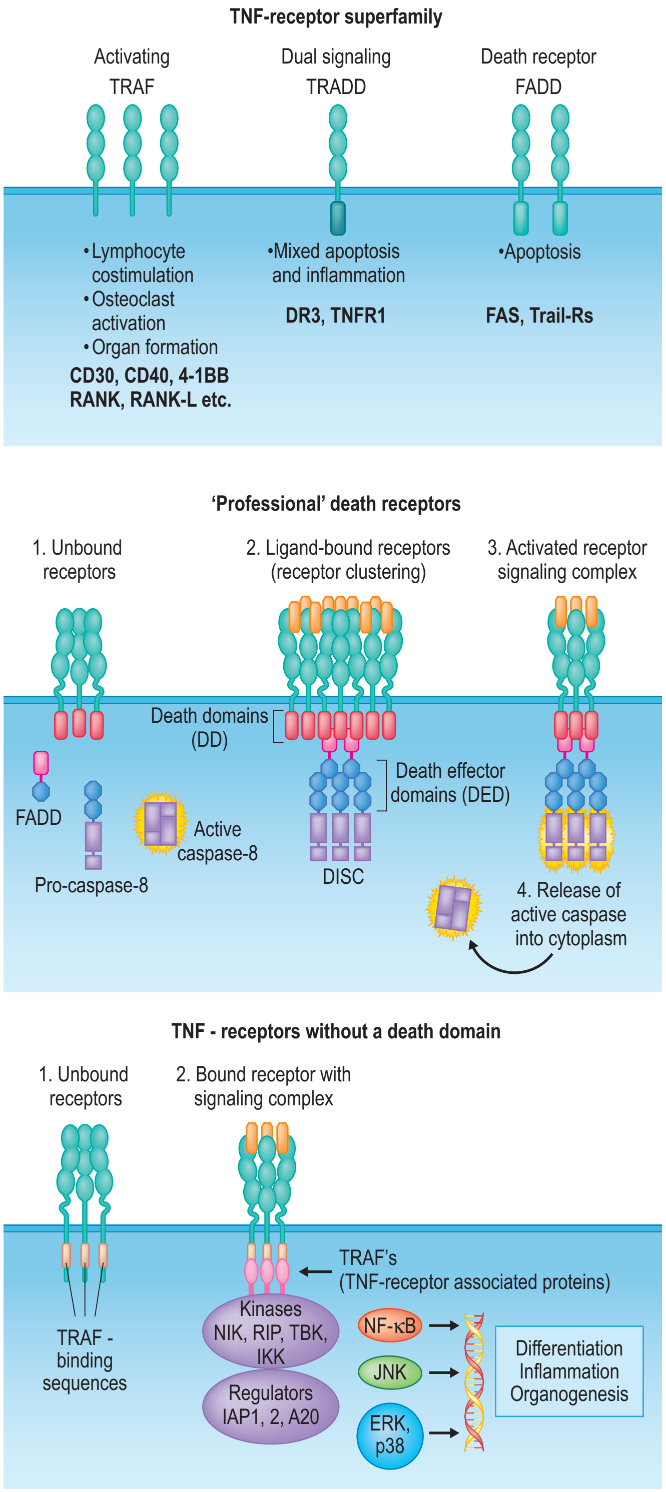

TNF-receptor superfamily
Activating
TRAF
Dual signaling
TRADD
Death receptor
FADD
•Lymphocyte costimulation
•Osteoclast activation
•Organ formation
CD30, CD40, 4-1BB RANK, RANK-L etc.
•Mixed apoptosis and inflammation
DR3, TNFR1
•Apoptosis
FAS, Trail-Rs
'Professional' death receptors
1. Unbound receptors
2. Ligand-bound receptors (receptor clustering)
3. Activated receptor signaling complex
Death domains (DD)
Death effector domains (DED)
FADD
Active caspase-8
Pro-caspase-8
DISC
4. Release of active caspase into cytoplasm
TNF - receptors without a death domain
1. Unbound receptors
2. Bound receptor with signaling complex
TRAF's (TNF-receptor associated proteins)
TRAF - binding sequences
Kinases NIK, RIP, TBK, IKK
Regulators IAP1, 2, A20
NF-κB
JNK
ERK, p38
Differentiation Inflammation Organogenesis

FIG 9.6 Signaling by tumor necrosis factor (TNF) family cytokines and their receptors.

The binding of TRADD to TNFR1 leads to the recruitment and activation of numerous associated signaling molecules. TNF-induced apoptosis is generally thought to be achieved by the interaction of TRADD with FADD (also known as MORT1), a ≈27-kDa protein that oligomerizes with TRADD through the DDs contained in both molecules. In turn, recruitment of FADD coordinately activates several members of the caspase family.[54]

Caspase-8, which is generally considered to be the apical caspase in the TNF and FAS pathways, is recruited to FADD in the activated complex and activated through increased local concentration. Cleaved caspase-8 can subsequently activate downstream caspases, notably caspase-3, which play a more proximal role in apoptosis. TRADD also has a TRAF-binding motif that leads to the recruitment of TRAF1 and -2 and the subsequent TRAF-dependent activation of proinflammatory signaling mediated by activation of NF-κB and mitogen-activated protein kinase (MAPK) pathways.

Although cell death in tumor cells can be induced by TNF, the most common result of TNFR1 ligation in primary immune cells is inflammation and sometimes protection from TNF-induced apoptosis. Recent evidence suggests that activation of proapoptotic and proinflammatory signaling by TNFR1 is not simultaneous but proceeds by sequential steps.

Unlike TNFR1, FAS can directly recruit FADD to its cytoplasmic DD, leading to the rapid formation of a death-inducing signaling complex (DISC), which contains FADD/MORT1 and caspase-8, thereby permitting activation of downstream caspases. FADD is recruited through interactions between charged residues in the DDs of FAS and FADD. Caspase-8 recruitment is accomplished through a structurally related domain termed the death-effector domain (DED). FADD DED contains two hydrophobic patches not present in the DD, which are vital for binding to the DEDs in the prodomain of caspase-8 and for apoptotic activity.[54]

Rather than DDs, the cytoplasmic domains of many receptors in the TNFR superfamily, including TNFR2, CD30, and CD40, contain short peptide consensus sequences that enable recruitment of a different family of adaptors, the TRAF proteins. A separate consensus sequence has been identified for TRAF6 versus other TRAF proteins, and other mechanisms probably operate to maintain the specificity of TRAF recruitment. Structural studies have revealed a mushroom-like structure for the TRAF proteins, with a trimer of the three TRAF subunits stabilized by a stalk-like coiled-coil domain.

TRAF proteins activate NF-κB and MAPK pathways through recruitment and activation of protein complexes that stimulate these signaling cascades. The exact mechanisms by which this occurs are not yet clear, but recent studies have called attention to the ability of TRAF proteins to catalyze ubiquitination of target signaling complexes, which can function as an activating step. TRAF6, which mediates NF-κB activation by a number of TNF-family receptors, associates with a protein complex that mediates K63-linked ubiquitination and activation of the inhibitor of κB kinase (IKK) complex, which consists of two catalytic subunits, IKKα and IKKβ, and a regulatory protein IKKγ or nuclear factor κB (NF-κB) essential modulator (NEMO).[55] Rather than causing degradation of IKK, K63-linked ubiquitination activates the kinase, leading to phosphorylation and degradation of IκB (inhibitor of NF-κB) and the release of active NF-κB subunits. Active NF-κB subunits translocate to the nucleus, where they regulate the expression of a wide variety of genes involved in the inflammatory response.

Some TNF-family receptors use other mechanisms to activate NF-κB. LTβ receptor activates the IKK complex via the serine–threonine kinase NIK, which was initially identified through its ability to associate with TRAF2. A naturally occurring mouse mutation termed *alymphoplasia* (*aly*) is the result of a point mutation of NIK. *Aly/aly* mice lack lymph nodes and Peyer patches and also exhibit disorganized splenic and thymic structures. This mutation, and the phenotype of LTβR knock-out mice, revealed the critical role of this receptor in normal lymph node development and the formation of "tertiary" lymphoid tissue in inflammation.

When a single TNF-family ligand, such as TNF, binds both a death receptor (TNFR1) and a non–death receptor (TNFR2), a number of mechanisms regulate receptor signaling and the cellular outcome. Rather than functioning in cell death, the physiological function of TNFR2 may be as a costimulator of lymphocyte proliferation.[56]

CLINICAL RELEVANCE

Tumor Necrosis Factor Receptor (TNFR) Superfamily Cytokines and Receptors and Disease

- Dominant mutations of in the gene encoding TNFR1 are associated with autosomal dominant periodic fever syndromes, known as TNFR1-associated periodic syndromes (TRAPS).
- Loss-of-function mutations in the gene encoding CD40L are associated with X-linked hyper-IgM syndrome (X-HIM).
- Dominant interfering mutations in *TNFRSF6,* encoding the Fas receptor are associated with autoimmune lymphoproliferative syndrome (ALPS).
- Rheumatoid arthritis (RA) often responds to therapeutic use of TNF antagonists.

Clinical Relevance

Mutations affecting the TNFR1 protein are associated with periodic fever syndromes (Chapter 60). Patients with the TNFR1-associated periodic syndrome (TRAPS) have missense mutations in exons encoding the extracellular regions of TNFR1 that affect normal TNFR1 function, prompting the designation. The mutated TNFR1 in TRAPS accumulates intracellularly and signals in a TNF-independent fashion to amplify inflammatory responses through the wild-type TNFR1.[57] Both TNF blockade with etanercept and IL-1 blockade have efficacy in reducing symptoms in TRAPS.[58]

The *in vivo* role of Fas signaling in the regulation of the immune system was confirmed when the naturally arising lpr and gld mouse strains were found to harbor homozygous mutations of Fas and Fas ligand, respectively. Both these mouse strains are characterized by lymphadenopathy and splenomegaly resulting from the accumulation of unusual $CD4^-CD8^-$ T cells, as well as the production of autoantibodies. Heterozygous dominant negative mutations in the gene encoding Fas cause the autoimmune lymphoproliferative syndrome (ALPS) (Chapter 35).

The gene encoding CD40 ligand is defective in X-linked hyper-IgM syndrome (X-HIM), a rare inherited disorder in which affected male children generate only IgM antibodies, many of which are autoantibodies (Chapter 34). Patients with X-HIM frequently suffer opportunistic infections, usually bacterial, and have an increased susceptibility to cancer. The physiological role of the B cell–activating factor (BAFF) receptor in mouse B-cell development is illustrated by BAFF-R mutations in the A/WySnJ mouse, which lacks peripheral B cells. TACI knock-out mice have hyperactive B cells, but in humans, dominant negative TACI mutations have been found in patients with common variable immunodeficiency affecting B-cell numbers and function (Chapter 34), which argues that in humans TACI serves as a positive modulator of B cells.

KEY CONCEPTS

Properties of the Interleukin-1 Receptor/Toll-Like Receptor (IL-1R/TLR) Family

- IL-1 plays a key role in fever and acute-phase responses.
- IL-18 augments T-helper 1 (Th1) cell differentiation.
- TLRs modulate proinflammatory signals in response to bacterial proteins.

INTERLEUKIN-1/TOLL-LIKE RECEPTOR FAMILY

Ligand and Receptor Structure

The IL-1/TLR family of receptors comprises at least 11 members, including the IL-1RI, IL-1RII, IL-1R-associated protein (IL-1RAcP), IL-18R, IL-18RAcP, IL-1Rrp2, IL-1RAPL, IL-33R(T1/ST2), TIGGIR, SIGGIR, and the mammalian TLRs.[59] The ligands for these receptors include IL-1, IL-18, IL-33, and IL-1 F5–IL-36 α, β, and γ; IL-36RA; IL-33 and IL-37; and IL-38.[60]

Family Members and Their Actions

Interleukin-1

There are two cell surface receptors for IL-1, type I (IL-1R1) and type II (IL-1R2). Both these bind ligand (see Fig. 9.4), but only IL-1R1 transduces signals. The extracellular domain of IL-1R1 has three Ig-like domains and a 200-amino acid cytoplasmic domain. Upon ligand binding, IL-1R1 associates with IL-1R accessory protein (IL-1RAcP), which is critical for the initiation of signaling. The IL-1R2 cytoplasmic domain is extremely short and has been suggested to be a "decoy" receptor, competing with IL-1RI for ligand binding and attenuating signaling. Both IL-1Rs are susceptible to proteolytic cleavage near the membrane surface. Thus they can be found as soluble proteins that can "buffer" IL-1 signaling. These soluble receptors are readily detectable in the circulation.

There are three members of the IL-1 gene family: two agonists, IL-1α and IL-1β, and one antagonist, IL-1 receptor antagonist (IL-1Ra). Both IL-1α and IL-1β are synthesized as precursor proteins. IL-1α and IL-1β are structurally similar and have similar actions, but they are regulated differently. IL-1α is processed by a calpain-like converting enzyme, but granzyme B and neutrophils and mast cell proteases can also cleave the precursor. The proform of IL-1α has biological activity, but cleavage results in increased biological activity. IL-1β is regulated at the levels of mRNA stabilization and translation, and it requires proteolytic activation. Pro–IL-1β remains in the cytoplasm until it is cleaved by proteases, such as caspase-1. It is then transported out of the cell.

The cleavage of IL-1β occurs in a multiprotein complex called the *inflammasome.* The key components of the inflammasome are caspase-1 and a recognition/assembly component nucleotide-binding oligomerization domain (NOD)–like receptor (NLR). Another protein called ASC (a simple adapter protein containing both pyrin and CARD domains) is required to facilitate assembly. NLR proteins are intracellular pattern-recognition receptors that contain three domains: *(i)* a segment with multiple leucine-rich repeats, whose role is to recognize the trigger for activation (whether directly or indirectly remains unclear); *(ii)* a portion called a NACHT domain, which leads to adenosine triphosphate (ATP)–dependent dimerization of the NLR after trigger recognition; and *(iii)* a protein–protein interaction domain, most commonly either a pyrin or a CARD domain, which recruits caspase-1.

The NLRP3 inflammasome is the best studied. Recognition of the trigger by NLRP3 causes its dimerization, recruitment of ASC via interaction of the pyrin domains of NLRP3 and ASC, and subsequent recruitment of caspase-1 via the CARD domains present in both ASC and caspase-1. Dimerization of caspase-1 upon inflammasome assembly allows for autoactivation by cleavage of the proform to generate active enzyme. The NLRP3 inflammasome can be activated by ATP, components of gram-positive bacterial cell walls, intracytoplasmic DNA, molecules resulting from tissue damage, uric acid crystals, alum, silica, asbestos, and amyloid-β among others. Cigarette smoke and cholesterol crystals also activate caspase-1.[61]

Similar to IL-1α, IL-1β can be cleaved and activated by other proteases, including neutrophil elastase, cathepsin, and proteinase-3 (PR-3). Mast cell proteases, as well as caspase-8, have also been shown to cleave IL-1β precursor. The mechanism of release of IL-1 and related family members from cells is somewhat mysterious because they lack a classic signal peptide and do not enter the secretory pathway. Although there is some evidence that IL-1β can be cleaved and released from cells without cell death, a large body of evidence suggests that IL-1 family proteins are released from cells upon lysis. Because of this, and the large numbers of cellular insults that result in release of IL-1 and related proteins and trigger inflammation, it has been proposed that they function as “alarmins” and act as sentinels of cellular damage.

The principal functions of IL-1 include the induction of acute-phase protein synthesis, cachexia, and fever. In fact, it was the first endogenous pyrogen to be identified. It induces the production of IL-6 and chemokines, promotes hematopoiesis, stimulates adhesion of vascular leukocytes to the endothelium, and has procoagulant effects. Importantly, IL-1 is a critical differentiation factor for Th17 cells, which underscores the role of this cytokine in inflammation and inflammatory diseases. Interestingly, it has recently been shown to activate ILC2, inducing proliferation and cytokine expression.[62] Unlike Fas and some of the other TNF-family cytokines, IL-1 does not directly induce cell death. Mononuclear phagocytes are the main, but not exclusive, source of IL-1. IL-1RI$^{-/-}$ and IL-1β$^{-/-}$ mice have blunted fever responses to some (but not all) stimuli. This indicates that despite the impressive actions of IL-1, it is evidently redundant to some extent in febrile responses.

Interleukin-18

A major action of IL-18, which is constitutively expressed by most cell types, is the induction of IFN-γ, and NK cells activation, a function it typically performs synergistically with IL-12. These functions are important for its antitumor activity. It has also been shown to promote angiogenesis and tumor progression.[63] IL-18 can induce IL-4 and IL-13 production, indicating a somewhat broader range of action. IL-18 precursor can be cleaved by the NRLP3 inflammasome as well as the NRLC4. IL-18–binding protein interacts with IL-18 and prevents association with IL-18R. IL-18R was first designated IL-1Rrp (IL-1R–related protein) before being recognized as the receptor for IL-18. The receptor is expressed predominantly on T cells, B cells, and NK cells. It associates with an accessory protein, IL-18RAcP.

Interleukin-33

IL-33 (previously known as IL-1 F11) appears to be a cytokine that acts by binding to a specific extracellular receptor, namely, the IL-1R –related protein ST2, also known as IL-33R. It also acts as a transcriptional repressor by associating with chromatin. Because of this dual effect and because of the expression of ST2 on different cell types, IL-33 acts on both immune and nonimmune cells. It acts on T and B cells, promoting Th2-associated cytokines, including IL-4, IL-5, and IL-13.[64] IL-33 acts on mast cells, promoting degranulation, an effect also shown on basophils and granulocytes in general, but also enhancing cell survival. Upon cleavage by mast cell proteases, IL-33 can also activate ILC2, inducing cytokines production and eosinophils recruitment.[65,66] Interestingly, IL-33–treated basophils have been shown to suppress arthritic inflammation. IL-33 can also act on endothelial and epithelial cells to induce angiogenesis and production of other cytokines and chemokines.

Interleukin-36

The three members of the IL-36 subfamily, IL-36α, β, and γ, are encoded by separate genes. They bind to a receptor composed of IL-36R and IL-1RacP. These cytokines are mostly expressed in skin. IL-36α transgenic mice exhibit skin inflammation, and these IL-36 cytokines can play a role in psoriasis.[67] IL-36 receptor antagonist (IL-36Rα, also known as IL-1 F5) acts as an antagonist for the three IL-36 family cytokines. *IL36RA* mutations are associated with generalized pustular psoriasis. Keratinocytes from patients with deficiency of IL-36 receptor antagonist (DITRA) have elevated levels of multiple inflammatory cytokines in lesional skin. They are highly expressed in skin and airway and may be involved in skin diseases, such as psoriasis (Chapter 64).

Interleukin-37

IL-37 (also known as IL-1 F7) exists in several splice variants of which IL-1F 7b has been the most studied. IL-37 binds to IL-18Rα chain, although with a lower affinity than IL-18. It uses TIR8 (also known as SIGIRR) as second chain of the receptor. Despite binding to the same receptor and its capacity to complex with IL-18Acp, IL-37 does not seem to act as a receptor antagonist for IL-18. IL-37 translocates to the nucleus and binds Smad3, enabling regulation of gene transcription. Cleavage of IL-37 appears to be dependent on caspase-1 and -4. Its antiinflammatory activity is dependent on ASC or NLRP3. IL-37 negatively regulates excessive inflammatory responses. Macrophages expressing IL-37 no longer secrete proinflammatory cytokines, and IL-37 transgenic mice are resistant to LPS-induced shock.[68]

Interleukin-38

IL-38 (IL-1 F10) has a 43% homology with IL-36Rα and, similarly, has an antiinflammatory activity. It is released by apoptotic cells to limit macrophage activation. Polymorphic IL-38 genes have been associated with increased susceptibility for autoimmune and autoinflammatory diseases (*e.g.,* RA, spondyloarthritis, and psoriatic arthritis).

Other Members of the Interleukin-1 Family

The remaining members of the IL-1 and IL-1R families are the receptor homologues APL and TIGIRR. These receptors have limited tissue distribution, with TIGIRR found almost exclusively in the brain and APL in the brain and a small number of other tissues.[69] Little is known of about the function of TIGIRR or

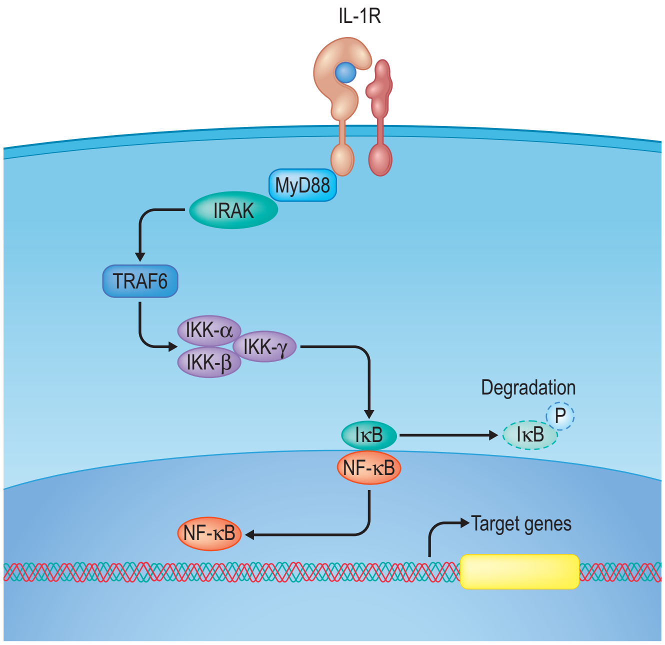


FIG 9.7 Mechanism of signal transduction by interleukin-1 receptor (IL-1R) and related receptors.

about APL in any organ other than the brain. Most insight into APL has come from the finding that deletions or mutations of this gene cause mental retardation.

Signaling

Ligand binding to IL-1R, IL-18R, and TLRs results in NF-κB activation (Fig. 9.7). These receptors all associate with the adapter protein MyD88 (Chapter 3). MyD88 has a C-terminal TIR domain and an N-terminal DD. MyD88 allows the recruitment of IL-1 receptor-associated kinase (IRAK), which also has an N-terminal DD. IRAK, in turn, permits the recruitment and activation of a member of the TRAF family, TRAF6. This leads to the activation of the serine kinases TAB2 and TAK1 (TGF-β–associated kinase-1), as well as that of the inhibitor of the κB kinases IKKα and IKKβ. With IKKγ or NEMO, these kinases phosphorylate IκB, which leads to its degradation within proteasomes, freeing bound NF-κB for nuclear translocation. Mice deficient in MyD88, IRAK, or TRAF6 have diminished responses to IL-1R/TLR family ligands. Other adapter molecules, including Mal and TRIF, are involved in TLR signaling.

CLINICAL RELEVANCE

Diseases Associated With the Interleukin-1 (IL-1) Family of Cytokines

- Mutations in genes involved in the formation of the inflammasome complexes result in increased secretion of IL-1 and other IL-1 family members.
- Mutations in the gene coding for IL-1 receptor antagonist cause a systemic autoinflammatory disease.

Clinical Relevance

The actions of IL-1 are tempered by the actions of a critical natural cytokine antagonist, IL-1Ra, which is encoded by the *IL1RN* gene. Mutations in *IL1RN* can cause a systemic autoinflammatory disease, which is denoted deficiency of IL-1Ra or DIRA.[70] Mutations in *NLRP3* contribute to several hereditary periodic fever syndromes (Chapter 60), including familial cold autoinflammatory syndrome (FCAS), Muckle-Wells syndrome, and neonatal-onset multisystem inflammatory disease (NOMID; in Europe called chronic infantile, neurological, cutaneous, and articular [CINCA] syndrome). NLRP3 is also called *cryopyrin;* thus these disorders have been collectively referred to as *cryopyrinopathies.*[71] Mutations in the alternative inflammasome component NRLC4 have also been showed to result in autoinflammation caused by excessive IL-18 production and, ultimately, macrophage activation syndrome.[72]

INTERLEUKIN-17 RECEPTORS

Although their precise functions are incompletely understood, IL-17 and related cytokines are major inducers of inflammation and serve to recruit inflammatory cells.

Ligand and Receptor Structure

The IL-17 receptor family is incompletely understood but comprises at least five receptors: IL-17AR and IL-17BR (IL-17RH1), IL-17RL (receptor like), IL-17RD, and IL-17RE, which are ubiquitously expressed.[73,74] These receptors have a single transmembrane domain and exceptionally large cytoplasmic tails. The ligands in the IL-17 family include IL-17A–F. The precise interactions between ligands and receptors have not been defined. A viral IL-17 homologue, designated HVS-13, is present in the *Herpesvirus saimiri* genome. Structurally, the IL-17 family contains cysteine knots. In this respect, these cytokines are related to other, better-known cytokines, such as nerve growth factor.

IL-17 (IL-17A) is located on human chromosome 6 (mouse chromosome 1) and is produced by activated CD4, γ/δ and CD8 T cells. Recent findings have suggested that CD4 T cells that preferentially produce IL-17 (Th17 cells) represent a distinct lineage of effector Th cells (Chapter 16). These cells can be differentiated via varied sets of stimuli that include IL-1β, IL-6, IL-21, and TGF-β. IL-23 appears to have a critical role in regulating the biological activity of these cells. IL-17F is located adjacent to IL-17. Although it seems to be regulated in a similar manner, it may be more widely expressed than IL-17. Less well studied are IL-17B, IL-17C, and IL-17D. All of these are thought to be expressed in a variety of nonhematopoietic tissues, although IL-17D is reported to be produced by CD4 T cells.

IL-17A and IL-17F evoke inflammation, largely by inducing the production of chemokines, G-CSF, and GM-CSF, and this leads to the subsequent recruitment of polymorphonuclear leukocytes. IL-17 also induces production of matrix metalloproteinase (MMP) by epithelial cells, which may be an important aspect of the proinflammatory effects. IL-17 family cytokines appear to be important in host defense against *Klebsiella pneumoniae* and *Mycobacterium tuberculosis.* Abundant data also point to the pathogenic roles of IL-17A in models of immune-mediated disease and in human autoimmune disorders. IL-17E, which is also known as IL-25, is produced by Th2 cells and mast cells. It evokes an inflammatory response characterized by overproduction of Th2 cytokines, mucus production, epithelial cell hyperplasia, and eosinophilia. This cytokine is essential for the elimination of helminthic parasites.[75]

Signaling

Engagement of the IL-17 receptor activates MAPKs, the PI3 kinase pathway and NF-κB. Signaling via IL-25 is reported to be dependent on the adapter molecule TRAF6. IL-17 acts synergistically with TNF.[76] The IL-17R associates with an adapter molecule, called *Act,* through the SEFIR domains, which are part of both the receptor and the adaptor.

CLINICAL RELEVANCE

Therapeutic Application of Interleukin-17 (IL-17) Blockade

Monoclonal antibodies (mAbs) neutralizing IL-17 or antagonizing the IL-17 receptor are effective in treating psoriasis and psoriatic arthritis.

Clinical Relevance

Many human diseases and animal models of autoimmune disease have been associated with increased levels of IL-17. The inflammatory effects of IL-23 and IL-17 also appear to be associated with malignant transformation. Targeting IL-17 ligands and receptors could thus be a useful strategy. However, IL-17A and IL-17F are essential for mucocutaneous immunity against *Candida albicans,* and IL-17 deficiency results in chronic mucocutaneous candidiasis.[77]

CYTOKINES ACTIVATING RECEPTOR TYROSINE KINASES

Ligand and Receptor Structure

Many growth factors, such as insulin and epidermal growth factor, utilize receptor tyrosine kinases (RTKs). Some, but not all, of these factors can be classified as cytokines. These include CSF-1 (colony-stimulating factor 1 or M-CSF), stem cell factor (SCF, c-KIT ligand, or steel factor), platelet-derived growth factor (PDGF), and FMS-like tyrosine kinase 3 ligand (FLT3-L). All of these have important hematological effects and tend to be included in discussions of cytokines. The structure of SCF and CSF-1 is similar to that of the cytokines that bind type I receptors, as they too form four α-helical bundles, even though their receptors are entirely distinct. The similarities in their three-dimensional structures point to a common evolutionary ancestor. The receptors in this subfamily typically have five Ig-like loops in their ligand-binding extracellular domains. The cytoplasmic domain contains a tyrosine kinase catalytic domain interrupted by an "insert region" that does not share homology with other tyrosine kinases. This segment is used to recruit various signaling molecules.

Family Members and Their Actions

Stem Cell Factor

Bone marrow stromal cells can synthesize SCF (c-KIT ligand or Steel factor) as either a secreted or a transmembrane protein. SCF is required to make stem cells responsive to other CSFs. SCF is widely expressed during embryogenesis and is also detectable in blood circulation in normal adults. It has effects on germ cells, melanocytes, and hematopoietic precursors and has important effects on the differentiation of mast cells as well. Naturally occurring mouse mutations of SCF (Steel) or its receptor (W) have been recognized for many years. These mice have defects in hematopoiesis and fertility, lack mast cells, and have absent coat pigmentation.

Colony-Stimulating Factor 1

CSF-1, also known as monocyte–macrophage–CSF or macrophage-CSF (M-CSF), is a hematopoietic growth factor that supports the survival and differentiation of monocytic cells. It is produced by a wide variety of cells, including monocytes, smooth muscle cells, endothelial cells, and fibroblasts. M-CSF–deficient mice manifest monocytopenia and osteopetrosis. IL-34 is a new cytokine that binds to the CSF-1 receptor.

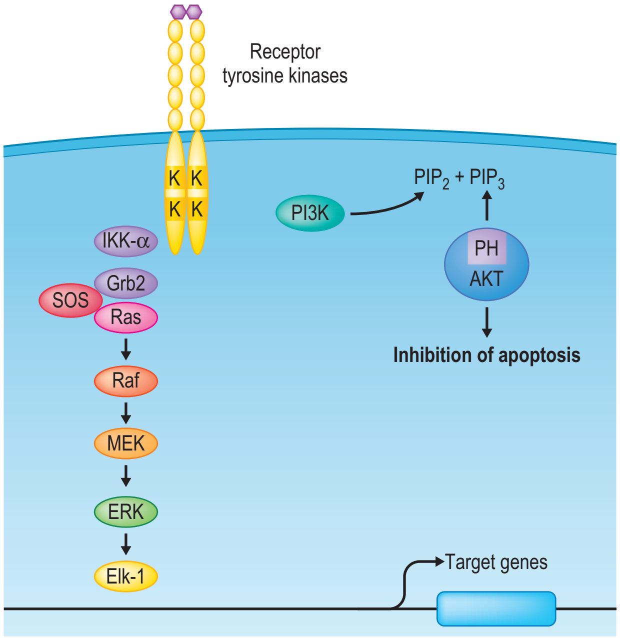


FIG 9.8 Mechanism of signal transduction by receptor tyrosine kinases.

FMS-like tyrosine kinase 3 ligand (FLT3-L) synergizes with other cytokines, including SCF, in inducing proliferation of hematopoietic precursors. FLT3-L is an important regulator of DCs.

Signaling

The first step in signaling by the RTKs is ligand-induced receptor dimerization (Fig. 9.8). Dimerization brings the two kinase domains into proximity and results in the activation of phosphotransferase activity. This leads to autophosphorylation of the receptor subunits on the tyrosine residues, which are then bound by a variety of signaling molecules, initiating signal transduction. During this important step, the signaling and adapter molecules recognize phosphotyrosine residues on the RTKs by virtue of either their SH2 domains or their PTB domains.

A major pathway activated by RTKs is the RAS/RAF/ERK pathway. RAS is a small G protein with intrinsic low guanosine diphosphate/guanosine triphosphate (GDP/GTP) exchange activity. In RTK signal transduction, two adapter proteins, GRB2 and SHC, have important functions in RAS activation. In some cases, GRB2 binds directly to phosphotyrosine residues on the cytoplasmic tail of the receptor via its SH2 domain. Alternatively, SHC can bind first and then recruit GRB2. In addition to an SH2 domain, GRB2 has two SH3 domains that bind proline-rich segments of the guanine nucleotide exchange factor son of sevenless (SOS) protein, recruiting it to the membrane and allowing it to activate RAS. Activated RAS binds and activates the serine/threonine kinase RAF, which, in turn, phosphorylates the dual-specificity kinase MEK. Activated MEK phosphorylates and activates ERK (extracellular signal–regulated kinase), which then translocates to the nucleus, where it phosphorylates and modulates the activity of various transcription factors, including ELK-1. Accordingly, mutations of RAS that lead to the constitutive activation of the ERK pathway have been found in a wide variety of human cancers.

Another important pathway activated by RTKs is the phosphatidylinositol 3′-OH kinase (PI3-kinase) pathway.[78] PI3-kinase catalyzes the formation of phosphatidylinositol-3,4,5-trisphosphate (PtdIns[3,4,5]P3) and PtdIns(3,4)P2. These phospholipids are recognized by proteins with pleckstrin homology (PH) domains. One such protein is protein kinase B (PKB or AKT), which has been implicated in the regulation of apoptosis. The PI3-kinase pathway is inhibited by the lipid phosphatase PTEN, which dephosphorylates PI3-kinase–generated phosphatidylinositides. Deletion of PTEN has been found in numerous tumor types, demonstrating a role for this protein as a tumor suppressor.

GOF mutations of c-kit result in a disorder termed *systemic mastocytosis* (Chapter 23). Mutations resulting in a fusion between the *PDGFRA* and *FIP1L1* genes underlie hypereosinophilic syndrome.[79,80]

KEY CONCEPT

Properties of the Transforming Growth Factor-β (TGF-β) Receptor Family

- TGF-β receptors play a key role in lymphoid homeostasis, with pro- and antiinflammatory actions.
- TGF-β promotes the differentiation of regulatory T cells and T-helper 17 (Th17) cells.
- TGF-β receptors transduce signals through SMAD proteins.
- TGF-β receptor function is dysregulated in many forms of human cancer.

TRANSFORMING GROWTH FACTOR-β LIGAND AND RECEPTOR FAMILIES

TGF-βs are a family of over 40 cytokines that inhibit cellular proliferation and can induce apoptosis of a variety of cell types. TGF-βs are involved in a number of biological processes, including tissue remodeling, wound repair, development, and hematopoiesis. Mutations of the elements in this pathway also contribute to malignant transformation. The mammalian ligands that belong to this family include TGF-β_1, -β_2, and -β_3, bone morphogenic proteins (BMPs), activins, inhibins, and müllerian-inhibiting substance. Despite their name, TGF-βs inhibit the growth of many other cells and, in combination with other cytokines and growth factors, may induce growth instead. TGF-βs also induce collagen and fibronectin production by fibroblasts, which are thought to be responsible, at least in part, for diseases characterized by fibrosis (*e.g.,* systemic sclerosis and pulmonary fibrosis). Functionally, TGF-β inhibits many aspects of lymphocyte function, including T-cell proliferation and CTL maturation. Transgenic mouse studies have been performed by using dominant negative forms of the TGF-β receptor. These mice exhibit massive expansion of lymphoid organs and develop T-cell lymphoproliferative disorders.[46]

Ligand and Receptor Structure

TGF-βs are expressed as biologically inactive disulfide-linked dimers that are cleaved to form active dimers. On translocation into the endoplasmic reticulum, the N-terminal leader peptide is cleaved, and the mature protein is subsequently generated by a second cleavage event that releases an N-terminal proregion. The proregion can remain associated with the biologically active C-terminal region, inhibiting its activity.

The biological effects of TGF-βs and their related ligands are mediated by two classes of receptor, designated type I (RI) and type II (RII).[81] A third group of receptors, denoted type III, also exists (*e.g.,* TGF-β RIII in the case of TGF-β). This latter group does not actively participate in signal transduction but is thought to function to present ligands to the functional receptors. Similar to the RTKs described previously, the cytoplasmic domains of TGF-β receptors possess intrinsic kinase activity. However, TGF-βRI and TGF-βRII encode serine/threonine kinases. The signaling cascade appears to be initiated by the binding of TGF-β to the type II receptor, inducing the assembly of a ternary complex containing TGF-β, TGF-βRII, and TGF-βRI.

TGF-β Family Members and Their Actions

The three known human TGF-βs—TGF-β_1, TGF-β_2, and TGF-β_3—all have expressed molecular weights of 25 kDa. These three isoforms of TGF-β are closely related and have very similar biological functions. TGF-β_1, the most abundant form, is the only isoform found in platelets. T cells and monocytes mainly synthesize TGF-β_1, a critical function of which is to antagonize lymphocyte responses.

Approximately half of TGF-$\beta_1^{-/-}$ mice survive until birth, and 3–4 weeks later, they typically succumb to an overwhelming autoimmune state characterized by lymphoid and mononuclear infiltration of heart, lung, and other tissues, and by autoantibody production. These studies, along with selective inhibition of TGF-β function in T cells, indicate that TGF-β plays a crucial role in T-cell homeostasis and the prevention of spontaneous T-cell differentiation. TGF-β_1 induces FoxP3, promotes adaptive Treg-cell differentiation, and inhibits IFN-γ production. Conversely, TGF-β_1 with IL-6 induces IL-17, a proinflammatory cytokine. Thus TGF-β_1 has both proinflammatory and antiinflammatory activities.

TGF-β_2 is the most abundant TGF-β isoform in body fluids, whereas TGF-β_3 is the least abundant of the three. TGF-β_2– and TGF-β_3–null mice exhibit defects distinct from those observed in TGF-β knock-out mice, particularly in bone and internal organ formation. Their deficiency is embryonically lethal, demonstrating that although the three isoforms functional similarly *in vitro,* they play distinct roles *in vivo.*

The human type II receptor is an 80-kDa glycoprotein that is the principal receptor for TGF-β. Upon binding of its ligand to RII, the type I receptor is recruited into the complex and activated through phosphorylation of its GS domain. The principal type I receptor in the TGF-β pathway is the ≈55-kDa activin-like kinase-5 (ALK-5). ALK-1 can also be recruited into the complex and can transduce TGF-β–mediated signals.

Signaling

Ligand binding to the type II receptor allows the recruitment of the type I receptor[81] (Fig. 9.9). Although the receptor subunits have some affinity for one another, the complex of the receptor subunits bound to ligand is more stable. The type II receptor is thought to be a constitutively active kinase. Upon ligand binding, the type II receptor phosphorylates the type I receptor. The type I receptor is structurally distinct from the type II receptor in having a glycine- and serine-rich sequence juxtamembrane domain that precedes the kinase domain, which is referred to as the *GS domain.* It is this site that is phosphorylated by the

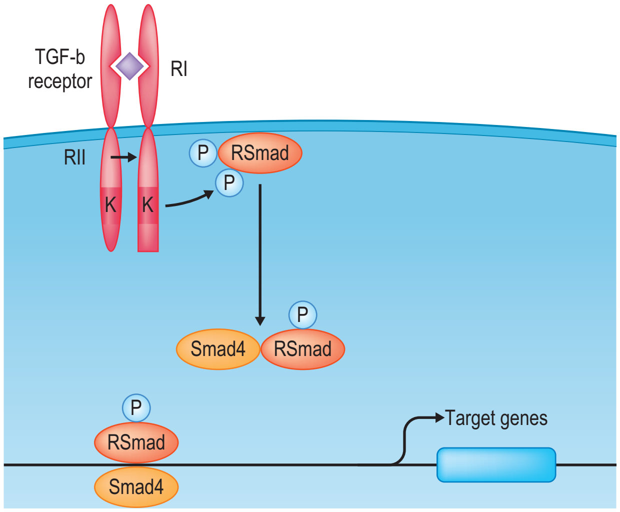


FIG 9.9 SMADs and signaling by transforming growth factor-β (TGF-β) superfamily receptor serine kinases.

type II receptor. In turn, the type I receptor is responsible for phosphorylating key signaling intermediates. It is not clear whether activation of the type I receptor results from enhancement of its kinase activity, the appearance of substrate-binding sites, or a combination of the two.

SMADs

The primary substrates activated by the type I receptors are SMADs, a group of related proteins that have been highly conserved throughout evolution and play a critical role in TGF-β signal transduction.[82] Eight mammalian SMADs have been identified. All exhibit a high degree of specificity for conserved motifs in the cytoplasmic tail of type I receptors. These proteins do not contain any previously known structural or enzymatic motifs. However, they have two homology domains, termed the MH1 and MH2 domains, at the N- and C-termini, respectively, which are separated by a central linker domain. The extreme C-terminus of some SMADs is a critical site of phosphorylation, as described below. SMADs have been subdivided into three classes based on functional distinctions.

R-SMADs. The receptor-regulated SMADs (R-SMADs) directly interact with, and are phosphorylated by, the type I receptor. These include SMAD1, -2, -3, -5, and -8. SMAD2 and SMAD3 are phosphorylated in response to TGF-β, whereas SMAD1, -5, and -8 are primarily activated in response to BMP activation. R-SMADs bind to the GS domain of type I receptor and are phosphorylated at a conserved SSXS motif in their C terminus. The interaction of R-SMADs with TGF-β receptors can also be regulated by another molecule termed SARA (SMAD anchor for receptor activation). SARA binds unphosphorylated SMAD2 and SMAD3. By virtue of its FYVE domain, SARA can bind phospholipids and localize SMADs to the plasma membrane, facilitating receptor binding. Phosphorylation of SMADs also permits dissociation from SARA.

Once phosphorylated, R-SMADs dissociate from the activated type I receptor and associate with SMAD4 in the cytoplasm. This is followed by nuclear translocation of the heteromeric SMAD complex, the binding to cognate DNA motifs in the promoters of TGF-β–responsive genes, and the concomitant induction of transcription. The SMAD MH1 motif mediates sequence-specific DNA binding, whereas the MH2 domain contains the transcriptional activation domain.

SMAD2-null mice lack anterior/posterior specification and fail to develop mesoderm. SMAD3-deficient mice have limb malformations and defective immune function. SMAD3-null mice exhibit defective TGF-β responses. However, these mice also display accelerated wound healing compared with normal mice, which seems to contradict the role of SMAD3 as a positive regulator of TGF-β in enhancing wound healing.

Upon recruitment to their cognate activated type I receptor, R-SMADs are phosphorylated on C-terminal serine residues, triggering homodimerization of R-SMADs or heterodimerization with another class of SMAD, the common SMAD or C-SMAD (see below).

C-SMADs. SMAD4 is the only known C-SMAD in vertebrates. It is thought to function as the central and essential downstream mediator of other SMADs in all TGF-β/BMP pathways. The SMAD MH2 domain is important for both receptor interaction and SMAD dimerization. SMAD4 lacks the SSXS element conserved in the R-SMADs and thus is not phosphorylated, rendering it unable to bind type I receptors directly. SMAD4-deficient mice die early in embryogenesis and exhibit severe defects in gastrulation.

I-SMADs. The third subfamily of SMADs is the inhibitory SMADs, or I-SMADs. In mammals, SMAD6 and SMAD7 are the I-SMADs. SMAD6 utilizes an alternative mechanism and is thought to function by competing with SMAD4 for binding to an R-SMAD. Studies with SMAD6-deficient mice indicate a role in the development and homeostasis of the cardiovascular system. SMAD7 is induced by TGF-β and binds to TGF-β receptors inhibiting the phosphorylation of R-SMADs, thus serving as a classic feedback inhibitor.

Consistent with their key functions in development, gene targeting of SMAD molecules has been shown to produce severe phenotypic abnormalities.

SMAD pathways. Although a canonical SMAD DNA-binding element (SBE) has been described (AGAC), comparison of the TGF-β responsiveness of synthetic promoters to natural promoters has revealed that SMADs can only partially account for the gene-regulatory effects of TGF-β signaling. SMADs also interact with a variety of other transcription factors, transcriptional coactivators, and transcriptional corepressors to regulate transcription of a select subset of complex promoters in a coordinated manner. For example, FAST-1 (forkhead activin signal transducer), a winged helix forkhead transcription factor, associates with the SMAD2/SMAD4 complex. SMADs also bind to c-JUN/c-FOS, and AP-1 sites frequently overlap with SBE sites in the naturally occurring promoters of TGF-β–responsive genes. SMADs can also bind ATF2, the vitamin D receptor, and other transcription factors and can recruit the coactivators CBP/p300. Additionally, the SKI and SNON corepressors can interact with SMADs and antagonize TGF-β signaling.

Other TGF-β–Activated Pathways

Other signaling pathways are known to mediate TGF-β signals. In particular, a number of members of the MAPK family are activated in response to TGF-β. TGF-β induces ERK, JNK, and p38 MAPK activity. Inhibition of several of these pathways inhibits TGF-β–mediated transcriptional activation. In addition, a MAPK family member, TAK1, has been implicated in TGF-β signaling. TAK1 is activated in response to ligand binding and has been shown to associate with another molecule, TAB1 (TAK1-associated binding protein), which activates TAK1 kinase activity. Together, they have been reported to activate the MAPK pathway(s), leading to activation of p38/MPK2 and c-JUN N-terminal kinase (JNK), with evidence that this may take place via MKK6 and MKK4, respectively.

Clinical Relevance

Although the relevance of TGF-β to clinical immunology remains unclear, defects in the TGF-β pathway have been identified in a range of human cancers. SMAD4 is deleted in half of all human pancreatic carcinomas. Mutations in SMAD2 have been identified in patients with colon cancer, and somatic mutations in TGF-β receptors have been identified in colon and gastric cancers. Loss of SMAD3 is associated with leukemia. In addition, oncogenic RAS has been shown to repress SMAD signaling by negatively regulating SMAD2 and SMAD3.

OTHER CYTOKINES

Interleukin-14

Several new cytokines have been identified, but their functions are less clear compared with those discussed above. IL-14 was identified as a high-molecular-weight B-cell growth factor produced by T cells and some B-cell tumors. The precise nature of this putative cytokine is still unclear.

Interleukin-16

IL-16 was formerly termed *lymphocyte chemoattractant factor* because of its ability to recruit CD4 T cells.[83] It is unrelated to other cytokines, and its only known receptor is the CD4 molecule. It was originally identified as a product of CD8 T cells, but its message is widely expressed. CD4 T cells, eosinophils, and mast cells can all secrete IL-16. It is present in bronchoalveolar lavage (BAL) fluids from patients with asthma and those with sarcoidosis. It has also been detected in blister fluid from bullous pemphigoid lesions. The physiological function of IL-16 has yet to be clarified.

Interleukin-32

IL-32 is one of the newest cytokines, and it, too, is structurally distinct.[84] It is inducible by the combination of IL-12 and IL-18. IL-32 induces the expression of various cytokines, including TNF, IL-1, IL-6, and chemokines. It can synergize with muramyl dipeptides. IL-32 signals via NF-κB and p38. IL-32 is present in the rheumatoid synovium, and injection of IL-32 induces inflammation and recruitment of inflammatory cells.[85]

THERAPEUTIC TARGETING OF CYTOKINES AND CYTOKINE RECEPTORS

Because of their important activities in numerous diseases outlined in this chapter, cytokines and their receptors have been actively studied as therapeutic targets or as therapeutic agents (Table 9.4; Chapter 89). IFNs were the first cytokines to be considered as antiviral agents. Currently, type I IFNs are used to treat a wide array of disease from viral hepatitis (IFN-α and IFN-λ) to MS (IFN-β). IL-2 is currently marketed (as aldesleukin) for the treatment of renal cancer and melanoma. TNF was investigated after its discovery as an antitumor agent, but its inflammatory properties made it a highly successful target for the treatment of autoimmune and inflammatory diseases.

With the advent of molecular biology techniques and mAb engineering, targeting of these molecules has become a highly successful strategy to modulate the immune response (Chapter 89). IL-2 is not only a therapeutic protein but also a target. Daclizumab and basiliximab are mAbs directed against the IL-2Rα chain and have been approved for use in transplantation, MS, uveitis, and ulcerative colitis. IL-6, one of the most important proinflammatory cytokines, has also been targeted and several anti–IL-6 antibodies are now under development. Tocilizumab, which targets the IL-6R, has been approved for the treatment of several autoimmune diseases ranging from RA to ankylosing spondylitis.

The effective targeting of the p40 subunit of IL-12/IL-23 with the human mAb ustekinumab has resulted in effective treatment of some forms of psoriasis, psoriatic arthritis, and Crohn disease. However, ustekinumab is not effective in MS.

Many studies have implicated TNF in the pathogenesis of the chronic inflammatory diseases, such as RA and Crohn disease. Levels of TNF are highly elevated in the synovial fluid and the serum of patients with RA, as well as in the gastrointestinal mucosa of patients with Crohn disease. This suggests that the proinflammatory effects of TNF might underpin the severe inflammatory symptoms observed in these diseases. Of particular relevance to RA, TNF both inhibits the synthesis of cartilage components, such as proteoglycan, and stimulates bone resorption. These findings prompted the development of specific TNF inhibitors that have demonstrated great promise in the treatment of RA and Crohn disease. The two best-characterized of these inhibitors are monoclonal anti-TNF antibodies (infliximab, adalimumab, and golimumab) and chimeric TNFR2-Fc protein (etanercept). Clinical studies of these compounds have demonstrated that they can induce striking improvement in patients with RA (see Table 9.4). The side effects of TNF antagonism include increased incidence of infection, particularly with *M. tuberculosis,* and a possible increased incidence of cancer.[86] Other TNF-family cytokines have been targeted for therapeutic purposes. Targeting of RANKL with denosumab has been used for the treatment of osteoporosis as well as the bone resorption and hypercalcemia that results from bone metastases in cancer.

Because of the importance of IL-1 in the pathogenesis of fever, it would be expected that agents that inhibit IL-1 would be therapeutically useful. Indeed, IL-1Ra (anakinra), the naturally produced antagonist, has been studied in a variety of settings. It has been found to be somewhat efficacious in the treatment of RA. Anakinra, canakinumab, an anti–IL-1 antibody, as well as rilonacept, a fusion protein containing the extracellular portions of the two IL-1R components IL-1R and AcP, have been more successful in autoinflammatory disorders, particularly those related to overactivity of inflammasome components, such as the cryopyrinopathies. Canakinumab has also been successful in the treatment of more common autoinflammatory conditions, such as juvenile-onset idiopathic arthritis, as well as in the treatment of gout and type II diabetes.[87-89] Similarly, the human mAb canakinumab has entered the clinical arena with overlapping applications, and others are in various stages of clinical development. One of the most exciting new antireceptor antibodies is belimumab, which targets BAFF, a B-cell stimulatory TNF family member. Belimumab is the first licensed drug for the treatment for SLE in more than 50 years, and its mechanism of action is believed to be interference with B-cell survival, differentiation, and autoantibody formation.[90]

Beyond their beneficial effects to patients, the clinical effect of specific cytokine blockade with these agents is the ultimate

TABLE 9.4 Recombinant Cytokines and Biological Agents Currently in Clinical Use or Being Tested for Clinical Application

Name	Target	Phase	Indications
Daclizumab	Interleukin (IL)-2Rα	FDA approved	Transplantation, multiple sclerosis
Basiliximab	IL-2Rα	FDA approved	Transplantation uveitis, ulcerative colitis,
Benralizumab	IL-5R	II	Asthma, chronic obstructive pulmonary disease (COPD)
Reslizumab	IL-5R	II/III	Asthma, inflammation of skin and gastrointestinal tract
Mepolizumab	IL-5	FDA approved	Asthma, nasal polyp
Tocilizumab	IL-6R	FDA approved	Rheumatoid arthritis (RA), juvenile idiopathic arthritis (JIA), Castleman disease, ankylosing spondylitis
Clazakizumab	IL-6	II	RA, psoriatic arthritis
Siltuximab	IL-6	FDA Approved	Castleman disease
Sarilumab	IL-6R	FDA approved	RA Uveitis
Olokizumab	IL-6	III	RA Crohn disease
Sirukinumab	IL-6	III	RA, systemic lupus erythematosus (SLE)
ALX-0061	IL-6		RA, SLE
Oprelvekin	IL-11R	II	Thrombocytopenia, von Willebrand disease
Ustekinumab	IL-12/23 p40	FDA approved	Psoriasis, psoriatic arthritis, Crohn disease, sarcoidosis, palmoplantar pustulosis
Briakinumab	IL-12/23 p40	III	Psoriasis, Crohn disease
Lebrikizumab	IL-13	II	Asthma
Tralokimumab	IL-13	II	Asthma, COPD
Anrukinzumab	IL-13	II	Ulcerative colitis
TNX-650	IL-13	I/II	Hodgkin lymphoma
Brodalumab	IL-17	FDA approved	RA, psoriasis
Ikzekizumab	IL-17	FDA approved	RA, psoriatic arthritis
Bimekizumab	IL-17	II	Psoriatic arthritis, rheumatoid arthritis
BCD-085	IL-17	II	Ankylosing spondylitis; plaque psoriasis
Secukinumab	IL-17	FDA approved	Psoriasis, psoriatic arthritis, ankylosing spondylitis, RA, Crohn disease, multiple sclerosis (MS), xerophthalmia
NNC-114-0005	IL-21	II	Type I diabetes, SLE
Tildrakizumab	IL-23	III	Plaque psoriasis
Risankizumab	IL-23	III	Plaque psoriasis
Guselkumab	IL-23	II	Plaque psoriasis
Anakinra	IL-1	FDA approved	Autoinflammatory syndromes, RA
IL-Trap/Rilonacpet	IL-1	FDA approved	Autoinflammatory syndromes, RA, gout, JIA
Canakinumab	IL-1β	FDA approved	Autoinflammatory syndromes, RA
MABp1	IL-1α	I/II/III	Cancer, cachexia, atherosclerosis, plaque psoriasis, acne vulgaris, type II diabetes
Gevokizumab	IL-1β	III	RA, type I and type II diabetes, giant cell arteritis, Behçet disease, uveitis
NI-0501	Interferon (IFN)-γ	II/III	Hemophagocytic lymphohistiocytosis
Anifrolumab	IFN-β	III	SLE
Etanercept	TNFα	FDA approved	RA, juvenile-onset rheumatoid arthritis (JRA), psoriatic arthritis, plaque psoriasis, ankylosing spondylitis
Infliximab	TNFα	FDA approved	RA, psoriasis, Crohn disease, ankylosing spondylitis, psoriatic arthritis, ulcerative colitis
Adalimumab	TNFα	FDA approved	RA, plaque psoriasis, Crohn disease, ankylosing spondylitis, psoriatic arthritis, ulcerative colitis, JIA
Golimumab	TNFα	FDA approved	RA, ankylosing spondylitis, psoriatic arthritis
Ozoralizumab	TNFα	II	RA
Pegsunercept	TNFα	FDA approved	RA, JRA, psoriatic arthritis, plaque psoriasis, ankylosing spondylitis
Certolizumab	TNFα	FDA approved	RA, JRA, psoriatic arthritis, plaque psoriasis, ankylosing spondylitis
Belimumab	BAFF/BLyS	FDA approved	SLE
Brentuximab vedotin	TNFRSF8 (CD30)	FDA approved	Lymphoma
Denosumab	RANKL	FDA approved	Osteoporosis, RA, hypercalcemia of malignancy, cancer

FDA, US Food and Drug Administration.
Source: Pipeline (Drug pipeline information database by Citeline, Inc.)

test of the pathogenic role of the targeted cytokine. Bringing cytokine blockade into the clinics has helped unravel the complex web of interactions between cytokines in human autoimmune and inflammatory conditions.

ON THE HORIZON

- Although the number of cytokines already seems vast, it is likely that more will be discovered in the future.
- As continued improvement in the imaging of cells and macromolecular complexes improves, it is likely that we will obtain new insights into how cytokines actually signal.
- Improved understanding of regulators like the suppressors of cytokine signaling (SOCS) proteins should offer new therapeutic opportunities.
- Additional therapeutic options for the future may include the use synthetic cytokines and the creation of bispecific antibodies to target more than one cytokine or cytokine receptor at the same time.

CONCLUSIONS AND SUMMARY

Cytokines encompass a wide range of molecules that are essential for communication between cells of the immune system and other nonimmune cells. Although the number of cytokines already seems vast, it is likely that more will be discovered in the future. Considerable progress has been made in defining the *in vivo* functions of various cytokines. Equally impressive have been advances in our understanding how dysregulation of cytokines and cytokine signaling contribute to human disease. Cytokine and anticytokine therapies are being successfully used in the clinic. It is likely that their use will increase with advances in the understanding of the immunobiology of these cytokines.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Dumonde DC, Wolstencroft RA, Panayi GS, et al. "Lymphokines": non-antibody mediators of cellular immunity generated by lymphocyte activation. Nature 1969;224:38–42.
2. Thompson AT, Lotze MT. The cytokine handbook. San Diego, CA: Academic Press; 2003.
3. Cohen S, Bigazzi PE, Yoshida T. Commentary. Similarities of T cell function in cell-mediated immunity and antibody production. Cell Immunol 1974;12:150–9.
4. Vilcek J. The cytokines: an overview. In: Thompson AW, Lotze M, editors. The cytokine handbook. San Diego, CA: Academic Press; 2003. p. 3–18.
5. Schwartz DM, Bonelli M, Gadina M, et al. Type I/II cytokines, JAKs, and new strategies for treating autoimmune diseases. Nat Rev Rheumatol 2016;12:25–36.
6. Boulay JL, O'Shea JJ, Paul WE. Molecular phylogeny within type I cytokines and their cognate receptors. Immunity 2003;19:159–63.
7. Taga T, Kishimoto T. Gp130 and the interleukin-6 family of cytokines. Annu Rev Immunol 1997;15:797–819.
8. Rose-John S. IL-6 trans-signaling via the soluble IL-6 receptor: importance for the pro-inflammatory activities of IL-6. Int J Biol Sci 2012;8:1237–47.
9. Hunter CA. New IL-12-family members: IL-23 and IL-27, cytokines with divergent functions. Nat Rev Immunol 2005;5:521–31.
10. Croxford AL, Lanzinger M, Hartmann FJ, et al. The cytokine GM-CSF Drives the inflammatory signature of CCR2+ Monocytes and licenses autoimmunity. Immunity 2015;43:502–14.
11. Codarri L, Gyulveszi G, Tosevski V, et al. RORgammat drives production of the cytokine GM-CSF in helper T cells, which is essential for the effector phase of autoimmune neuroinflammation. Nat Immunol 2011;12:560–7.
12. Hamilton JA. GM-CSF as a target in inflammatory/autoimmune disease: current evidence and future therapeutic potential. Expert Rev Clin Immunol 2015;11:457–65.
13. Kovanen PE, Leonard WJ. Cytokines and immunodeficiency diseases: critical roles of the gamma(c)-dependent cytokines interleukins 2, 4, 7, 9, 15, and 21, and their signaling pathways. Immunol Rev 2004;202:67–83.
14. Malek TR, Bayer AL. Tolerance, not immunity, crucially depends on IL-2. Nat Rev Immunol 2004;4:665–74.
15. Sakaguchi S, Sakaguchi N. Regulatory T cells in immunologic self-tolerance and autoimmune disease. Int Rev Immunol 2005;24:211–26.
16. Laurence A, Tato CM, Davidson TS, et al. Interleukin-2 signaling via STAT5 constrains T helper 17 cell generation. Immunity 2007;26:371–81.
17. Johnston RJ, Choi YS, Diamond JA, et al. STAT5 is a potent negative regulator of TFH cell differentiation. J Exp Med 2012;209:243–50.
18. Lutterotti A, Martin R. Getting specific: monoclonal antibodies in multiple sclerosis. Lancet Neurol 2008;7:538–47.
19. Nelms K, Keegan AD, Zamorano J, et al. The IL-4 receptor: signaling mechanisms and biologic functions. Annu Rev Immunol 1999;17:701–38.
20. Risma KA, Wang N, Andrews RP, et al. V75R576 IL-4 receptor alpha is associated with allergic asthma and enhanced IL-4 receptor function. J Immunol 2002;169:1604–10.
21. Mosmann TR, Coffman RL. TH1 and TH2 cells: different patterns of lymphokine secretion lead to different functional properties. Annu Rev Immunol 1989;7:145–73.
22. Lee GR, Kim ST, Spilianakis CG, et al. T helper cell differentiation: regulation by cis elements and epigenetics. Immunity 2006;24:369–79.
23. Thaci D, Simpson EL, Beck LA, et al. Efficacy and safety of dupilumab in adults with moderate-to-severe atopic dermatitis inadequately controlled by topical treatments: a randomised, placebo-controlled, dose-ranging phase 2b trial. Lancet 2016;387:40–52.
24. Waldmann TA, Dubois S, Tagaya Y. Contrasting roles of IL-2 and IL-15 in the life and death of lymphocytes: implications for immunotherapy. Immunity 2001;14:105–10.
25. Crotty S. Follicular helper CD4 T cells (TFH). Annu Rev Immunol 2011;29:621–63.
26. Chapuis AG, Lee SM, Thompson JA, et al. Combined IL-21-primed polyclonal CTL plus CTLA4 blockade controls refractory metastatic melanoma in a patient. J Exp Med 2016;213:1133–9.
27. He H, Wisner P, Yang G, et al. Combined IL-21 and low-dose IL-2 therapy induces anti-tumor immunity and long-term curative effects in a murine melanoma tumor model. J Transl Med 2006;4:24.
28. Trinchieri G, Pflanz S, Kastelein RA. The IL-12 family of heterodimeric cytokines: new players in the regulation of T cell responses. Immunity 2003;19:641–4.
29. Watford WT, Hissong BD, Bream JH, et al. Signaling by IL-12 and IL-23 and the immunoregulatory roles of STAT4. Immunol Rev 2004;202:139–56.
30. Smale ST, Fisher AG. Chromatin structure and gene regulation in the immune system. Annu Rev Immunol 2002;20:427–62.
31. Brightling CE, Chanez P, Leigh R, et al. Efficacy and safety of tralokinumab in patients with severe uncontrolled asthma: a randomised, double-blind, placebo-controlled, phase 2b trial. Lancet Respir Med 2015;3:692–701.
32. Corren J, Lemanske RF, Hanania NA, et al. Lebrikizumab treatment in adults with asthma. N Engl J Med 2011;365:1088–98.
33. Dillon SR, Sprecher C, Hammond A, et al. Interleukin 31, a cytokine produced by activated T cells, induces dermatitis in mice. Nat Immunol 2004;5:752–60.
34. Perrigoue JG, Zaph C, Guild K, et al. IL-31-IL-31R interactions limit the magnitude of Th2 cytokine-dependent immunity and inflammation following intestinal helminth infection. J Immunol 2009;182:6088–94.
35. Watanabe N, Hanabuchi S, Soumelis V, et al. Human thymic stromal lymphopoietin promotes dendritic cell-mediated CD4+ T cell homeostatic expansion. Nat Immunol 2004;5:426–34.

36. Siracusa MC, Saenz SA, Hill DA, et al. TSLP promotes interleukin-3-independent basophil haematopoiesis and type 2 inflammation. Nature 2011;477:229–33.
37. Garcia-Sastre A, Biron CA. Type 1 interferons and the virus-host relationship: a lesson in detente. Science 2006;312:879–82.
38. Lu HT, Riley JL, Babcock GT, et al. Interferon (IFN) beta acts downstream of IFN-gamma-induced class II transactivator messenger RNA accumulation to block major histocompatibility complex class II gene expression and requires the 48-kD DNA-binding protein, ISGF3-gamma. J Exp Med 1995;182:1517–25.
39. Asselin-Paturel C, Boonstra A, Dalod M, et al. Mouse type I IFN-producing cells are immature APCs with plasmacytoid morphology. Nat Immunol 2001;2:1144–50.
40. Siegal FP, Kadowaki N, Shodell M, et al. The nature of the principal type 1 interferon-producing cells in human blood. Science 1999;284:1835–7.
41. Bach EA, Aguet M, Schreiber RD. The IFN gamma receptor: a paradigm for cytokine receptor signaling. Annu Rev Immunol 1997;15:563–91.
42. Donnelly RP, Sheikh F, Kotenko SV, et al. The expanded family of class II cytokines that share the IL-10 receptor-2 (IL-10R2) chain. J Leukoc Biol 2004;76:314–21.
43. Pesu M, Candotti F, Husa M, et al. Jak3, severe combined immunodeficiency, and a new class of immunosuppressive drugs. Immunol Rev 2005;203:127–42.
44. Wormald S, Hilton DJ. Inhibitors of cytokine signal transduction. J Biol Chem 2004;279:821–4.
45. Chen Z, Laurence A, Kanno Y, et al. Selective regulatory function of Socs3 in the formation of IL-17-secreting T cells. Proc Natl Acad Sci USA 2006;103:8137–42.
46. Ghoreschi K, Laurence A, Yang XP, et al. T helper 17 cell heterogeneity and pathogenicity in autoimmune disease. Trends Immunol 2011;32:395–401.
47. Knosp CA, Carroll HP, Elliott J, et al. SOCS2 regulates T helper type 2 differentiation and the generation of type 2 allergic responses. J Exp Med 2011;208:1523–31.
48. Beutler BA. The role of tumor necrosis factor in health and disease. J Rheumatol Suppl 1999;57:16–21.
49. Klebanoff CA, Scott CD, Leonardi AJ, et al. Memory T cell-driven differentiation of naive cells impairs adoptive immunotherapy. J Clin Invest 2016;126:318–34.
50. Park SM, Schickel R, Peter ME. Nonapoptotic functions of FADD-binding death receptors and their signaling molecules. Curr Opin Cell Biol 2005;17:610–16.
51. Sidiropoulos PI, Boumpas DT. Lessons learned from anti-CD40L treatment in systemic lupus erythematosus patients. Lupus 2004;13:391–7.
52. Cordoba F, Wieczorek G, Audet M, et al. A novel, blocking, Fc-silent anti-CD40 monoclonal antibody prolongs nonhuman primate renal allograft survival in the absence of B cell depletion. Am J Transplant 2015;15:2825–36.
53. Siegel RM, Muppidi J, Roberts M, et al. Death receptor signaling and autoimmunity. Immunol Res 2003;27:499–512.
54. Siegel RM. Caspases at the crossroads of immune-cell life and death. Nat Rev Immunol 2006;6:308–17.
55. Deng L, Wang C, Spencer E, et al. Activation of the IkappaB kinase complex by TRAF6 requires a dimeric ubiquitin-conjugating enzyme complex and a unique polyubiquitin chain. Cell 2000;103:351–61.
56. Kim EY, Priatel JJ, Teh SJ, et al. TNF receptor type 2 (p75) functions as a costimulator for antigen-driven T cell responses in vivo. J Immunol 2006;176:1026–35.
57. Simon A, Park H, Maddipati R, et al. Concerted action of wild-type and mutant TNF receptors enhances inflammation in TNF receptor 1-associated periodic fever syndrome. Proc Natl Acad Sci USA 2010;107:9801–6.
58. Bulua AC, Mogul DB, Aksentijevich I, et al. Efficacy of etanercept in the tumor necrosis factor receptor-associated periodic syndrome: a prospective, open-label, dose-escalation study. Arthritis Rheum 2012;64:908–13.
59. Dunne A, O'Neill LA. The interleukin-1 receptor/Toll-like receptor superfamily: signal transduction during inflammation and host defense. Sci STKE 2003;2003:re3.
60. Dinarello C, Arend W, Sims J, et al. IL-1 family nomenclature. Nat Immunol 2010;11:973.
61. Martinon F, Mayor A, Tschopp J. The inflammasomes: guardians of the body. Annu Rev Immunol 2009;27:229–65.
62. Ohne Y, Silver JS, Thompson-Snipes L, et al. IL-1 is a critical regulator of group 2 innate lymphoid cell function and plasticity. Nat Immunol 2016;17:646–55.
63. Fabbi M, Carbotti G, Ferrini S. Context-dependent role of IL-18 in cancer biology and counter-regulation by IL-18BP. J Leukoc Biol 2015;97:665–75.
64. Schmitz J, Owyang A, Oldham E, et al. IL-33, an interleukin-1-like cytokine that signals via the IL-1 receptor-related protein ST2 and induces T helper type 2-associated cytokines. Immunity 2005;23:479–90.
65. Halim TY, Krauss RH, Sun AC, et al. Lung natural helper cells are a critical source of Th2 cell-type cytokines in protease allergen-induced airway inflammation. Immunity 2012;36:451–63.
66. Lefrancais E, Duval A, Mirey E, et al. Central domain of IL-33 is cleaved by mast cell proteases for potent activation of group-2 innate lymphoid cells. Proc Natl Acad Sci USA 2014;111:15502–7.
67. Johnston A, Xing X, Guzman AM, et al. IL-1F5, -F6, -F8, and -F9: a novel IL-1 family signaling system that is active in psoriasis and promotes keratinocyte antimicrobial peptide expression. J Immunol 2011;186:2613–22.
68. Sims J, Towne J, Blumberg H. 11 IL-1 family members in inflammatory skin disease. Ernst Schering Res Found Workshop 2006;(56):187–91.
69. Carrie A, Jun L, Bienvenu T, et al. A new member of the IL-1 receptor family highly expressed in hippocampus and involved in X-linked mental retardation. Nat Genet 1999;23:25–31.
70. Aksentijevich I, Masters SL, Ferguson PJ, et al. An autoinflammatory disease with deficiency of the interleukin-1-receptor antagonist. N Engl J Med 2009;360:2426–37.
71. Neven B, Prieur AM, Quartier dit Maire P. Cryopyrinopathies: update on pathogenesis and treatment. Nat Clin Pract Rheumatol 2008;4:481–9.
72. Canna SW, de Jesus AA, Gouni S, et al. An activating NLRC4 inflammasome mutation causes autoinflammation with recurrent macrophage activation syndrome. Nat Genet 2014;46:1140–6.
73. McKenzie BS, Kastelein RA, Cua DJ. Understanding the IL-23-IL-17 immune pathway. Trends Immunol 2006;27:17–23.
74. Kawaguchi M, Adachi M, Oda N, et al. IL-17 cytokine family. J Allergy Clin Immunol 2004;114:1265–73, quiz 1274.
75. Fort MM, Cheung J, Yen D, et al. IL-25 induces IL-4, IL-5, and IL-13 and Th2-associated pathologies in vivo. Immunity 2001;15:985–95.
76. Maezawa Y, Nakajima H, Suzuki K, et al. Involvement of TNF receptor-associated factor 6 in IL-25 receptor signaling. J Immunol 2006;176:1013–18.
77. Puel A, Cypowyj S, Bustamante J, et al. Chronic mucocutaneous candidiasis in humans with inborn errors of interleukin-17 immunity. Science 2011;332:65–8.
78. Cantley LC. The phosphoinositide 3-kinase pathway. Science 2002;296:1655–7.
79. Longley BJ, Reguera MJ, Ma Y. Classes of c-KIT activating mutations: proposed mechanisms of action and implications for disease classification and therapy. Leuk Res 2001;25:571–6.
80. Cools J, DeAngelo DJ, Gotlib J, et al. A tyrosine kinase created by fusion of the PDGFRA and FIP1L1 genes as a therapeutic target of imatinib in idiopathic hypereosinophilic syndrome. N Engl J Med 2003;348:1201–14.
81. Heldin CH, Miyazono K, ten Dijke P. TGF-beta signalling from cell membrane to nucleus through SMAD proteins. Nature 1997;390:465–71.
82. Massague J, Gomis RR. The logic of TGFbeta signaling. FEBS Lett 2006;580:2811–20.
83. Wilson KC, Center DM, Cruikshank WW. The effect of interleukin-16 and its precursor on T lymphocyte activation and growth. Growth Factors 2004;22:97–104.

84. Heinhuis B, Netea MG, van den Berg WB, et al. Interleukin-32: a predominantly intracellular proinflammatory mediator that controls cell activation and cell death. Cytokine 2012;60:321–7.
85. Kim S. Interleukin-32 in inflammatory autoimmune diseases. Immune Netw 2014;14:123–7.
86. Bongartz T, Sutton AJ, Sweeting MJ, et al. Anti-TNF antibody therapy in rheumatoid arthritis and the risk of serious infections and malignancies: systematic review and meta-analysis of rare harmful effects in randomized controlled trials. JAMA 2006;295:2275–85.
87. Ruperto N, Brunner HI, Quartier P, et al. Two randomized trials of canakinumab in systemic juvenile idiopathic arthritis. N Engl J Med 2012;367:2396–406.
88. Tran AP, Edelman J. Interleukin-1 inhibition by anakinra in refractory chronic tophaceous gout. Int J Rheum Dis 2011;14:e33–7.
89. Goldbach-Mansky R. Blocking interleukin-1 in rheumatic diseases. Ann N Y Acad Sci 2009;1182:111–23.
90. Hahn BH. Belimumab for systemic lupus erythematosus. N Engl J Med 2013;368:1528–35.

MULTIPLE-CHOICE QUESTIONS

1. The phenomenon of shared biological activities (cytokine redundancy) between cytokines that belong to the same subfamily reflects which of the following biological properties of type 1 family cytokines?
   - **A.** The cytoplasmic portions of the type I family receptors all contain membrane proximal domains that bind Janus kinases (JAKs).
   - **B.** Type I family cytokines act on all cells to inhibit viral replication, as well as cellular proliferation.
   - **C.** Heterodimeric type I cytokine receptors often use a shared receptor submit in conjunction with a ligand specific subunit.
   - **D.** There can be more than one receptor for each cytokine.

2. Which of the following cytokines promotes allergic responses and inhibits cell-mediated immune responses?
   - **A.** Interleukin-1 (IL-1)
   - **B.** IL-2
   - **C.** IL-4
   - **D.** IL-17

3. Which of the following cytokines is essential for mucocutaneous immunity against *Candida albicans*?
   - **A.** Transforming growth factor-β (TGFβ)
   - **B.** IL-3
   - **C.** IL-6
   - **D.** IL-17

4. Blockade of which of the following cytokines ameliorates rheumatoid arthritis (RA) and endotoxic shock?
   - **A.** IL-7
   - **B.** IL-12
   - **C.** IL-21
   - **D.** Tumor necrosis factor (TNF)

10

Chemokines and Chemokine Receptors

Philip M. Murphy

INTRODUCTION

Chemokines form a large family composed primarily of small secreted cytokine proteins that coordinate leukocyte trafficking by binding to seven-transmembrane (7TM) domain receptors. Chemokines mediate normal host defense and tissue repair; however, they may also support pathological immune responses, including chronic inflammation, autoimmunity, and cancer. The chemokine system is also a major target for immune system evasion or exploitation by pathogens (e.g., human immunodeficiency virus [HIV] and Plasmodium vivax). Increasingly, additional immunological and nonimmunological chemokine functions are being recognized. The nonimmunological functions can be beneficial, as in embryogenesis, or harmful, as in cancer. WHIM (warts, hypogammaglobulinemia, infections, and myelokathexis) syndrome is the only known mendelian condition caused by mutations in a chemokine or chemokine receptor (the receptor CXCR4). Two chemokine receptor antagonists have been approved by the U.S. Food and Drug Administration (FDA) so far: maraviroc, a CCR5 antagonist, used in the treatment of human immunodeficiency virus/acquired immunodeficiency syndrome (HIV/AIDS), and plerixafor, a CXCR4 antagonist, approved for hematopoietic stem cell mobilization for transplantation in cancer. This chapter will expand on these and other basic principles and clinical correlates of chemokine regulation of the immune system.

KEY CONCEPTS

Chemokine and Chemokine Receptors at a Glance

- *Definition:* Chemokines are defined by a common structure, the chemokine fold. Chemokine receptors are defined by a common biochemical function: chemokine binding-dependent cell signaling. Most chemokine receptors catalyze guanine nucleotide exchange on Gi-type G proteins. A small group of atypical receptors signal through an arrestin-dependent pathway.
- *Classification:* Chemokines form four main structural subclasses (C, CC, CXC, and CX3C) and two main immunological subclasses (inflammatory and homeostatic).
- *Evolution:* Chemokines and chemokine receptors arose in vertebrates and have been copied or mimicked by viruses. Chemokines and chemokine receptors rapidly evolve; the repertoires can differ among species and among individuals of the same species.
- *Ligand–receptor promiscuity:* The majority of chemokine receptors pair promiscuously with chemokine ligands, usually restricted to a single chemokine subclass; these typically mediate inflammatory responses.
- *Cell biology:* Chemokines coordinate leukocyte trafficking but can have prominent nontrafficking functions (*e.g.,* lymphocyte proliferation/apoptosis/differentiation/activation, granulocyte degranulation/superoxide production, direct antimicrobial activity), and nonimmunological effects on other cell types (*e.g.,* development, cancer, angiogenesis).
- *Biology:* Chemokines act redundantly or nonredundantly *in vivo,* depending on the context. Host chemokine receptors mediate antimicrobial defense, but certain pathogens (*e.g.,* HIV and *Plasmodium vivax* infection) can exploit chemokine receptors to infect the host. Excessive or inappropriate chemokine expression may pathologically amplify immunologically mediated disease.

MOLECULAR ORGANIZATION OF THE CHEMOKINE SYSTEM

Chemokines are defined by structure rather than function.[1] This reflects the highly conserved nature of their tertiary folded structure that is created by uniformly spaced, disulfide-bonded cysteines (Fig. 10.1). Chemokines contain three β-sheets arranged in the shape of a Greek key, overlaid by a *C*-terminal α-helical domain and flanked by an *N*-terminal domain that lacks order. All chemokines have at least two cysteines, and all but two have at least four. In the four-cysteine group, the first two are either adjacent (CC motif, $n = 24$) or separated by either one amino acid (CXC motif, $n = 16$) or three amino acids (CX3C motif, $n = 1$). Disulfide bonds link C-1 to C-3 and C-2 to C-4. C chemokines ($n = 2$) have only two cysteines, corresponding to C-2 and C-4 in the other groups. Sequence identity is less than 30% for any two chemokines from different groups but ranges between 30% and 99% for any two chemokines from the same group. The group names are used as roots followed by the letter "L" and a number (*e.g.,* CXCL1) in a systematic nomenclature that was established to resolve competing aliases.[2]

Several additional motifs enable subclassification of CC and CXC chemokines. The seven CXC chemokines with glu-leu-arg (ELR) *N*-terminal to C-1 share greater than 40% sequence identity, attract neutrophils, bind the same CXCR2 receptor, and are angiogenic (Table 10.1). CXCL9, CXCL10, and CXCL11 share a receptor (CXCR3) as well as greater than 40% identity, but they are angiostatic.

Two CC subgroups have two additional cysteines. One is in the *C*-terminal domain, and the other can be found either in the *C*-terminal domain or between C-2 and C-3 (Table 10.2). CXCL16 and CX3CL1 cross classes to form a unique multimodular subgroup. Each has a classic chemokine domain, a mucin-like stalk, a transmembrane domain, and a *C*-terminal cytoplasmic module. Each can exist as either a membrane-bound form or a shed form, enabling either direct cell–cell adhesion or chemotaxis, respectively. Chemokine monomer, dimer, and tetramer structures

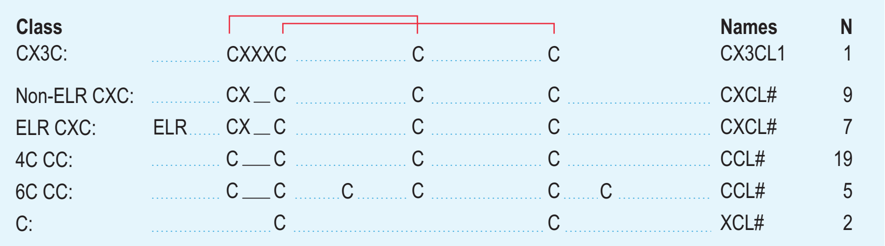

Class
Names
N
CX3C: CXXXC C C CX3CL1 1
Non-ELR CXC: CX_C C C CXCL# 9
ELR CXC: ELR CX_C C C CXCL# 7
4C CC: C__C C C CCL# 19
6C CC: C__C C C C C CCL# 5
C: C C XCL# 2

FIG 10.1 Chemokine Classification and Nomenclature. Chemokine classes are defined by the number and arrangement of conserved cysteines, as shown. Brackets link cysteines that form disulfide bonds. ELR refers to the amino acids glu-leu-arg. X refers to an amino acid other than cysteine. The underscore is a spacer used to optimize the alignment. The *N*- and *C*-termini can vary considerably in length (not illustrated). For molecules with four cysteines, there are approximately 24 amino acids between Cys-2 and Cys-3, and 15 amino acids between Cys-3 and Cys-4. At right are listed the nomenclature system and the number of human chemokines known in each class (N).

TABLE 10.1 Human CXC, CX3C, and C Chemokines

Chemokine	Common Aliases	Main Sources	G Protein–Coupled Receptors	Main Immunological Roles
CXCL1, 2 and 3	GROα, β, and γ	Inducible in most hematopoietic and tissue cells	CXCR2	Neutrophil trafficking
CXCL4	PF4	Preformed in platelets	CXCR3	Procoagulant
CXCL5	ENA-78	Induced in epithelial cells of gut and lung; N, Mo, Plts, EC	CXCR2	Neutrophil trafficking
CXCL6	GCP-2	Induced in lung microvascular EC; Mo; alveolar epithelial cells, mesothelial cells, EC and MΦ	CXCR1, CXCR2	Neutrophil trafficking
CXCL7	NAP-2	Preformed in platelets	CXCR2	Neutrophil trafficking
CXCL8	IL-8	Induced in most cell types	CXCR1, CXCR2	Neutrophil trafficking
CXCL9	Mig	Induced in PMN, MΦ, T cells, astrocytes, microglial cells, hepatocytes, EC, fibroblasts, keratinocytes, thymic stromal cells	CXCR3	Th1 response
CXCL10	IP-10	Induced in ECs, Mo, keratinocytes, respiratory and intestinal epithelial cells, astrocytes, microglia, mesangial cells, smooth muscle cells	CXCR3	Th1 response
CXCL11	I-TAC	ECs, Mo	CXCR3	Th1 response
CXCL12	SDF-1, PBSF	Constitutive in bone marrow stromal cells; most tissues	CXCR4	Myelopoiesis; HPC, neutrophil homing to BM; B lymphopoiesis
CXCL13	BCA-1	Constitutive in follicular HEV of secondary lymphoid tissue	CXCR5	Naïve B- and T-cell homing to follicles; B_1-cell homing to peritoneum; Natural Ab production
CXCL14	BRAK	Constitutive in most tissues, breast and kidney tumors	ND	Macrophage migration
(CXCL15)	(Mouse only)	Constitutive in lung epithelial cells	ND	Neutrophil trafficking
CXCL16	Sexckine	Constitutive in spleen; DCs of the T cell zone	CXCR6	T-cell and DC homing to spleen
CXCL17		Lung, heart, tumor cells	CXCR8	Monocyte and myeloid DC trafficking
CX3CL1	Fractalkine	EC, neurons, Mo, DC	CX3CR1	NK, monocyte, MΦ, and Th1-cell migration
XCL1 and XCL2	Lymphotactin α and β	γδ epidermal T cells, NK, NK-T, activated CD8 and Th1 CD4 T cells	XCR1	$CD62L^{lo}$ T effector cell migration

NA, not applicable; Mo, monocyte; PMN, neutrophil; DC, dendritic cell; EC, endothelial cell; HEV, high endothelial venule; MPC, myeloid progenitor cell; plt, platelet; MΦ, macrophage; GRO, growth-related oncogene; PF4, platelet factor 4; GCP, granulocyte chemoattactant protein; ENA-78, 78 amino acid epithelial cell-derived neutrophil activator; NAP, neutrophil activating protein; IL-8, interleukin-8; Mig, monokine induced by IFN-γ; I-TAC; interferon-inducible T-cell αchemoattractant; SDF, stromal cell-derived factor; BCA, B cell–activating chemokine; BRAK, breast and kidney-associated chemokine.

TABLE 10.2 CC Chemokines

Chemokine	Common Aliases	Sources	Main G Protein–Coupled Receptors	Main Immunological Roles
CCL1	I-309	Inducible in Mo and CD4$^+$ and CD8$^+$ αβ and CD4$^-$CD8$^-$ γδ T cells	CCR8	Th2 response
CCL2	MCP-1	Inducible in Mo, fibroblasts, keratinocytes, EC, PMN, synoviocytes, mesangial cells, astrocytes, lung epithelial cells, and MΦ. Constitutively made in splenic arteriolar lymphatic sheath and medullary region of lymph node, many tumors, and arterial plaque EC	CCR2	Innate immunity Th2 response CD4$^+$ T-cell differentiation
CCL3	MIP-1α LD78α MIP-1αS	Inducible in Mo/MΦ, CD8 T cells, B cells, plts, PMN, Eo, Ba, DC, NK, mast cells, keratinocytes, fibroblasts, mesangial cells, astrocytes, microglial cells, epithelial cells	CCR1, CCR5	Innate immunity Th1 response CD4 T-cell differentiation
CCL3L1	LD78β MIP-1αP	Similar to CCL3	CCR1, CCR5	Probably similar to CCL3
CCL4	MIP-1β	Similar to CCL3	CCR5	Innate immunity Th1 response
CCL5	RANTES	Inducible in EC, T cells, epithelial cells, Mo, fibroblasts, mesangial cells, NK cells Constitutively expressed and stored in plt and Eo granules	CCR1, CCR3, CCR5	Innate immunity Th1 and Th2 response
(CCL6)	Mouse only	Inducible in bone marrow and peritoneal-derived MΦ	CCR1	ND
CCL7	MCP-3	Inducible in Mo, plts, fibroblasts, EC, skin, bronchial epithelial cells, astrocytes	CCR1, CCR2, CCR3, CCR5	Th2 response
CCL8	MCP-2	Inducible in fibroblasts, PMN, astrocytes Constitutively expressed in colon, small intestine, heart, lung, thymus, pancreas, spinal cord, ovary, placenta	CCR8	Th2 response
(CCL9/10)	Mouse only	Constitutively expressed in all mouse organs except brain; highest in lung, liver and thymus Induced in heart and lung	CCR1	ND
CCL11	Eotaxin	Epithelial cells, EC, smooth muscle, cardiac muscle, Eo, dermal fibroblasts, mast cells, MΦ, Reed–Sternberg cells	CCR3	Th2 response Eosinophil trafficking Mast cell trafficking Basophil trafficking, degranulation
(CCL12)	Mouse only	Inducible in lung alveolar MΦ and smooth muscle cells; spinal cord. Constitutive expression in Lymph node and thymic stromal cells	CCR2	Allergic inflammation
CCL13	MCP-4	Inducible in nasal and bronchial epithelial cells; dermal fibroblasts; PBMCs; atherosclerotic plaque EC and MΦ Constitutively expressed in small intestine, colon, thymus, heart, and placenta	CCR1, CCR2, CCR3	Th2 response
CCL14a	HCC-1	Constitutively expressed in most organs; high plasma levels	CCR1, CCR5	ND
CCL14b	HCC-3	Same as CCL14b except absent from skeletal muscle and pancreas	CCR1, CCR5	ND
CCL15	HCC-2; Lkn-1	Inducible in Mo and DC Constitutive RNA expression in liver, gut, heart, and skeletal muscle, adrenal gland and lung leukocytes	CCR1, CCR3	ND
CCL16	HCC-4; LEC	Constitutively expressed in liver, possibly many other organs. Also, Mo, T cells, and NK cells express mRNA	CCR1, CCR2, CCR5	ND
CCL17	TARC	Constitutive in normal DC and Reed-Sternberg cells of Hodgkin disease	CCR4	Th2 response
CCL18	DC-CK1, PARC	Constitutive in Mo/MΦ, germinal center DCs	CCR8	DC attraction of naïve T cells
CCL19	ELC, MIP-3β	Constitutive on interdigitating DCs in secondary lymphoid tissue	CCR7	Naïve and memory T-cell and DCs homing to lymph node
CCL20	LARC MIP-3α	Constitutive in lymph nodes, peripheral blood leukocytes, thymus, and appendix Inducible in PBMC, HUVEC	CCR6	DCs homing to Peyer patch Humoral response
CCL21	SLC, 6Ckine	Constitutive in lymphatic EC, HEV and interdigitating DCs in T areas of 2° lymphoid tissue, thymic medullary epithelial cells and EC	CCR7	Naïve and memory T cells and DCs homing to lymph node
CCL22	MDC	Constitutive in DC and MΦ Inducible in Mo, T, and B cells	CCR4	Th2 response
CCL23	MPIF-1	Constitutive in pancreas and skeletal muscle	CCR1	ND
CCL24	Eotaxin-2	Inducible in Mo	CCR3	Eosinophil migration
CCL25	TECK	Constitutive in thymic stromal cells and small intestine	CCR9	Thymocyte migration Homing of memory T cells to gut
CCL26	Eotaxin-3	Constitutive in heart and ovary Inducible on dermal fibroblasts and EC	CCR3	Th2 response

Continued

TABLE 10.2 CC Chemokines—cont'd

Chemokine	Common Aliases	Sources	Main G Protein–Coupled Receptors	Main Immunological Roles
CCL27	CTACK, Eskine	Constitutive in placenta, keratinocytes, testis, and brain	CCR10	Homing of memory and effector T cells to skin
CCL28	MEC	Constitutive in epithelial cells of gut, airway	CCR10	Homing of T cells to mucosal surfaces

NA, not applicable; Mo, monocyte; PMN, polymorphonuclear neutrophil; DC, dendritic cell; EC, endothelial cell; HEV, high endothelial venule; MPC, myeloid progenitor cell; plt, platelet; MΦ, macrophage; MCP, monocyte chemoattractant protein; MIP, macrophage inflammatory protein; RANTES, regulated upon activation normal T cell expressed and secreted; MRP, MIP-related protein; HCC, hemofiltrate CC chemokine; TARC, thymus and activation-related chemokine; PARC, pulmonary and activation-related chemokine; ELC, Epstein-Barr virus-induced receptor ligand chemokine; LARC, liver and activation-related chemokine; SLC, secondary lymphoid tissue chemokine; MDC, macrophage-derived chemokine; MPIF, myeloid progenitor inhibitory factor; TECK, thymus-expressed chemokine; CTACK, cutaneous T-cell-associated chemokine; MEC, mucosa-associated epithelial cell chemokine.

can occur, and complex quaternary structures bound to glycosaminoglycans (GAGs) on the surface of cells may be important for function *in vivo*.[1]

Chemokine Receptors

Chemokine receptors are defined as mediators that activate cellular responses upon binding of chemokines. Twenty-three subtypes of human chemokine receptors have been identified, all of which are members of the seven-transmembrane (7TM) domain superfamily of receptors.[3] They can be divided into two main groups: the G protein–coupled chemotactic chemokine receptors ($n = 19$) and the atypical chemokine receptors ($n = 4$). Chemokine binding, membrane anchoring, and signaling domains for receptors from both groups come from a single polypeptide chain. Structural and biochemical evidence exists that these receptors form homo- and heterodimers.

Some chemokine receptors pair monogamously with their chemokine ligands. Most, however, are promiscuous but restricted to one chemokine structural group (Fig. 10.2). The G protein–coupled receptors are named by ligand group specificity. Each chemokine has a unique receptor specificity profile, and each receptor has a unique chemokine specificity profile. Almost all chemokines are chemotactic agonists, and a few may be agonists at one G protein–coupled chemokine receptor and antagonists at another, in addition to binding to atypical receptors. Differential receptor usage, differential regulation of expression, and biased agonism may all account for nonredundant function observed *in vivo* for chemokines acting at the same G protein–coupled receptor.

Atypical Chemokine System Components

There are three classes of atypical chemokine system components: (1) the atypical 7TM chemokine receptors: as described previously, these bind chemokines promiscuously without signaling or with atypical signaling. These proteins are thought to function as chemokine scavengers but may also facilitate chemokine transcytosis across endothelial barriers;[4] (2) endogenous nonchemokine agonists: these act at chemokine receptors (*e.g.*, β-defensin-2 at CCR6); (3) virally encoded chemokines, 7TM chemokine receptors, structurally unique chemokine-binding proteins (scavengers), and nonchemokine chemokine receptor ligands (agonists or antagonists);[5] examples of the last one include two viral chemokine mimics encoded by HIV: the gp120 and tat proteins, which possess chemokine agonist and antagonist activity, respectively. Viral chemokine elements can function to evade the immune system, recruit new target cells, reprogram gene expression for cell proliferation and angiogenesis, or target cell entry (*e.g.*, HIV gp120).[6] Secreted chemokine-binding proteins have even been identified in tick saliva, which could explain, in part, the lack of inflammation associated with tick bites.[7]

KEY CONCEPTS

Immunological Classification of the Chemokine System

- *Homeostatic system:* Constitutively expressed ligands and receptors. Important in hematopoiesis and immune surveillance. Key receptors: CXCR4 on all leukocytes, especially hematopoietic stem and progenitor cells; CXCR5 on B cells; CCR7 on dendritic cells and T cells; and gut and skin-specific T-cell homing receptors CCR9 and CCR10, respectively.
- *Inflammatory system:* In innate immunity, inducible ligands and constitutively expressed receptors (*e.g.*, neutrophil CXCR2, monocyte/macrophage CCR2 and CX3CR1, eosinophil CCR3, and NK cell CX3CR1). In adaptive immunity, inducible ligands and inducible receptors (*e.g.*, CXCR3, CCR4, and CCR6 on Th1, Th2, and Th17 subsets of CD4 T cells, respectively).

Immunological Classification

Each of the leukocyte subtypes responds to chemokines via a characteristic subset of chemokine receptors. The chemokine system can be subclassified into two main subsystems, *homeostatic* and *inflammatory*, based on receptor expression patterns. Homeostatic chemokines are differentially and constitutively expressed in specific microenvironments within primary and secondary immune organs. They recruit both immature and mature leukocytes via constitutively expressed receptors. Noxious stimuli induce inflammatory chemokines in diverse tissue cells and leukocytes. Inflammatory chemokine receptors are constitutively expressed on myeloid and natural killer (NK) cells but must be induced on activated effector lymphocytes. Dynamic shifts in receptor expression occur during dendritic cell (DC) and NK cell maturation, as well as during lymphocyte maturation, activation, and differentiation.

Inflammatory CXC and CC chemokine genes are found in two main clusters on human chromosomes 4q12-q21 and 17q11-q21, respectively. Conversely, homeostatic chemokine genes have undergone a diaspora resulting in small clusters of genes on multiple chromosomes. Thirteen of the 19 human chemokine receptor genes are clustered at 3p21-23, and *CXCR1* and *CXCR2* are adjacent at 2q34-q35. The chemokine and chemokine receptor repertoire may vary even among closely related species. Gene copy number (*e.g.*, for *CCL3*) and sequence

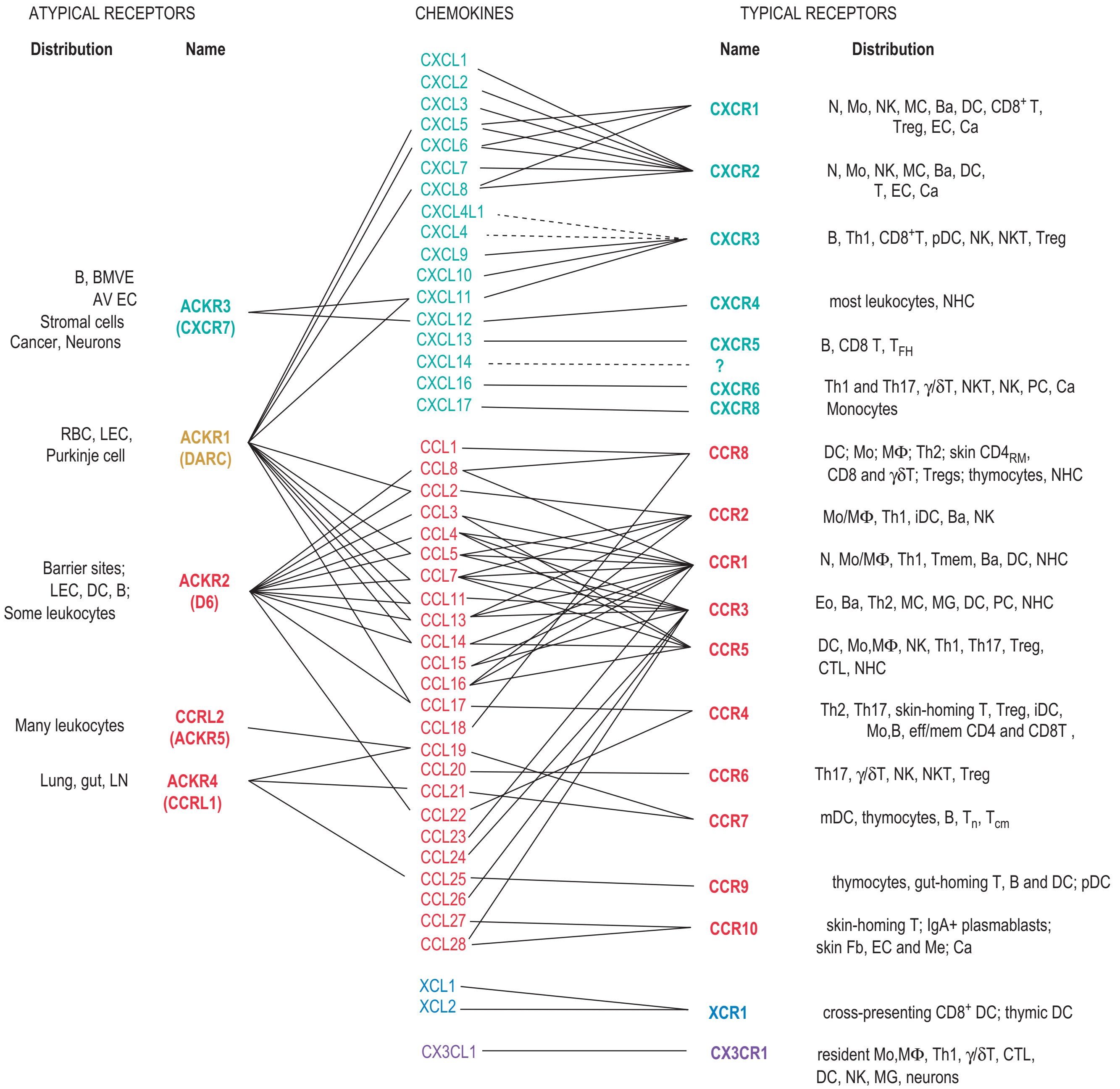

ATYPICAL RECEPTORS
Distribution
Name
CHEMOKINES
TYPICAL RECEPTORS
Name
Distribution
B, BMVE
AV EC
Stromal cells
Cancer, Neurons
ACKR3
(CXCR7)
RBC, LEC,
Purkinje cell
ACKR1
(DARC)
Barrier sites;
LEC, DC, B;
Some leukocytes
ACKR2
(D6)
Many leukocytes
CCRL2
(ACKR5)
Lung, gut, LN
ACKR4
(CCRL1)
CXCL1
CXCL2
CXCL3
CXCL5
CXCL6
CXCL7
CXCL8
CXCL4L1
CXCL4
CXCL9
CXCL10
CXCL11
CXCL12
CXCL13
CXCL14
CXCL16
CXCL17
CCL1
CCL8
CCL2
CCL3
CCL4
CCL5
CCL7
CCL11
CCL13
CCL14
CCL15
CCL16
CCL17
CCL18
CCL19
CCL20
CCL21
CCL22
CCL23
CCL24
CCL25
CCL26
CCL27
CCL28
XCL1
XCL2
CX3CL1
CXCR1
N, Mo, NK, MC, Ba, DC, CD8$^+$ T, Treg, EC, Ca
CXCR2
N, Mo, NK, MC, Ba, DC, T, EC, Ca
CXCR3
B, Th1, CD8$^+$T, pDC, NK, NKT, Treg
CXCR4
most leukocytes, NHC
CXCR5
B, CD8 T, T_{FH}
?
CXCR6
Th1 and Th17, γ/δT, NKT, NK, PC, Ca
CXCR8
Monocytes
CCR8
DC; Mo; MΦ; Th2; skin $CD4_{RM}$, CD8 and γδT; Tregs; thymocytes, NHC
CCR2
Mo/MΦ, Th1, iDC, Ba, NK
CCR1
N, Mo/MΦ, Th1, Tmem, Ba, DC, NHC
CCR3
Eo, Ba, Th2, MC, MG, DC, PC, NHC
CCR5
DC, Mo,MΦ, NK, Th1, Th17, Treg, CTL, NHC
CCR4
Th2, Th17, skin-homing T, Treg, iDC, Mo,B, eff/mem CD4 and CD8T ,
CCR6
Th17, γ/δT, NK, NKT, Treg
CCR7
mDC, thymocytes, B, T_n, T_{cm}
CCR9
thymocytes, gut-homing T, B and DC; pDC
CCR10
skin-homing T; IgA+ plasmablasts; skin Fb, EC and Me; Ca
XCR1
cross-presenting CD8$^+$ DC; thymic DC
CX3CR1
resident Mo,MΦ, Th1, γ/δT, CTL, DC, NK, MG, neurons

FIG 10.2 Human Chemokine Specificities for Chemokine Receptors. Abbreviations: N, neutrophil; NK, natural killer cell; NKT, NK T cell; Mo, monocyte; DC, dendritic cell; B, B lymphocyte; BMVEC, bone marrow vascular endothelial cell; AVEC, aortic vascular endothelial cell; RBC, red blood cell; LEC, lymphatic endothelial cell; LN, lymph node; MG, microglial cells; CTL, cytotoxic T lymphocytes; MΦ, macrophage; pDC, plasmacytoid dendritic cells; iDC, immature dendritic cells; Ba, basophil; Treg, regulatory T cells; EC, endothelial cells; Ca, cancer; NHC, nonhematopoietic cells; T_{fh}, T-follicular helper cells; PC, plasma cells; Eo, eosinophil; T_{mem}, memory T cells; T_{cm}, central memory T cells; Fb, fibroblasts.

can vary among individuals of a species. This variation may influence the risk for acquiring certain diseases.

Chemokine Presentation Mechanisms

Chemokines act locally. They appear to be tethered to extracellular matrix proteins or to endothelial cells by binding to glycosaminoglycans or, for CX3CL1 and CXCL16, by transmembrane domains.[1] The tethering cell may have produced the chemokine or else imported it by transcytosis from neighbors. The ligand-binding site includes the receptor *N*-terminus and one or more extracellular loops, which allow docking of the chemokine *N*-loop domain, and multiple 7TM domains, which accept the chemokine's *N*-terminus and are critical for triggering.

Leukocyte Responses to Chemokines

All leukocyte subtypes migrate in response to chemokines. However, each subtype can also respond in additional stereotypical ways. Lymphocytes may proliferate, undergo apoptosis, or release immunoregulatory and cytotoxic factors. Granulocytes may release antimicrobial and inflammatory mediators (*e.g.*, superoxide,

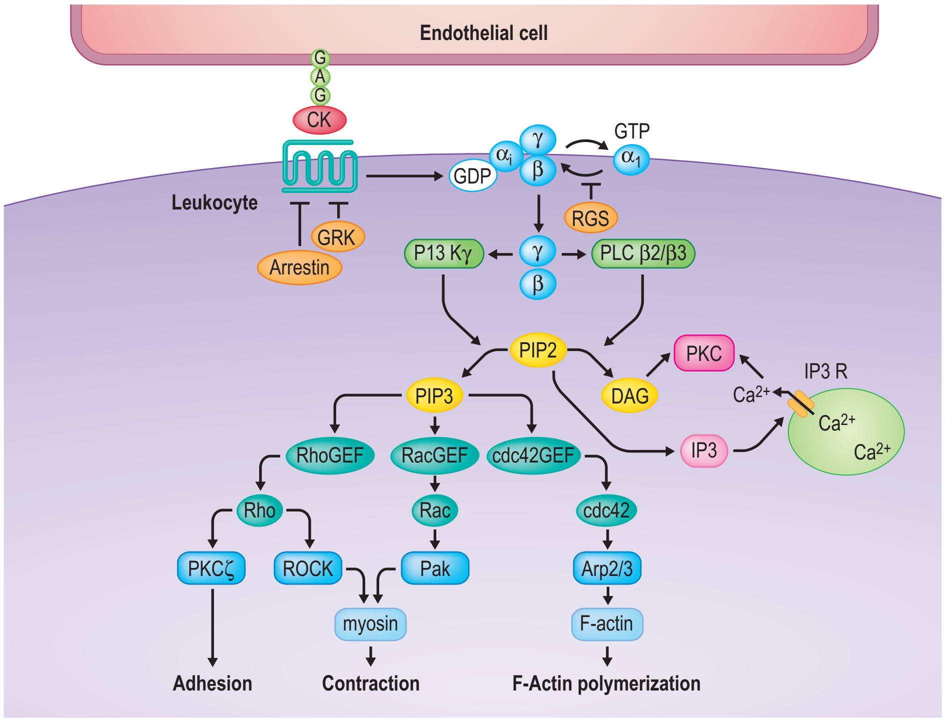


FIG 10.3 **Chemokine Signal Transduction in Chemotaxis.** Depicted are key steps in two of the main pathways induced by most chemokines. The PI3Kγ pathway is particularly important for cell migration. Chemokines are able to activate other pathways as well, including non-Gi-type G proteins, protein tyrosine kinases, and MAP kinases. These pathways influence cell proliferation and activation. The model is modified from the Alliance for Cell Signaling (http://www.signaling-gateway.org). PLC, phospholipase C; PI3K, phosphatidylinositol-3-kinase; RGS, regulator of G protein signaling; DAG, diacylglycerol; IP_3, inositol trisphosphate; PIP, phosphatidylinsol phosphate; GAG, glycosaminoglycan; CK, chemokine; PKC, protein kinase C; GRK, G protein–coupled receptor kinase; GEF, guanine nucleotide exchange factor.

defensins, proteases, histamine, eicosanoids). The mechanism of leukocyte migration can vary, depending on the leukocyte and the environment.[8] The mechanism is best understood for transendothelial migration (Chapter 11).[9] In an initial chemokine-independent step, leukocytes roll on inflamed endothelium in a selectin-dependent manner. Next, chemokines posted on endothelium stimulate rolling leukocytes to express activated β_2 integrins, which mediate firm adhesion via endothelial intercellular adhesion molecules (ICAMs). Leukocytes sense chemokine gradients, polarize, and then become poised to crawl. Motion involves shear-dependent coordinated cytoskeletal remodeling. This includes expansion of the leading edge (lamellipodium), myosin-based contraction at the trailing edge (uropod), release of the uropod from substrate, and membrane lipid movement.[10] Navigation through tissue may require relays of chemokines and adhesion molecules.

CHEMOKINE SIGNALING PATHWAYS

Chemokines trigger G protein–coupled chemokine receptors, mainly heterotrimeric Gi-type G proteins, to act as guanine nucleotide exchange factors (GEFs). Binding either releases guanosine diphosphate (GDP) from or binds guanosine triphosphate (GTP) to the Gi α subunit.[11] This results in G protein dissociation into α and βγ subunits, which, in turn, activates diverse G protein–dependent effectors, including phospholipases A2, C (subtypes β_2 and β_3) and D, phosphatidylinositol-3-kinase γ (PI3Kγ), protein tyrosine kinases (PTKs) and phosphatases, low-molecular-weight guanosine triphosphatases (GTPases), and mitogen-activated protein kinases (Fig. 10.3).

Phospholipase C (PLC) hydrolyzes PI bisphosphate (PIP_2) to form 1,2-diacylglycerol (DAG) and inositol 1,4,5-trisphosphate (IP_3). IP_3 induces Ca^{2+} release from intracellular stores, which acts with 1,2-diacylglycerol (DAG) to activate protein kinase C (PKC). PI3Kγ phosphorylates PIP_2 to form PIP_3, which recruits proteins containing pleckstrin homology (PH) or PHOX (PX) domains to lamellipodium, thereby converting shallow analogue extracellular chemokine gradients into steep digital intracellular effector gradients. Four PH domain-containing targets—Akt, and GEFs for Rac, Rho, and Cdc42—modulate distinct phases of cell movement in various model systems. Rho regulates cell adhesion, chemotaxis, and myosin contraction. Rac and Cdc42 control lamellipodia and filopodia formation, respectively. Downstream targets of Rac include Pak1, which also regulates myosin contraction.

REGULATION OF CHEMOKINE ACTION

Chemokine and chemokine receptor expression can be positively or negatively regulated at the transcriptional level by diverse factors, including proinflammatory cytokines, oxidant stress, viruses, bacterial products (*e.g.,* lipopolysaccharide [LPS] and

N-formyl peptides), cell adhesion, antigen uptake, T-cell costimulation, and diverse transcription factors (*e.g.,* NF-κB). In innate immunity, proinflammatory cytokines, such as interleukin-1 (IL-1), tumor necrosis factor (TNF), and IL-15 induce expression of inflammatory chemokines important for recruitment of myeloid and NK cells. In adaptive immunity, signature cytokines of polarized helper T cells establish positive feedback loops for production of signature chemokines able to specifically recruit these cells. For example, T helper-1 (Th1) cells produce interferon-γ (IFN-γ), which induces expression of CXCL9, CXCL10, and CXCL11, the chemokine agonists specific for the signature Th1 cell chemokine receptor CXCR3. This amplifies Th1 cell recruitment. Similar loops may exist for Th2 cells involving CCL17/CCR4 and IL-4, as well as for Th17 cells involving CCL20/CCR6 and IL-17. IFNs, glucocorticoids, and antiinflammatory cytokines (*e.g.,* IL-10 and transforming growth factor-β [TGF-β]) can inhibit inflammatory chemokine gene expression. Chemokines can also be regulated at the level of messenger ribonucleic acid (mRNA) stability.

A chemokine gene can generate families of proteins varying in activity and potency by alternative splicing and posttranslational modification, especially *N*- and *C*-terminal proteolytic trimming.[12] Proteases can target many chemokines (*e.g.,* CD26 [dipeptidyl peptidase IV] and matrix metalloproteinases [MMPs]), or few or only one (*e.g.,* TACE [the TNF-α–converting enzyme], plasmin, urokinase plasminogen activator, and cathepsin G). Chemokine action can be blocked by atypical chemokine receptors (*e.g.,* ACKR1), endogenous receptor antagonists, receptor decoys, and autoantibodies. Cytokines may also convert a signaling receptor into a decoy (*e.g.,* IL-10 inactivates CCR2 on monocytes).

CHEMOKINE REGULATION OF HEMATOPOIESIS

Bone Marrow

Most chemokines that modulate hematopoietic progenitor cell (HPC) proliferation *ex vivo* act early during hematopoiesis and are inhibitory. CXCL12, the most abundant chemokine in bone marrow, is an important exception. CXCL12 signaling through its receptor CXCR4 is critical for bone marrow myelopoiesis and B-cell lymphopoiesis. CXCR2 and CXCR4 both regulate neutrophil egress from bone marrow, and CXCR4 is also critical for mobilization of hematopoietic stem cells (HSCs) and HPCs from bone marrow.[13] A second CXCL12 receptor, ACKR3 (previously known as CXCR7), does not regulate hematopoiesis, but interestingly it is an essential factor responsible for cardiac valve development apparently due to its ability to bind the nonchemokine ligand adrenomedullin. ACKR3 also regulates marginal zone B-cell positioning in the spleen. CCR2 is important for monocyte release from bone marrow.

Thymus

During development, T cells must migrate from the thymic cortex to the medulla (Chapter 8). Chemokines and chemokine receptors are differentially expressed in thymus and coordinate migration.[14] CCR9 and its ligand, CCL25, may be important, since competitive transplantation of CCR9$^{-/-}$ bone marrow is less efficient than normal marrow at repopulating the thymus of lethally irradiated Rag-1$^{-/-}$ mice.

CCL25 is expressed by medullary DCs and both cortical and medullary epithelial cells. CCR9 is expressed on the majority of immature CD4$^+$CD8$^+$ thymocytes but is downregulated during transition to the CD4$^+$ or CD8$^+$ single-positive stage (Fig. 10.4). Just before thymic egress, thymocytes become CCR9 negative and upregulate L-selectin. Transition from CD4$^+$CD8$^+$ thymocytes in the cortex to CD4$^+$ or CD8$^+$ single-positive thymocytes in the medulla is associated with upregulation of CCR4 and CCR7, receptors for CCL22, and CCL19 and CCL21, respectively, which are expressed in the medullary stroma. Accordingly, these chemokines attract thymocytes between the late cortical and medullary stages of development *in vitro.* Neutralization studies suggest that egress of newly formed T cells from the fetal thymus to the circulation is mediated by CCL19, which is selectively localized on endothelial cells of medullary venules and acts at CCR7 on mature thymocytes.

Tissue

Once myeloid cells are released from bone marrow, they undergo specific trafficking itineraries and in some cases become resident in tissue. This is regulated, in part, by specific chemokines. For example, CXCL14 is important for macrophage positioning in the lung, and CCL11 and its receptor CCR3 for eosinophils in the spleen and the gastrointestinal tract. CCR6 regulates positioning of immature myeloid CD11c$^+$CD11b$^+$ DCs in the subepithelial dome of Peyer patches. CX3CR1 regulates localization of myeloid DCs in Peyer patches and may be important for antigen sampling from the intestine.

CHEMOKINE REGULATION OF THE IMMUNE RESPONSE

The innate and adaptive immune systems are deployed separately but are assembled, in part, by specific sets of chemokines and chemokine receptors (see Fig. 10.4).[15]

Innate Immunity

Platelet-Derived Chemokines

Made primarily during platelet development, stored in platelet α granules, and rapidly released during platelet degranulation, CXCL4 and CXCL7 are among the first chemokines to appear at sites of tissue injury and infection. When there is hemorrhage and vascular damage, chemokine concentrations can reach elevated levels.[16] CXCL7 can function as an immediate-early mediator of neutrophil recruitment released from platelets at sites of inflammation. Although it is not a prominent leukocyte chemoattractant and does not induce degranulation of neutrophil lysosomal enzymes, CXCL4 is able to induce neutrophil secondary granule exocytosis and release of matrix-degrading enzymes, which may facilitate neutrophil penetration of infected or injured tissues.

CXCL8 and CXCR2

All seven ELR$^+$ CXC chemokines preferentially recruit neutrophils *in vitro* by binding to CXCR2. Two of these, CXCL6 and CXCL8, are also potent agonists at CXCR1, which is coexpressed at similar levels on neutrophils.[17] These seven chemokines are rapidly inducible but may differ biologically because of temporal and spatial differences in expression. These differences provide a mechanism for the graded navigation of neutrophils through tissue. Blocking studies in multiple animal models have demonstrated the importance of CXCL8 and CXCR2 in neutrophil accumulation in response to infectious and noninfectious stimuli. Intradermal injection of CXCL8 in humans causes rapid (<30 min) and selective accumulation of large numbers of

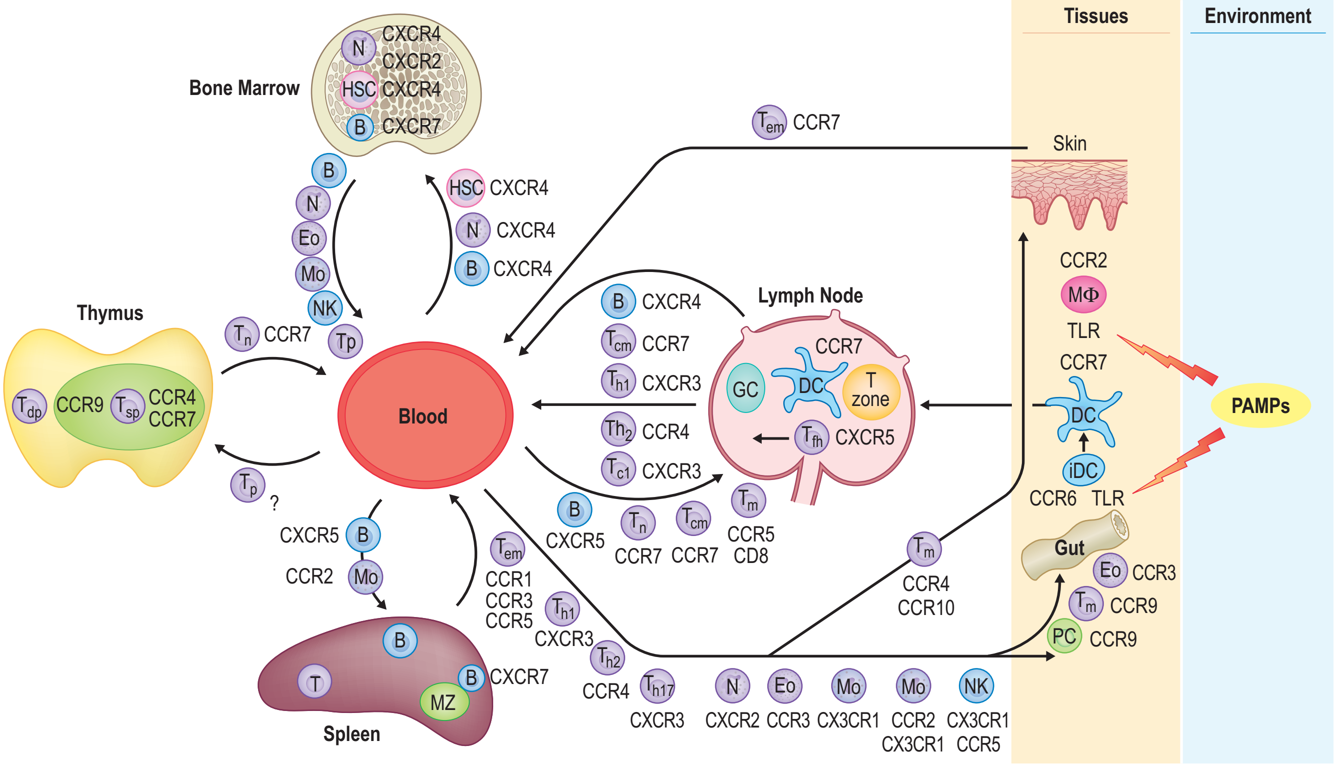

Tissues
Environment
Bone Marrow
N CXCR4
CXCR2
HSC CXCR4
B CXCR7
T_{em} CCR7
Skin
B
N
Eo
Mo
NK
HSC CXCR4
N CXCR4
B CXCR4
Thymus
T_n CCR7
Tp
T_{dp} CCR9 T_{sp} CCR4 CCR7
Blood
B CXCR4
T_{cm} CCR7
T_{h1} CXCR3
Th_2 CCR4
T_{c1} CXCR3
Lymph Node
CCR7
GC
DC
T zone
T_{fh} CXCR5
CCR2
MΦ
TLR
CCR7
DC
PAMPs
iDC
CCR6 TLR
Gut
Eo CCR3
T_m CCR9
PC CCR9
T_p ?
CXCR5 B
CCR2 Mo
B CXCR5
T_n CCR7
T_{cm} CCR7
T_m CCR5 CD8
T_{em} CCR1 CCR3 CCR5
T_{h1} CXCR3
T_{h2} CCR4
T_{h17} CXCR3
T_m CCR4 CCR10
B
T
MZ
B CXCR7
Spleen
N CXCR2
Eo CCR3
Mo CX3CR1
Mo CCR2 CX3CR1
NK CX3CR1 CCR5

FIG 10.4 Chemokine Receptor Control of Leukocyte Trafficking. Shown are routes among primary and secondary immune organs and the periphery, leukocyte subtypes trafficking on those routes and some of the chemokine receptors that appear to be most important in regulating each route. Tn, naïve T cells; Tp, precursor T cells; Tm, memory T cells; TEM, effector memory T cells; TCM, central memory T cells; TFH, T-follicular helper cells; iDC, immature dendritic cells; N, neutrophil; Eo, eosinophil; MΦ, macrophage; Mo, monocyte; NK, natural killer cell; PC, plasma cell; HSC, hematopoietic stem cell; GC, germinal center. The model is based primarily on studies of mice where the relevant gene has been inactivated by gene targeting.

neutrophils in perivascular regions of skin without causing edema.

Tissue-specific transgenic overexpression of mouse CXCL8 paralogues KC and MIP-2 suggests that *in vivo* these factors may recruit cells, but not independently activate cytotoxic mechanisms. In a human blister model, endogenous CXCL8 peaks at ~24 hours, whereas C5a and leukotriene B4, which also recruit neutrophils, appear earlier. Thus the primary role of CXCL8 may be to amplify early inflammatory responses initiated by immediate-early chemoattractants, such as leukotriene B4.[18] CXCL1, -2, -3, -7, and 8 have also been reported to induce basophil chemotaxis and histamine release *in vitro*, which, together with other factors, such as complement-derived anaphylatoxins, promote vasodilatation during the early stages of the innate immune response.

Monocyte recruitment typically follows neutrophil accumulation with delayed kinetics. Recruitment can be mediated by multiple inflammatory CC receptors and CX3CR1. CCR2 and CX3CR1 are particularly important and define two monocyte subsets, $CX3CR1^{hi}CCR2^{-}$ and $CX3CR1^{lo}CCR2^{+}$, which are referred to as "resident" and "inflammatory" monocytes because of their distinct trafficking characteristics.[19]

Natural Killer Cells

Human NK cell subsets express unique repertoires of chemokine receptors.[20] The $CD56^{dim}CD16^{+}$ subset, which is associated with high cytotoxic capacity and low cytokine production, expresses primarily CXCR1 and CX3CR1. The minor $CD56^{bright}CD16^{dim}$ subset, which produces large amounts of cytokines but has low killing capacity, preferentially expresses CCR7. The exact profile of chemokine receptor expression can be modulated by adherence and stimulation *ex vivo* with IL-2. Cognate chemokines chemoattract NK cells and promote degranulation and killing.

The importance of chemokines in NK cell function *in vivo* is evident in mouse cytomegalovirus (MCMV) infection, a cause of hepatitis. MCMV induces CCL3 production in the liver, and this is required for recruitment of NK cells. NK cells are the major source of IFN-γ in this model. IFN-γ induces CXCL9, which is required for protection. Thus a cytokine–chemokine–cytokine cascade is required for NK cell–mediated host defense against this pathogen.

Dendritic Cells and Transition to the Adaptive Immune Response

Transition from innate to adaptive phases of the immune response involves antigen uptake by antigen-presenting cells (APCs), especially DCs, mediated by Fc and complement phagocytic receptors, as well as pattern recognition receptors (PRRs), including DC-SIGN and Toll-like receptors (TLRs) (Chapter 3). Both TLR2 and TLR4 signaling induce expression of CCL3, CCL4, and CCL5. However, TLR2 selectively induces CXCL8, whereas TLR4 selectively induces CXCL10. CXCL8 increases neutrophil migration to the site, whereas CXCL10 enhances NK cell or Th1 effector T-cell trafficking. Pathogens can skew the nature and magnitude of the immune response in a specific direction by means of specific pattern recognition receptor (PRR) ligands.

The chemokine receptors expressed on DCs vary, depending on the nature of the inflammatory stimulus and type. For example, blood-derived plasmacytoid and myeloid DCs express a similar repertoire of inflammatory chemoattractant receptors, but they are functional only on myeloid DCs. CCL3, CCL4, and CCL5 may be particularly important for recruiting additional mononuclear phagocytes and DCs to sites of infection. This can amplify the late stage of the innate immune response. In the extreme, this can devolve into endotoxic shock. Consistent with this idea, genetic disruption of the CCL3/CCL4/CCL5 receptor CCR5 renders mice relatively resistant to LPS-induced endotoxemia.

Adaptive Immunity

Afferent Trafficking to Secondary Lymphoid Tissue

The homeostatic receptors CXCR5 and CCR7 and their ligands are major regulators of the immune response. They act at the level of B and T lymphocytes and DCs trafficking to and within secondary lymphoid tissue.[21,22] DC maturation in peripheral tissues is associated with downregulation of inflammatory receptors, which is important for recruitment, migration, and retention in the periphery, and reciprocal upregulation of CCR7, which mediates mature DC migration to draining lymph nodes. Inflammatory receptors, such as CCR2, may also contribute to afferent trafficking. CCR7 is also a major lymph node trafficking receptor for naïve T cells and can mediate activated T-cell exit from inflamed tissue.

The CCR7 ligand CCL21 is constitutively expressed on afferent lymphatic endothelium, high endothelial venules (HEVs), stromal cells, and interdigitating DCs in the T zones of the lymph node, Peyer patch, mucosa-associated lymphoreticular tissue, and spleen.[8] It is not expressed in B-cell zones or sinuses. CCL19, another CCR7 ligand, is also restricted to the T-cell zone and is expressed on interdigitating DCs.

CCR7$^{-/-}$ mice and the *plt* mouse, which is naturally deficient in CCL19 and a CCL21 isoform expressed in secondary lymphoid organs, have similar phenotypes: atrophic T-cell zones populated by a paucity of naïve T cells. This and the failure of activated DCs to migrate to the lymph node from the skin of these mice explain why contact sensitivity, delayed type hypersensitivity, and antibody production are severely impaired. However, lymph node trafficking is not completely abolished in these mice, which may develop autoimmune phenomena.

CXCR5 is expressed on all peripheral blood and lymph node B cells as well as on some T cells. Its ligand, CXCL13, is expressed constitutively on follicular HEVs and controls trafficking of CXCR5 positive B and T cells from blood into the follicles. In *CXCR5*$^{-/-}$ mice, B cells do not migrate to the lymph node, Peyer patches are abnormal, and inguinal lymph nodes are absent. CXCL13 is also required for B1 cell homing, natural antibody production, and body cavity immunity. *CXCR5*$^{-/-}$ mice still can produce antibody, perhaps, in part, because B cells and follicular DCs can form ectopic germinal centers within T-cell zones of the periarteriolar lymphocyte sheath of spleen.

Migration Within Lymph Node Microenvironments

CXCR5 is expressed on a majority of memory CD4 T cells in the follicles of inflamed tonsils. T-follicular helper cells (T_{FH}), a CD57$^+$ subset of CXCR5$^+$ T cells, lack CCR7, which licenses them to move from the T zone following activation to the follicles where they provide help for B-cell maturation and antibody production. Reciprocally, B cells activated by antigen in the follicles upregulate CCR7 and move toward the T-cell zone. Thus B–T interaction can be facilitated by the reciprocal movement of these cells, which may be influenced by the balance of chemokines made in adjacent lymphoid zones.[23] CXCR4 signaling is also important in naïve and memory B-cell trafficking to germinal centers. CCR5 ligands can guide naïve CD8 T cells to sites of CD4 T cell–DC interaction in lymph nodes.

Efferent Trafficking

Naïve lymphocytes that do not encounter antigen continue to recirculate between blood and secondary lymphoid tissue without acquiring any tissue-specific homing properties. Most antigen-stimulated T cells die by apoptosis. The survivors can be divided into functionally distinct subsets marked by characteristic patterns of chemokine receptor expression. In general, the trafficking properties of these cells are not well understood.

Among CD4 T cells, three memory subsets and two main effector subsets have been proposed. The memory subsets include T_{FH} cells described above as well as effector memory cells (T_{EM}), central memory cells (T_{CM}), and tissue resident memory cells (T_{RM}). T_{EM} do not express L-selectin or CCR7. They traffic through peripheral tissues as immune surveillance cells, rapidly releasing cytokines in response to activation by recall antigens. T_{CM} cells express CCR7 and lymph node homing receptors. They traffic between blood and secondary lymphoid organs but are not polarized and lack immediate effector function. Instead, they efficiently interact with DCs in the lymph node and differentiate into effector cells upon secondary stimulation. Highly heterogeneous, multifunctional human T_{EM} cells are produced early in differentiation and may become stable resting cells.[24]

Upon antigen activation, the classic effector CD4 T-cell subsets, Th1, Th2, and Th17, downregulate CCR7 and upregulate inflammatory chemokine receptors. Exit from lymph node via efferent lymphatics is mediated by additional mechanisms, including S1P1 (sphingosine 1 phosphate type 1 receptor) signaling, which can be blocked by the drug FTY720. *In vivo,* Th1 cells more frequently express CXCR3, CXCR6, CCR2, CCR5, and CX3CR1 than do Th2 cells, and the pattern of receptor expression is a function of maturation and age.[25] In contrast, Th2 cells more frequently express CCR3, CCR4, and CCR8 than do Th1 cells.

CXCR3 expression has been most consistently associated with Th1 immune responses and Th1-associated diseases. Its agonists CXCL9-11 are highly induced by IFN-γ but not IL-4. "Th1 chemokines" help maintain Th1 dominance through their ability to block CCR3. Similarly, in Th2 immunity, IL-4 and IL-13 made at inflamed sites in the periphery induce production of CCL7, CCL11 and other CCR3 ligands, the CCR4 ligands CCL17 and CCL22, and the CCR8 ligands CCL1 and, in the mouse, CCL8. CCR3 is expressed on a subset of Th2 lymphocytes as well as on eosinophils and basophils, the three major cell types associated with Th2-type allergic inflammation. Th2 cells are also associated with CCR4 expression. Arrival of Th2 cells amplifies a positive feedback loop through secretion of additional IL-4. CCL7 and CCL11 can block Th1 responses by antagonizing CCR2, CXCR3, and CCR5. Th17 cells are all found within the CCR6$^+$ subset of CD4 T cells.[26] There may also be a positive feedback loop for IL-17 induction of the CCR6 ligand CCL20 as a mechanism to recruit additional Th17 cells to inflamed sites.

Tissue-Specific Lymphocyte Homing

Cutaneous lymphocyte-associated antigen (CLA$^+$) T lymphocytes, which home to skin, preferentially express CCR4 and CCR10.[24] The CCR4 ligand CCL22 is made by resident dermal macrophages and DCs, whereas the CCR10 ligand CCL27 is made by keratinocytes. Blocking both these pathways, but not either one alone, has been reported to inhibit lymphocyte recruitment to skin in a delayed-type hypersensitivity model. This implies that these two molecules act redundantly as well as independently of inflammatory chemokines.

Homing to the small intestine is determined, in part, by T-lymphocyte expression of the integrin $\alpha_4\beta_7$ and CCR9.[27] The $\alpha_4\beta_7$ ligand MAdCAM-1 and the CCR9 ligand CCL25 colocalize on normal as well as on inflamed endothelium of the small intestine. Most T cells in the intraepithelial and lamina propria zones of the small intestine express CCR9. These cells, which are mainly TCR$\gamma\delta^+$ or TCR$\alpha\beta^+$CD8$\alpha\beta^+$, are reduced in the small intestine from *CCR9*$^{-/-}$ mice.

As B cells differentiate into plasma cells, they downregulate CXCR5 and CCR7 and exit the lymph node. B immunoblasts expressing immunoglobulin G (IgG) coordinately upregulate CXCR4, which promotes homing to bone marrow, whereas B immunoblasts expressing IgA specifically migrate to mucosal sites. Like gut-homing T cells, B immunoblasts that home to the small intestine express $\alpha_4\beta_7$ integrin and CCR9 and respond to CCL25.

Our understanding of the chemokine-dependent mechanisms for leukocyte subset trafficking in most tissues remains rudimentary. There are hundreds of subsets that need to be evaluated in both healthy tissues and disease states. At stake is the potential identification of important new targets for therapeutic development of specific, safe, and effective mechanism-based drugs in human disease.

CHEMOKINES AND DISEASE

There is a vast amount of literature addressing the presence and potential clinical relevance of chemokines in human disease. This section provides only a sampling of this work. Diseases for which evidence is strongest for a clear role of chemokines in pathogenesis in humans are highlighted. In some cases, understanding the role of chemokines has led to specific therapeutic interventions.

Opposite Effects of CCR5 in HIV and West Nile Virus Infection

HIV envelope glycoprotein gp120 mediates fusion of viral envelope with the target cell membrane by binding to CD4 and a specific chemokine receptor, which is referred to in this context as an HIV coreceptor.[6] CCR5 and CXCR4, the most important HIV coreceptors, have been shown to be physically associated with CD4 and gp120. HIV viruses are classified into three main subtypes by preferred chemokine receptor usage: X4 strains, which use CXCR4 for entry; R5 strains, which use CCR5; and R5X4 strains, which can use either coreceptor. Coreceptor preference is the main determinant of distinct cytotropisms between primary macrophages and cultured T-cell lines that have been observed for these strains.

The importance of CCR5 in clinical HIV/AIDS is revealed most clearly by *CCR5Δ32,* a nonfunctional allele that occurs in ~20% of North American Caucasians. Homozygotes are highly resistant to R5 HIV, the main transmitting strain, and HIV-infected heterozygotes have slower disease progression. Homozygotes appear healthy, as do unstressed CCR5 knockout mice. This has led to the development and the US Food and Drug Administration (FDA) approval of the specific CCR5 antagonist maraviroc for treatment of patients with HIV/AIDS.[28] The clinical importance of CCR5 in HIV disease is also illustrated by the outcome of an HIV$^+$ patient from Germany who had leukemia and fortuitously received bone marrow transplantation from a *CCR5Δ32* homozygote after cytoreductive therapy for leukemia. Remarkably, this patient has remained well, with undetectable viral load without antiretroviral therapy, suggesting a functional "cure."[29] This has spurred on efforts to develop zinc finger nuclease, CRISPR (clustered regularly interspersed short palindromic repeats), and other genetic methods to inactivate the *CCR5* gene as a cure strategy in infected patients. Positive preliminary results have been reported for a CCR5 zinc finger nuclease.

CLINICAL RELEVANCE

Examples of Chemokine and Chemokine Receptor Determinants of Human Disease

- *WHIM syndrome:* Mendelian combined immunodeficiency disorder caused by gain-of-function mutations in *CXCR4*
- Plasmodium vivax *malaria:* Protection conferred by a nonfunctional promoter variant in ACKR1/Duffy that abrogates expression of the receptor on red blood cells preventing use of the receptor by the parasite for cell entry
- *HIV infection:* Protection from cell entry by the virus conferred by homozygous mutation *CCR5Δ32*
- *AIDS disease progression rate:* Slowed by heterozygous *CCR5Δ32*
- *West Nile virus disease:* Increased risk with homozygous *CCR5Δ32*
- *Kaposi's sarcoma:* Human herpesvirus 8 vGPCR
- *Age-related macular degeneration:* Increased risk with *CX3CR1 M280* allele
- *Cardiovascular disease:* Decreased risk with *CX3CR1 M280* allele
- *Autoimmune heparin-induced thrombocytopenia:* Caused by CXCL4 autoantibodies
- *Chronic renal allograft rejection:* Reduced risk with homozygous *CCR5Δ32*
- *Rheumatoid arthritis:* Reduced risk with *CCR5Δ32*
- *Eosinophilic esophagitis:* Associated with CCL26 variant

CLINICAL RELEVANCE

FDA-Approved Drugs Targeting the Chemokine System

- *Maraviroc:* CCR5 antagonist for human immunodeficiency virus/acquired immunodeficiency syndrome (HIV/AIDS) that works by blocking HIV target cell entry
- *Plerixafor:* CXCR4 antagonist indicated together with granulocyte–colony-stimulating factor (G-CSF) for hematopoietic stem cell (HSC) mobilization in autologous stem cell transplantation of patients with multiple myeloma and non-Hodgkin lymphoma receiving cytoreductive therapy

CCR5 is also important in the pathogenesis of West Nile virus (WNV) infection, but in this case, it plays a protective role.[30] The mechanism appears to be increased antiviral defense by increasing recruitment of CCR5$^+$ leukocytes to the brain. Theoretically, therefore CCR5 blocking agents could increase the risk of WNV disease in infected individuals, particularly in the setting of HIV/AIDS, where the immune system is already compromised.

Malaria

Plasmodium vivax also uses a chemokine receptor for target cell entry.[31] The parasite ligand, named the *Plasmodium vivax* Duffy Binding Protein (PvDBP), is expressed in micronemes of merozoites and binds to the *N*-terminal domain of ACKR1 (originally known as Duffy and later as DARC [Duffy Antigen Receptor for Chemokines]) on erythrocytes via a cysteine-rich domain. This interaction is required for junction formation during invasion, but not for initial binding or parasite orientation. *P. vivax*-malaria is rare in sub-Saharan Africa because of genetic deficiency in ACKR1/Duffy caused by a single nucleotide substitution in the gene promoter (–46 C) affecting an erythroid-specific GATA-1 site. Fixation of the mutation in Africa presumably occurred because of positive selective pressure from malaria. ACKR1/Duffy deficiency in humans and in *Ackr1* knockout mice is not associated with any known health problems. ACKR1/Duffy is an obvious drug target, but to date, no candidates have been reported.

WHIM Syndrome

Truncating mutations in the *C*-tail of CXCR4 that increase receptor signaling are the cause of most cases of WHIM syndrome, a rare disease characterized by *w*arts, *h*ypogammaglobulinemia, *i*nfection, and *m*yelokathexis (neutropenia without maturation arrest).[32] Myelokathexis and infection can be explained by exaggeration of the normal bone marrow retention function of CXCR4 for myeloid cells, inhibiting their egress to blood. The exact mechanism underlying the selective predisposition to human papillomavirus infection is unknown. Dialing down the pathologically increased CXCR4 activity with a specific CXCR4 antagonist is a direct and rational strategy for treatment of patients with WHIM syndrome. In this regard, clinical trials are in progress to repurpose plerixafor (Mozobil, AMD3100), a CXCR4 antagonist approved by the FDA for use with granulocyte colony-stimulating factor (G-CSF) to mobilize HSCs for autologous transplantation in patients with multiple myeloma or lymphoma undergoing cytoreductive therapy.[33,34]

Remarkably, a patient who was spontaneously cured of WHIM syndrome by chromothripsis (chromosome shattering) on chromosome 2, which harbors the *CXCR4* gene, has been identified.[35] Chromothripsis is a newly described and uncommon example of genomic instability first described in cancer. The chromothriptic event in the patient apparently occurred in a single HSC that not only survived and lost the WHIM allele (but not the normal allele of *CXCR4*) but also acquired a selective advantage that allowed it to repopulate bone marrow, resulting in correction of panleukopenia and clinical cure. This fortuitous experiment of nature has been recapitulated in experimental mice, indicating directly that CXCR4 haploinsufficiency promotes HSC engraftment and therefore may have been an important or even the sole mechanism of cure in the patient. Together, these findings point to a potential genetic cure strategy for WHIM syndrome that might also be applied more broadly as an adjuvant to aid engraftment of gene-corrected cells in other genetic diseases of blood.

Atherosclerosis

Macrophages are the dominant leukocytes present in atherosclerotic lesions and are associated with the presence of macrophage-targeted chemokines, such as CCL2, CCL5, and CX3CL1.[36] *Ccl2*$^{-/-}$, *Ccr2*$^{-/-}$, *Cx3cl1*$^{-/-}$, and *Cx3cr1*$^{-/-}$ mice on the atherogenic *ApoE*$^{-/-}$ genetic background demonstrate smaller lesions and reduced accumulation of macrophages in the vessel wall. Adoptive transfer studies with bone marrow from *Cxcr2*$^{-/-}$ mice have also revealed a role for CXCR2 in promoting atherosclerosis in mouse models, apparently by promoting monocyte adhesion to early atherosclerotic endothelium through interaction with its mouse ligand KC and activation of the VLA-4/VCAM-1 adhesion system.

The CX3CR1 genetic variant CX3CR1-M280, which lacks normal CX3CL1-dependent adhesive function under conditions of physiological flow, has been associated with reduced risk of atherosclerotic vascular disease in humans. Mechanistic studies suggest that CX3CL1 on coronary artery smooth muscle cells anchors macrophages via CX3CR1.

ON THE HORIZON

The Viral Chemokine System

Studies of the viral chemokine system can yield insight into the role of chemokines in immunoregulation. Viral chemokine elements may also have applications as therapeutic agents.

Structural Classification

- Chemokines (*e.g.,* HHV8 vMIP-I, II, and III)
- 7TM chemokine receptors (*e.g.,* HCMV US28, HHV8 vGPCR)
- Chemokine-binding proteins (*e.g.,* γHV68 vCKBP-III)
- Chemokine mimics (*e.g.,* HIV gp120)

Functional Classification

- Cell entry factors (*e.g.,* HIV gp120)
- Leukocyte chemoattractants (*e.g.,* HHV8 vMIP-II)
- Immune evasion
  - Chemokine scavengers (*e.g.,* γHV68 vCKBP-III)
  - Chemokine receptor antagonists (*e.g.,* MCV MC148-R)
- Angiogenic factors (*e.g.,* HHV8 vGPCR)
- Growth factors (*e.g.,* HHV8 vGPCR)
  - In mice, overexpression of this viral chemokine receptor induces Kaposi's sarcoma (KS)-like lesions.

Kaposi's Sarcoma

HHV8 is an example of a virus laden with genes encoding pirated chemokines and chemokine receptors. It encodes three CC chemokines, vMIP-I, -II, and -III, as well as a constitutively active CC/CXC chemokine receptor named vGPCR, encoded by ORF74. All of these factors are angiogenic and may contribute to the pathogenesis of Kaposi's sarcoma (KS), a highly vascular multicentric nonclonal tumor caused by HHV8, typically in the setting of immunosuppression, such as in HIV/AIDS. Consistent with this, vGPCR induces KS-like tumors when expressed in transgenic mice. The mechanism may involve activation of NF-κB and induction of angiogenic factors and proinflammatory cytokines. This virus appears to have converted a hijacked receptor, probably CXCR2, into a regulator of gene expression.[37]

Autoimmunity

Two human diseases in which chemokines act as autoantigens for autoantibodies have been identified. The first, heparin-induced thrombocytopenia (HIT), is the only human autoimmune disease directly linked mechanistically to chemokines.[38] An established risk factor for thromboembolic complications of heparin therapy, HIT occurs in 1–5% of patients treated with heparin and is the result of autoantibodies that bind specifically to CXCL4-heparin complexes in plasma. The second, autoimmune myositis, is associated with autoantibodies to histidyl transfer RNA (tRNA) synthetase, a protein synthesis factor that is also able to induce DC chemotaxis, apparently by acting as an agonist at CCR5. Its exact importance in promoting inflammation in myositis has not been established.

In general, T cell–dependent autoimmune diseases in human, such as psoriasis, multiple sclerosis (MS), rheumatoid arthritis (RA), and type 1 diabetes mellitus, are associated with inflammatory chemokines and tissue infiltration by T lymphocytes and monocytes expressing inflammatory chemokine receptors. In a mouse model of immune complex–induced arthritis, the specific contributions of Ccr1, Cxcr2, Blt1 (a leukotriene B4 receptor), and the C5a receptor have been dissected in detail in joint venules at the level of adhesion and transendothelial migration. A dominant negative antagonist of CCL2 inhibits arthritis in the MRL-*lpr* mouse model of RA, suggesting a potential role for CCL2 and CCR2. Met-RANTES, a chemically modified variant of CCL5 that blocks CCR1, CCR3, and CCR5, was shown to be beneficial in a collagen-induced arthritis model in DBA/I mice.

The relative importance of these in human disease is not known. Drugs targeting CCR1 have failed in clinical trials in RA. Likewise, although a gene association meta-analysis suggested a protective effect for *CCR5Δ32* in RA, the CCR5 antagonist maraviroc was ineffective in a clinical trial in patients with RA.

Paradoxically, in some cases, blocking of chemokine receptors can lead to increased inflammation, as shown for CCR1 and CCR2 in mouse models of nephrotoxic nephritis and glomerulonephritis. This is associated with increased renal recruitment of CD4 and CD8 T cells, macrophages, and enhanced Th1 immune responses. The mechanism remains unclear.

Acute Neutrophil-Mediated Inflammatory Disorders

Many neutrophil-mediated human diseases, including psoriasis, gout, acute glomerulonephritis, acute respiratory distress syndrome, RA, and ischemia reperfusion injury, have been associated with the presence of CXCL8. Systemic administration of neutralizing anti-CXCL8 antibodies is protective in diverse models of neutrophil-mediated acute inflammation in the rabbit (skin, airway, pleura, glomeruli), providing proof-of-concept that CXCL8 is a nonredundant mediator of innate immunity and acute pathological inflammation in these settings. CXCR2 knockout mice are less susceptible to acute urate crystal-induced gouty synovitis. In addition, SB-265610, a nonpeptide small molecule antagonist with exquisite selectivity for CXCR2, prevents neutrophil accumulation in the lungs of hyperoxia-exposed newborn rats. Together, these results identify CXCL8 and its receptors as candidate drug targets for diseases mediated by acute neutrophilic inflammation. CXCR2 knockout mice also have shown delayed wound healing. Interestingly, studies of *Candida albicans* and *Pseudomonas aeruginosa* infection in mice have suggested that the highly related neutrophil receptor Cxcr1 does not mediate neutrophil recruitment into infected sites but is, instead, involved in activation of antimicrobial effector mechanisms.

Transplant Rejection

In the case of transplant rejection, the advantage over other animal models of human disease is that in both human and animal situations, the time of antigenic challenge is precisely known. The most extensive analysis of the role of chemokines in transplant rejection has been carried out in an major histocompatibility complex (MHC) class I/II-mismatched cardiac allograft rejection model in the mouse, which is mediated by a Th1 immune response. Similar sets of inflammatory chemokines are found in the mouse model as in human disease and appear in a strict temporal sequence.

Analysis of knockout mice has demonstrated that multiple chemokine receptors contribute to rejection in this model but that there is a marked rank order: Cxcr3 > > Ccr5 > Ccr1 = Cx3cr1 = Ccr2. Most impressively, rejection and graft arteriosclerosis do not occur if the recipient mouse, treated with a brief, subtherapeutic course of cyclosporine, is *Cxcr3*$^{-/-}$ or if the donor heart is *Cxcl10*$^{-/-}$. This identifies the CXCR3/CXCL10 axis as a potential drug target. Neutralization of Cxcl9, a Cxcr3 ligand that appears later than Cxcl10, can also prolong cardiac allograft survival.

In humans, CCR5 may be important in chronic kidney allograft rejection, since individuals homozygous for *CCR5Δ32* are underrepresented among patients with this outcome in a large German kidney transplantation cohort.

Allergic Airway and Intestinal Disease

Chemokine receptors associated with asthma include CXCR2, CCR3, CCR4, and CCR8.[39] CCR3 is present on eosinophils, basophils, mast cells, and some Th2 T cells. CCR4 and CCR8 identify airway T cells of allergen-challenged patients with atopic asthma.

Ccr8 knockout mice have reduced allergic airway inflammation in response to three different Th2-polarizing antigens: *Schistosoma mansoni* soluble egg antigen, ovalbumin, and cockroach antigen.

A role for the CCR3 axis in asthma has been supported by Ccl11 neutralization in the guinea pig and *Ccr3* gene knockout in the mouse. The effect of *Ccr3* knockout depends dramatically on the specific method of sensitization and challenge because of complex and opposite effects on eosinophil and mast cell trafficking. Thus *Ccr3*$^{-/-}$ mice sensitized intraperitoneally have reduced eosinophil extravasation into the lung but increased mast cell homing to the trachea. The net result is a paradoxical increase in airway responsiveness to cholinergic stimulation. Mast cell mobilization is not seen after epicutaneous sensitization, and these animals have reduced airway eosinophilia on challenge and no increase in airway hyperresponsiveness.

Ccr6$^{-/-}$ mice have decreased allergic airway inflammation in response to sensitization and challenge with cockroach antigen, which is consistent with the induction of its ligand Ccl20 in this model. *Ccr6*$^{-/-}$ mice are also protected from a mouse model of psoriasis, which is related to Ccr6 expression on monocytes and IL-17 expression in γδ T cells. Eosinophilic esophagitis has been associated with a CCL26 variant. Although there is no CCL26 homologue in the mouse, other mouse CCR3 ligands have been implicated in a mouse model of this disease.

CCR9 is an attractive drug target in Crohn's disease because of its important role in T-cell homing to gut; however, a CCR9 antagonist entered into clinical trials in this disease did not meet the primary endpoint for efficacy.

Cancer

Many chemokines and leukocyte subtypes have been detected *in situ* in tumors, and cancer cells have been shown to produce chemokines and express chemokine receptors. However, the role played by endogenous tumor-associated chemokines in recruiting tumor-infiltrating lymphocytes and tumor-associated macrophages and in promoting an antitumor immune response has not been clearly delineated. On the contrary, there are data from mouse models suggesting that the overall effect may be to promote tumorigenesis through additional effects on cell growth, angiogenesis, apoptosis, immune evasion, and metastasis.[39] Controlling the balance of angiogenic and angiostatic chemokines may be particularly important. Chemokine receptors on tumor cells have been shown to directly mediate chemokine-dependent proliferation. In Japan, a monoclonal antibody mogamulizumab, which blocks CCR4, has been approved for the treatment of adult T-cell leukemia or lymphoma, which expresses this receptor at high levels.

THERAPEUTIC APPLICATIONS

Chemokines and Chemokine Receptors as Targets for Drug Development

Efforts to develop chemokine blocking agents over the past 20 years have so far led to FDA approval of two drugs: maraviroc, targeting CCR5 as a viral entry factor in HIV/AIDS; and plerixafor, targeting CXCR4 in HSC mobilization from bone marrow for transplantation in cancer patients (see above).[28,33] Chemokine receptors are the first cytokine receptors for which potent, selective nonpeptide small molecule antagonists have been identified that work *in vivo*. Targeting host determinants, as in the case of CCR5 in HIV/AIDS, is a new approach in the development of antimicrobial agents. Other candidate disease indications are ACKR1/Duffy in *P. vivax* malaria; CXCR4 in WHIM syndrome; CXCR2 in acute neutrophil-mediated inflammation; CXCR3 and CCR2 in Th1-driven disease; CCR2 and CX3CR1 in atherosclerosis; CCR3 and possibly CCR4 and CCR8 in Th2 diseases, such as asthma; CCR9 in inflammatory bowel disease; and CCR6 in Th17 diseases, such as psoriasis.

Potent and selective nonpeptide small molecule antagonists of chemokine receptors, including CXCR2, CXCR3, CXCR4, CCR1, CCR2, CCR3, CCR5, and CCR9 have been reported.[40] These molecules share a nitrogen-rich core and appear to block ligand binding by acting at a conserved allosteric site analogous to the retinal-binding site in the transmembrane region of rhodopsin. Although small molecules taken as pills are the main goal, other blocking strategies are also under consideration, such as (1) ribozymes; (2) modified chemokines (*e.g.,* amino-terminally modified versions of CCL5); (3) intrakines, which are modified forms of chemokines delivered by gene therapy that remain in the endoplasmic reticulum and block surface expression of newly synthesized receptors; (4) monoclonal antibodies; and (5) specific gene editing, using the zinc finger nucleases, TALENs and CRISPR/Cas9.

The fact that viral and tick antichemokines typically block multiple chemokines acting at multiple receptors suggests that the most clinically effective chemokine-targeted antiinflammatory strategy may need to provide broad-spectrum coverage. In this regard, viral and tick antichemokines themselves may have a place therapeutically, although issues of antigenicity may be limiting for chronic inflammatory disease applications.

Chemokines as Biological Response Modifiers

Both inflammatory and homeostatic chemokines are being evaluated for therapeutic potential as biological response modifiers, acting mainly as immunomodulators or as regulators of angiogenesis. Studies, to date have not revealed major problems with toxicity, and efficacy has been noted in models of cancer, inflammation, and infection. Clinical trials in cancer and stem cell protection have shown disappointing results. Chemokines are being developed as vaccine adjuvants as well. Chemokine gene administration has also been shown to induce neutralizing antibody against the encoded chemokine, which is able to block immune responses and to alleviate experimental allergic encephalomyelitis and arthritis in rodent models.

Many chemokines, when delivered pharmacologically as recombinant proteins, by plasmid DNA, or in transfected tumor cells, are able to induce immunologically mediated antitumor effects in mouse models and could be clinically useful. Mechanisms may differ, depending on the model, but may involve recruitment of monocytes, NK cells, and cytotoxic CD8 T cells to tumor. Chemokines could also be useful as adjuvants in tumor antigen vaccines. Chemokine–tumor antigen fusion proteins represent a novel variation of this approach, facilitating uptake of tumor antigens by APCs via the normal process of ligand–receptor internalization. Non-ELR CXC chemokines, such as CXCL4, also exert antitumor effects through angiostatic mechanisms. As a final point, the use of chemokine-based immunotoxins to target disease cell–associated chemokine receptors, including cells infected with viral chemokine receptors, is quite conceivable.

CONCLUSION

The chemokine system occupies a central place in immunoregulation and is an attractive source of potential drug targets for any disease with an innate or adaptive immune component. A basic outline of how the system works has been established by using mouse models, and there has been tangible progress in the application of this knowledge in clinical practice. Two major pioneering successes, CCR5 inhibition in HIV/AIDS and CXCR4 inhibition for HSC mobilization, have demonstrated the feasibility of targeting the chemokine system therapeutically in humans. However, both are eccentric indications that address unusual, niche roles of chemokine receptors. Satisfactory treatments for chronic immune-mediated diseases remain a major unmet medical need, and in this context, the chemokine system offers many drug targets, as well as difficult challenges for the future.

ACKNOWLEDGMENTS

This review was supported with funding from the Division of Intramural Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health (NIH).

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Kufareva I, Salanga CL, Handel TM. Chemokine and chemokine receptor structure and interactions: implications for therapeutic strategies. Immunol Cell Biol 2015;93:372.
2. Zlotnik A, Yoshie O, Nomiyama H. The chemokine and chemokine receptor superfamilies and their molecular evolution. Genome Biol 2006;7:243.
3. Bachelerie F, Ben-Baruch A, Burkhardt AM, et al. International Union of Basic and Clinical Pharmacology. LXXXIX. Update on the extended family of chemokine receptors and introducing a new nomenclature for atypical chemokine receptors. Pharmacol Rev 2013;66:1.
4. Ulvmar MH, Hub E, Rot A. Atypical chemokine receptors. Exp Cell Res 2011;317:556–68.
5. Slinger E, Langemeijer E, Siderius M, et al. Herpesvirus-encoded GPCRs rewire cellular signaling. Mol Cell Endocrinol 2011;331:179–84.

6. Berger EA, Murphy PM, Farber JM. Chemokine receptors as HIV-1 coreceptors: roles in viral entry, tropism, and disease. Annu Rev Immunol 1999;17:657–700.
7. Deruaz M, Frauenschuh A, Alessandri AL, et al. Ticks produce highly selective chemokine binding proteins with antiinflammatory activity. J Exp Med 2008;205:2019–31.
8. Schumann K, Lammermann T, Bruckner M, et al. Immobilized chemokine fields and soluble chemokine gradients cooperatively shape migration patterns of dendritic cells. Immunity 2010;32:703–13.
9. Butcher EC. Leukocyte-endothelial cell recognition: three (or more) steps to specificity and diversity. Cell 1991;67:1033–6.
10. Nourshargh S, Alon R. Leukocyte migration into inflamed tissues. Immunity 2014;41:694.
11. Thelen M, Stein JV. How chemokines invite leukocytes to dance. Nat Immunol 2008;9:953–9.
12. Mortier A, Gouwy M, Van Damme J, et al. Effect of posttranslational processing on the in vitro and in vivo activity of chemokines. Exp Cell Res 2011;317:642–54.
13. Martin C, Burdon PC, Bridger G, et al. Chemokines acting via CXCR2 and CXCR4 control the release of neutrophils from the bone marrow and their return following senescence. Immunity 2003;19:583–93.
14. Schulz O, Hammerschmidt SI, Moschovakis GL, et al. Chemokines and chemokine receptors in lymphoid tissue dynamics. Annu Rev Immunol 2016;34:203.
15. Charo IF, Ransohoff RM. The many roles of chemokines and chemokine receptors in inflammation. N Engl J Med 2006;354:610–21.
16. Vandercappellen J, Van Damme J, Struyf S. The role of the CXC chemokines platelet factor-4 (CXCL4/PF-4) and its variant (CXCL4L1/PF- 4var) in inflammation, angiogenesis and cancer. Cytokine Growth Factor Rev 2011;22:1–18.
17. Stillie R, Farooq SM, Gordon JR, et al. The functional significance behind expressing two IL-8 receptor types on PMN. J Leukoc Biol 2009;86:529–43.
18. Chou RC, Kim ND, Sadik CD, et al. Lipid-cytokine-chemokine cascade drives neutrophil recruitment in a murine model of inflammatory arthritis. Immunity 2010;33:266–78.
19. Geissmann F, Mass E. A stratified myeloid system, the challenge of understanding macrophage diversity. Semin Immunol 2015;27:353.
20. Griffith JW, Sokol CL, Luster AD. Chemokines and chemokine receptors: positioning cells for host defense and immunity. Annu Rev Immunol 2014;32:659.
21. Forster R, Davalos-Misslitz AC, Rot A. CCR7 and its ligands: balancing immunity and tolerance. Nat Rev Immunol 2008;8:362–71.
22. Muller G, Hopken UE, Lipp M. The impact of CCR7 and CXCR5 on lymphoid organ development and systemic immunity. Immunol Rev 2003;195:117–35.
23. Cyster JG. Chemokines, sphingosine-1-phosphate, and cell migration in secondary lymphoid organs. Annu Rev Immunol 2005;23:127–59.
24. Tubo NJ, McLachlan JB, Campbell JJ. Chemokine receptor requirements for epidermal T-cell trafficking. Am J Pathol 2011;178:2496–503.
25. Zhang HH, Song K, Rabin RL, et al. CCR2 identifies a stable population of human effector memory CD4β T cells equipped for rapid recall response. J Immunol 2010;185:6646–63.
26. Singh SP, Zhang HH, Foley JF, et al. Human T cells that are able to produce IL-17 express the chemokine receptor CCR6. J Immunol 2008;180:214–21.
27. Koenecke C, Forster R. CCR9 and inflammatory bowel disease. Expert Opin Ther Targets 2009;13:297–306.
28. Singh IP, Chauthe SK. Small molecule HIV entry inhibitors: part I. Chemokine receptor antagonists: 2004–2010. Expert Opin Ther Pat 2011;21:227–69.
29. Hutter G, Nowak D, Mossner M, et al. Long-term control of HIV by CCR5 Delta32/Delta32 stem-cell transplantation. N Engl J Med 2009;360:692–8.
30. Lim JK, Murphy PM. Chemokine control of West Nile virus infection. Exp Cell Res 2011;317:569–74.
31. Gaur D, Mayer DC, Miller LH. Parasite ligand-host receptor interactions during invasion of erythrocytes by *Plasmodium* merozoites. Int J Parasitol 2004;34:1413–29.
32. Kawai T, Malech HL. WHIM syndrome: congenital immune deficiency disease. Curr Opin Hematol 2009;16:20–6.
33. Pusic I, DiPersio JF. Update on clinical experience with AMD3100, an SDF-1/CXCL12-CXCR4 inhibitor, in mobilization of hematopoietic stem and progenitor cells. Curr Opin Hematol 2010;17:319–26.
34. McDermott DH, Liu Q, Velez D, et al. A phase 1 clinical trial of long-term, low-dose treatment of WHIM syndrome with the CXCR4 antagonist plerixafor. Blood 2014;123:2308.
35. McDermott DH, Gao JL, Liu Q, et al. Chromothriptic cure of WHIM syndrome. Cell 2015;160:686.
36. Koenen RR, Weber C. Chemokines: established and novel targets in atherosclerosis. EMBO Mol Med 2011;3:713.
37. Grisotto MG, Garin A, Martin AP, et al. The human herpesvirus 8 chemokine receptor vGPCR triggers autonomous proliferation of endothelial cells. J Clin Invest 2006;116:1264–73.
38. Cuker A. Recent advances in heparin-induced thrombocytopenia. Curr Opin Hematol 2011;18:315–22.
39. Allavena P, Germano G, Marchesi F, et al. Chemokines in cancer related inflammation. Exp Cell Res 2011;317:664.
40. Proudfoot AE, Bonvin P, Power CA. Targeting chemokines: pathogens can, why can't we? Cytokine 2015;74:259.

MULTIPLE-CHOICE QUESTIONS

1. The major signal transduction pathway triggered by chemokines includes:
   - **A.** Pertussis toxin sensitive heterotrimeric Gi proteins
   - **B.** Phospholipase C β
   - **C.** Diacylglycerol
   - **D.** Phosphoinositide 3 kinase
   - **E.** All of the above

2. Chemokine receptors are:
   - **A.** Rhodopsin-like G protein–coupled receptors
   - **B.** Localized in the nucleus
   - **C.** Encoded by certain bacterial species
   - **D.** Used by intracellular bacteria for cell entry

3. The chemokine system includes:
   - **A.** Chemokines
   - **B.** Chemokine receptors
   - **C.** Seven-transmembrane (7TM) domain chemokine-binding proteins
   - **D.** Chemokine-binding proteins that lack both chemokine and 7TM structure
   - **E.** All of the above

4. The chemokine system is involved in the pathogenesis of which of the following diseases?
   - **A.** WHIM syndrome
   - **B.** HIV/AIDS
   - **C.** Malaria
   - **D.** All of the above

5. The major homeostatic chemokine receptors include:
   - **A.** CCR7
   - **B.** CXCR3
   - **C.** CCR6
   - **D.** CCR1

11

Lymphocyte Adhesion and Trafficking

Sirpa Jalkanen, Marko Salmi

EARLY LYMPHOCYTE PRECURSOR TRAFFICKING TO THE PRIMARY LYMPHOID ORGANS

Lymphocyte trafficking begins at an early stage of human ontogeny when lymphocyte precursor cells first appear and migrate into the primary lymphoid organs.[1-3] The multipotent hematopoietic progenitor and stem cells from the yolk sac and from the aorta–gonad–mesonephros migrate via the circulation to the liver and the spleen, which are important organs that support lymphocyte production in the embryo (Chapters 2, 7), and then into bone marrow. Thereafter, the developmental maturation of B cells takes place solely in bone marrow. T cells, in contrast, require an additional migratory event, which involves the entry of marrow-derived T-cell progenitors into the thymus (Chapter 8).[4] These early T-cell progenitors enter the thymus via the vessels in the cortical region. Concomitant with their differentiation and maturation via positive and negative selection, they pass from the cortex into the medulla.

MIGRATION OF NAIVE MATURE LYMPHOCYTES FROM BLOOD TO THE SECONDARY LYMPHOID ORGANS

After completing their initial course of development, newly arisen naïve B and T cells exit the primary lymphoid organs, travel through blood, and extravasate selectively to the secondary lymphoid organs.[2,5,6] These include peripheral lymph nodes, organized lymphoid tissues of the gut (*e.g.,* Peyer patches and the appendix), and the spleen (Chapter 2). In lymph nodes, most lymphocyte trafficking from blood to tissues takes place in specialized postcapillary venules. The endothelial cells of the venule exhibit a characteristic high cuboidal morphology that has given them their name: high endothelial venules (HEVs). The protrusion of the surface of these endothelial cells into the vascular lumen promotes the interaction of leukocytes in the relatively low-shear venular part of the circulatory system with the endothelial surface membrane (Fig. 11.1). HEVs carry many unique adhesion molecules that enable the capture of passing lymphocytes. They also have special intercellular connections that facilitate penetration of the vessel walls by these emigrating lymphocytes. It has been estimated that more than 50% of incoming lymphocytes make transient contacts with the vascular lining in the lymph nodes and that as many as one passing cell in four adheres to the endothelium and then extravasates into the tissue.

KEY CONCEPTS

Lymphocyte Recirculation

- Lymphocytes recirculate continuously between blood and lymphoid organs.
- 80% of lymphocytes enter the lymph nodes via specialized vessels called high endothelial venules (HEVs).
- The remaining lymphocytes enter the lymph nodes together with dendritic cells (DCs) and antigens via afferent lymphatics.
- Lymphocytes leave the lymph nodes via efferent lymphatics.
- Lymphocyte recirculation allows the lymphocytes to meet their cognate antigens and other leukocyte subsets to evoke an efficient immune response.

Antigens gather into these secondary lymphoid organs by a different route. Most antigens in the periphery can be taken up by dendritic cells (DCs) (Chapter 6), which subsequently migrate into the secondary lymphoid organs via the afferent lymphatics.[2] These afferent lymph channels open into the subcapsular sinus of the lymph node. Individual DCs subsequently penetrate the lymphatic endothelium and migrate into the stroma. Unbound, or free, antigens that are being carried via the afferent lymphatics through the body can diffuse into these secondary lymphoid organs, and then they can be captured by the professional antigen-presenting cells (APCs) of the lymph nodes. Lymph nodes thus serve as traps for the immune system, collecting lymphocytes from blood and antigens from lymph (Fig. 11.2). In these organs, lymphocytes percolate through the tissue in search of their cognate antigens.[2,7] If a given lymphocyte does not find its antigen, it will leave the organ by entering the efferent lymphatics that drain the medullary sinuses and is then transported via a major lymphatic trunk, such as the thoracic duct, back into the large systemic veins. After reentering the circulation, the cell can randomly gain access to another lymph node, where it has another chance to extravasate into tissue and find its cognate antigen. One round of recirculation from blood to the lymph node stroma, to a lymphatic vessel, and then back to blood takes about 1 day. Naïve lymphocytes continue recirculating until they either find their cognate antigen or die.

ACTIVATED LYMPHOCYTES DISPLAY SELECTIVE TISSUE HOMING PATTERNS

In a secondary lymphoid organ, a successful encounter between a lymphocyte and its cognate antigen leads to the proliferation

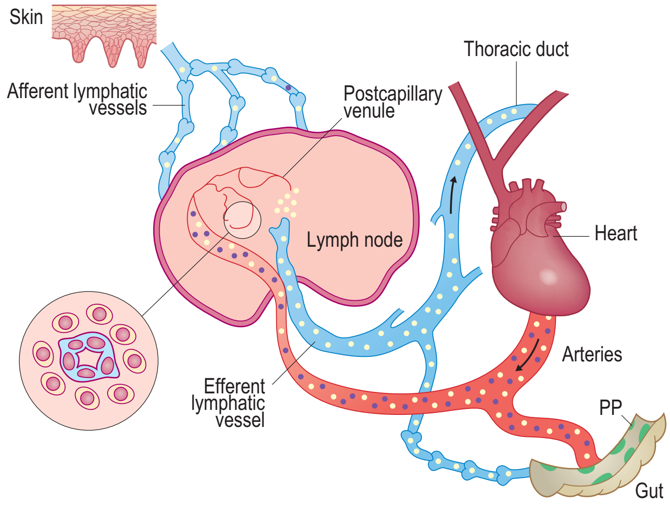


FIG 11.1 Lymphocyte Recirculation Routes Under Physiological Conditions. A low level of continuous antigenic transport into lymphoid organs takes place via the afferent lymphatics draining the skin and epithelium of the gut. Bloodborne lymphocytes enter the organized lymphatic tissues (lymph nodes and Peyer patches [PP]) from the circulation via the arterial tree, flow through the capillary bed, and then extravasate in the postcapillary high endothelial venules (HEVs). The extravasated lymphocytes percolate through the tissue parenchyma, enter the lymphatic vessels, and are then carried via the efferent lymphatics back to the systemic circulation. (Most of the venous circulation has been omitted from the figure.) Inset: an HEV. (From Salmi M, Jalkanen S. How do lymphocytes know where to go: current concepts and enigmas of lymphocyte homing. Adv Immunol 1997;64:139, with permission from Elsevier.)

of the cell within the organ and the maturation of its progeny.[2,6,7] Simultaneously, a process called *imprinting* leads to a profound change in the subsequent pattern of migration of the antigen-responsive cell. During the imprinting, local DCs give educational clues, such as vitamin A and D metabolites, to the lymphocytes. These lead to changes in the chemoattractant and adhesion receptor repertoire of the lymphocytes.[8] Although these responding cells leave the node via the lymphatics and are carried back to the systemic circulation, unlike naïve cells, they no longer migrate randomly to any lymphoid tissue. Instead, imprinting primes cells to preferentially seek the peripheral tissues in which the inciting antigen was originally ingested by the DC. In this way, selective homing of lymphocytes according to their previous history allows the organism to focus the immune response to the tissues where the effector cells can do the most good.

Among activated T cells, distinct pools of short-lived T effector cells, long-lived central memory T cells, effector memory T cells, and tissue resident memory T cells can be distinguished.[6,9] The different profile of adhesion molecule and chemokine receptor expression allows central memory T cells to continue migration through lymph nodes, whereas effector memory T cells are dispersed to patrol the peripheral tissues. In contrast, tissue resident lymphocytes, including tissue resident memory T cells, intraepithelial lymphocytes, γδ T cells, and innate lymphoid cells, mostly remain sessile in barrier tissues without the pattern of constant recirculation that is typical of the other lymphocyte subpopulations.

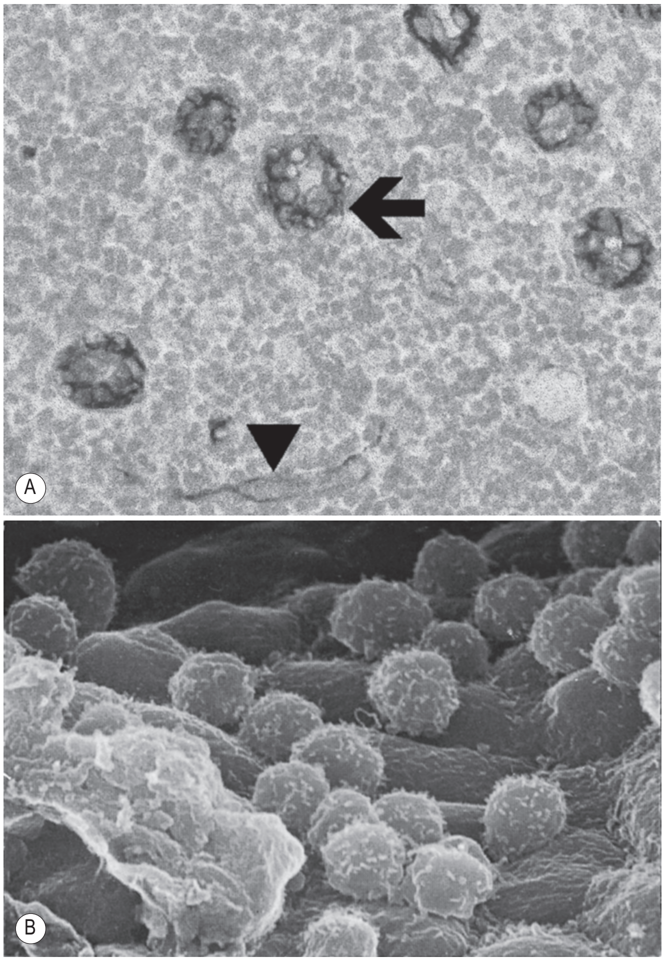


FIG 11.2 Characteristic Features of High Endothelial Venules (HEVs). (A) In this immunoperoxidase staining with anti-CD31 antibody, six HEVs are seen with typical plump endothelial cells. One HEV is identified with an arrow; a vessel with flat endothelium is also seen in this Figure *(arrowhead)*. (B) In this scanning electron micrograph lymphocytes are adhering to the HEV.

Under normal conditions, two distinct routes of lymphocyte recirculation can be discerned.[2,6] One targets lymphoid cells to the peripheral lymph nodes, and the second targets them to gut-associated lymphoreticular tissues (GALT; Chapter 20). Although the common gut-associated lymphoreticular system has long been considered to include both the respiratory and genitourinary tracts, differences in the fine specificity of lymphocyte homing between these targets do exist.

DISTINCT RECIRCULATION ROUTES IN THE SPLEEN

The spleen holds a unique place in the panoply of secondary lymphoid tissues.[10] It contains more lymphocytes than all of the peripheral lymph nodes combined, and the number of lymphocytes recirculating through it daily is the equivalent of the total pool of circulating lymphocytes. In the spleen, which lacks HEVs, many lymphocytes enter the tissue through the marginal zone sinuses, where macrophage-like cells may play an important role. Once in the splenic parenchyma, T cells accumulate in the regions surrounding the central arteriole in the white pulp, a location known as the *periarteriolar sheath*. B cells are scattered in the corona that surrounds these T-cell areas. Most splenic lymphocytes leave the spleen via the splenic vein. The mechanisms that control the entry and exit of lymphocytes from the spleen remain incompletely understood but include chemokines, oxysterols, and sphingosine 1 phosphate.

KEY CONCEPTS

Adhesion Molecules in Inflammation

- Adhesion molecules are important in directing leukocyte traffic to sites of inflammation.
- Numerous inflammatory mediators upregulate and/or induce expression of several endothelial cell adhesion molecules.
- Harmful inflammation can be prevented and cured by blocking the function of adhesion molecules.

INFLAMMATION-INDUCED CHANGES IN LEUKOCYTE TRAFFICKING

During an acute inflammatory response to an antigenic insult, leukocytes can migrate into all nonlymphoid. The inflammation-induced leukocyte migration takes place in characteristic waves.[11] First, the polymorphonuclear leukocytes rapidly (typically within 1–4 hours) infiltrate into the inflammatory focus. They are then followed by mononuclear cells (monocytes and lymphocytes). In a primary challenge, it may take 3 days or more before antigen-specific immunoblasts are seen at the peripheral site of inflammation. However, a secondary response by memory lymphocytes typically has a much shorter lag period. Different CD4 T-helper subpopulations, including Th1, Th2, Th17, and regulatory T cells (Tregs; Chapters 8, 16, and 18), CD8 T-cytotoxic cells (Chapter 17), and B cells (Chapter 7) can all enter the inflamed tissue by using basically the same mechanisms. However, the ratio of these populations and the individual molecules employed can vary.[6,12] Successful resolution of an inflammatory reaction is also dependent on a coordinated program involving specialized proresolving lipid mediators (lipoxins, resolvins, protectins, and maresins), proteins (annexin A1), purines (adenosine), and gaseous mediators (*e.g.*, hydrogen sulfide), all of which serve to halt the inflammatory cell recruitment and to initiate multiple antiinflammatory, tissue-repairing mechanisms.[13]

The normal vascular endothelium in nonlymphoid tissues has a flat, inactive morphology. With inflammation, a series of events renders postcapillary venules in these tissues capable of binding lymphocytes. The most important changes result from the proadhesive effects of a multitude of proinflammatory cytokines that are released by a variety of cell types after being subjected to inflammatory stimuli.[11] If inflammation becomes chronic, marked histological manifestations become apparent in the affected nonlymphoid tissues.[14] Most notably, the venules in these chronically inflamed tissues acquire many of the characteristics of HEVs. Immigrating lymphocytes can form lymphoid follicles that resemble those seen in lymph nodes. These alterations have consequences for lymphocyte recirculation pathways. For example, inflamed skin displays characteristics of lymphocyte homing that are clearly distinct from those of either mucosal or peripheral lymph node systems.[8,15]

MOLECULAR MECHANISMS INVOLVED IN LEUKOCYTE EXTRAVASATION FROM BLOOD INTO TISSUES

The Adhesion Cascade

Dynamic interactions between leukocytes and endothelial cells can be observed both *in vitro* and *in vivo*. For example, leukocyte adhesion can be followed *in vivo* in experimental animals and even in human tissues by intravital microscopy (Fig. 11.3). Leukocyte–endothelial cell interaction during the extravasation cascade can be divided into a series of phases, or steps, that all leukocyte subtypes, including lymphocytes, are thought to follow (Fig. 11.4).[2,6,11,16]

First, the leukocytes marginate out of the main bloodstream and begin to tether and roll on the endothelial cell surface. This step is mediated primarily by selectins and their mucin-like counterreceptors. This slow-velocity movement culminates in an activation phase, during which leukocyte chemokine receptors transmit activation signals by recognizing their chemokine ligands presented on the endothelial cell surface. This leads to avidity and/or affinity changes in leukocyte integrins, which bind leukocytes firmly to their immunoglobulin superfamily ligands on endothelial cells. Leukocytes then begin to crawl on the endothelium. After finding a proper place, they transmigrate through the endothelium. This transmigration process begins with interactions between leukocyte integrins, their counterreceptors, and other molecules. This step is followed by complex signaling events that lead to protein phosphorylation and dynamic clustering of the cytoskeleton.

Leukocytes transmigrate between endothelial cells at the endothelial cell junction, which opens transiently and subsequently closes by localized stimuli.[16] This process requires proteinases, such as matrix metalloproteinase-2 (MMP-2), which is induced in T cells upon adhesion to endothelial cells, as well as other repair mechanisms that are currently not well known. This subsequently closes the path of transmigration. Interestingly, leukocytes can also migrate through endothelial cells in a subtype-specific fashion. For example, polymorphonuclear leukocytes prefer entering via the interendothelial junctions, whereas nonactivated lymphocytes may choose the transcellular route.[11,16]

KEY CONCEPTS

Leukocyte–Endothelial Cell Interactions

- Leukocytes interact with the vessel wall in a multistep fashion, using several leukocyte surface molecules that recognize their counterreceptors on endothelial cells.
- Selectins mediate the rolling and tethering of leukocytes on the vessel wall.
- Chemokines and their receptors activate leukocyte integrins.
- Only activated integrins are able to mediate firm adhesion between leukocytes and endothelium.
- The transmigration of leukocytes into the tissues requires proteinases and repair mechanisms.

Certain endothelial molecules involved in the adhesion cascade show organ-specific expression patterns. Analogously, leukocyte-associated homing molecules display subtype-specific expression profiles. Only those leukocytes that have the proper set of molecules on their surface can enter a particular tissue because

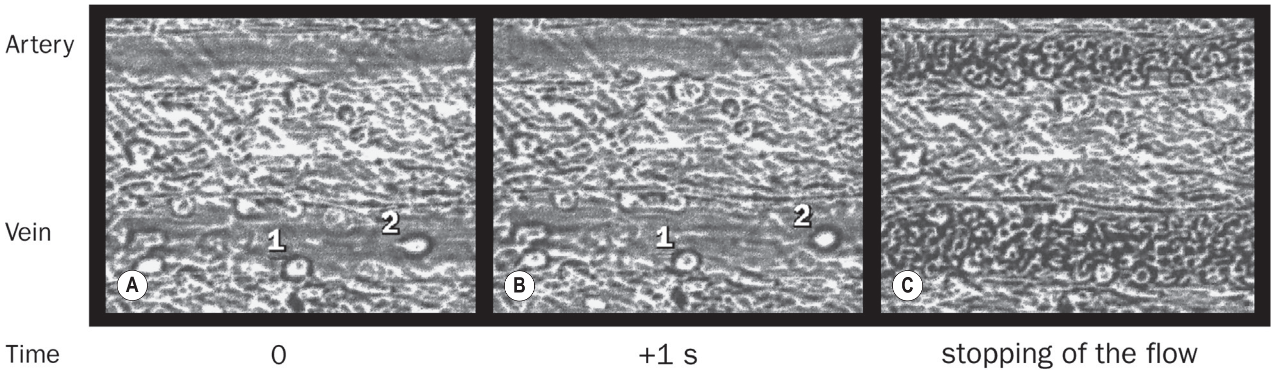


FIG 11.3 Intravital Microscopy of Mesenteric Vessels. In these video frames taken at indicated intervals from the same field, a vein, an artery, leukocytes, and the transparent mesenteric membrane are seen. Within the vein, rolling and adherent leukocytes are visible, whereas no such cells are seen in the artery. Leukocyte 1 is attached to the vessel wall and leukocyte 2 is slowly rolling. Compare the locations of these cells in A and B to stationary leukocytes outside the vessels. Leukocyte 1 is at the same location in both panels, whereas leukocyte 2 has moved a distance corresponding to roughly the length of its diameter. Under normal conditions, freely flowing cells move so fast that they cannot be visualized. However, in panel C, the flow has been transiently stopped. Under this static condition, the large number of hematopoietic cells (mainly erythrocytes) traveling within the bloodstream can be seen. (Courtesy of S. Tohka, University of Turku, Finland.)

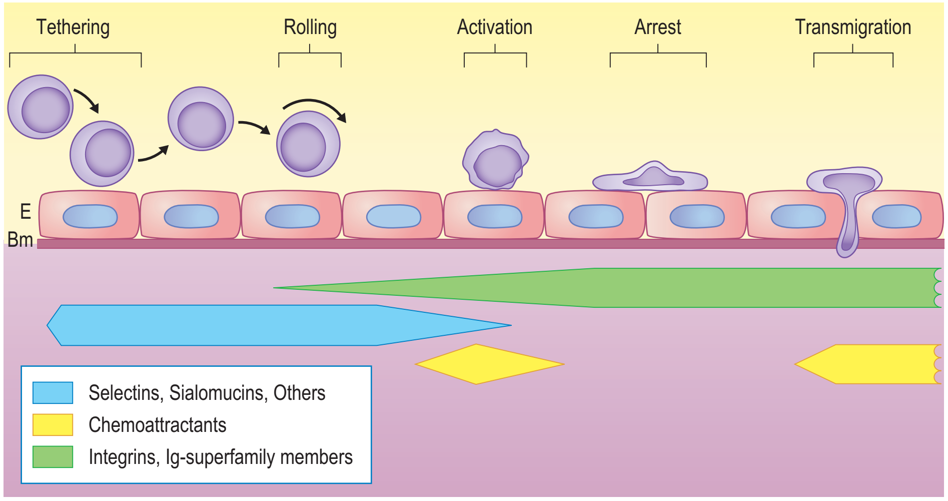


FIG 11.4 The Multistep Cascade of Lymphocyte Extravasation. The bloodborne cell makes transient initial contacts with endothelial cells (tethering), which leads to the cell rolling along the vascular lining. If the cell becomes activated, it can subsequently adhere firmly to the endothelial cells (arrest). The adherent cell can then penetrate through the endothelial cell cytoplasm (transcellular migration) or seek for interendothelial junctions (crawling). It can then migrate between two endothelial cells (paracellular migration) and finally penetrate the basement membrane to enter the tissue. The contribution of major superfamilies of adhesion-associated molecules at each step is depicted below. E, endothelial layer; Bm, basement membrane.

the entering leukocyte must find a correct endothelial partner molecule at each step of the adhesion cascade. Thus leukocyte–endothelial cell interactions take place in a well-coordinated multistep fashion, in which every step must be properly executed before the leukocyte can be guided into the tissue. The multistep nature of the leukocyte adhesion cascade is reminiscent of the cascades involved in blood clotting and complement-mediated killing.

Receptors and Their Ligands in Leukocyte–Endothelial Cell Interaction

Various molecules belonging to several molecular families are expressed on both leukocyte and endothelial cell surfaces and participate in the complex extravasation process.[16] Most of these molecules exert their function in successive, but overlapping, phases of the adhesion cascade. In addition, proper functioning

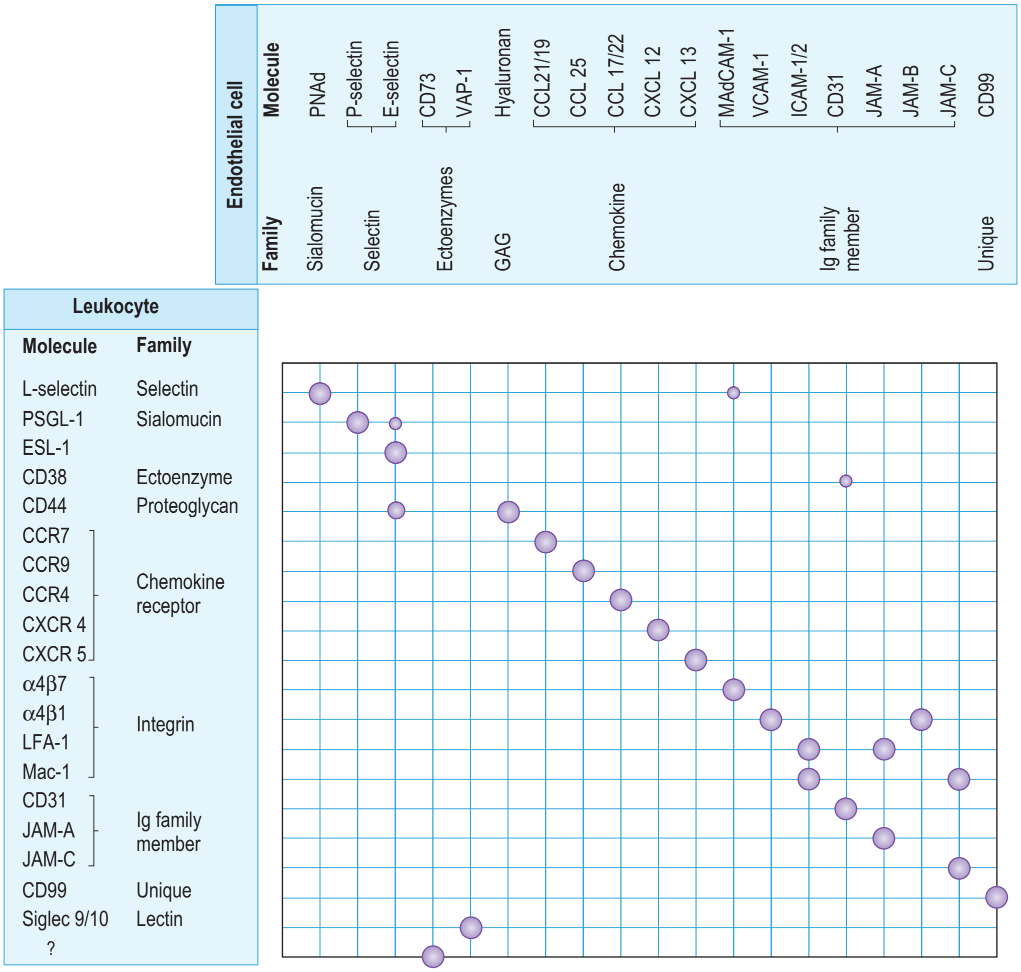

Endothelial cell
Family
Molecule
Sialomucin
PNAd
Selectin
P-selectin
E-selectin
Ectoenzymes
CD73
VAP-1
GAG
Hyaluronan
Chemokine
CCL21/19
CCL 25
CCL 17/22
CXCL 12
CXCL 13
Ig family member
MAdCAM-1
VCAM-1
ICAM-1/2
CD31
JAM-A
JAM-B
JAM-C
Unique
CD99
Leukocyte
Molecule
Family
L-selectin
Selectin
PSGL-1
Sialomucin
ESL-1
CD38
Ectoenzyme
CD44
Proteoglycan
CCR7
CCR9
CCR4
CXCR 4
CXCR 5
Chemokine receptor
α4β7
α4β1
LFA-1
Mac-1
Integrin
CD31
JAM-A
JAM-C
Ig family member
CD99
Unique
Siglec 9/10
Lectin
?

FIG 11.5 Adhesion Molecules Mediating Leukocyte Traffic. The most relevant proteins involved in leukocyte–endothelial cell interactions are shown as receptor–ligand pairs. GAG, glycosaminoglycan.

of multiple intracellular signaling molecules is needed for execution of successful extravasation.[2,16] In the following section, only the best-known surface molecules in this process are discussed (Fig. 11.5).

Selectins and Their Ligands

Three members of the selectin family mediate leukocyte trafficking. L-selectin is expressed on several leukocyte subpopulations. E-selectin is expressed on the endothelium, and P-selectin is expressed on both platelets and the endothelium. An important structural feature of selectins is the presence of a terminal lectin domain that is used to bind to their counterreceptors. The counterreceptor is typically decorated by a sialyl Lewis X (sLeX) carbohydrate, which is a prototype recognition domain for selectins in general.[17] The interaction between selectins and their counterreceptors is transient and weak, which allows leukocytes to form and break contacts with the endothelium effectively during tethering and rolling under shear stress.

L-selectin preferentially mediates lymphocyte migration to the peripheral lymph nodes. However, it also participates in the homing of lymphocytes to organized MALT (*e.g.*, Peyer patches) (Chapter 20). L-selectin is also an important contributor to the process of leukocyte trafficking to sites of inflammation. Peripheral lymph node addressins (PNAds) are the best-characterized counterreceptors for L-selectin. This group consists of at least six different molecules that are decorated with a sulfated and fucosylated sLeX, which serves as a recognition motif for L-selectin. PNAds include glycosylation-dependent cell adhesion molecule-1 (GlyCAM-1), CD34, podocalyxin, endomucin, nepmucin, and mucosal addressin cell adhesion molecule-1 (MAdCAM-1). MAdCAM-1 is decorated with recognition epitopes for L-selectin on HEVs in the organized lymphoid areas of the gut, but not on the flat-walled vessels in the lamina propria.

The importance of posttranslational carbohydrate modifications for the function of selectin ligands has been well demonstrated in knockout mice deficient in fucosyl, glucosaminyl, galactosyl, sialyl, or sulfotransferases. For example, mice with a targeted disruption of fucosyltransferase VII (Fuc-TVII) are unable to glycosylate L-selectin ligands appropriately. These mice exhibit severe defects in lymphocyte homing and in the extravasation of leukocytes to sites of inflammation. In contrast, core 2 β 1,6-N-acetylglucosaminyltransferase (C2 β GlcNAcT) knockout mice that have deficient glycosylation of their P- and E-selectin ligands demonstrate normal lymphocyte homing to lymph nodes, but impaired leukocyte trafficking to sites of inflammation. Interestingly, L-selectin also facilitates entry of leukocytes into tissues by mediating leukocyte tethering and rolling of endothelium-bound leukocytes by binding to P-selectin glycoprotein ligand-1 (PSGL-1).

E-selectin and P-selectin are inflammation-inducible molecules. Within minutes, P-selectin can be translocated from intracellular storage granules onto the endothelial cell surface, where it binds to its leukocyte receptor, PSGL-1 (see Fig. 11.5). P-selectin and PSGL-1 mediate rolling at early time points during inflammation. Platelet P-selectin can facilitate lymphocyte entry into tissues because it can simultaneously bind to PNAd on the endothelium and PSGL-1 on lymphocytes. E-selectin upregulation requires new protein synthesis. It is maximally expressed about 4 hours after induction of inflammation. E-selectin is needed for slow rolling of leukocytes. It also has an affinity toward PSGL-1, as well as a specific glycoform of PSGL-1, cutaneous lymphocyte antigen (CLA). CLA specifically directs lymphocyte trafficking to inflamed skin. In addition to PSGL-1, E-selectin has more private leukocyte receptors, CD44 and E-selectin ligand-1 (ESL-1) (see Fig. 11.5).[17] Mice deficient in both E- and P-selectins have more drastic defects in their rolling and leukocyte migration to sites of inflammation than could have been anticipated from mice deficient in only one of the two selectins. This suggests that E- and P-selectins have overlapping functions and can compensate for each other.

Chemokines and Their Receptors

Chemotactic cytokines and their receptors (Chapter 10) are grouped into four different families based on their primary protein structure. The families are defined by a cysteine signature motif—CXC, CC, C, and CX3C—where C is a cysteine, and X any amino acid residue.[18] Most chemokines are small, soluble heparin-binding chemoattractants. The relevant chemokines for leukocyte extravasation are presented to bloodborne lymphocytes by proteoglycan molecules present on endothelial cell surfaces. During the adhesion cascade, they activate integrins by signaling via serpentine receptors, which are pertussis toxin sensitive and G-protein linked.[19]

Different leukocyte subsets bear their own distinct sets of receptors, which enable them to respond to chemokines presented on the vascular endothelium as well as within tissues. For example, CCL21 and CCL19 are preferentially expressed by HEVs that are found in the interfollicular areas within a lymph node. They preferentially attract T cells bearing the CCR7 receptor and thus draw T lymphocytes from blood and into these areas. Fractalkine, a CX3C chemokine, can be produced in either a soluble or a membrane-anchored form. It can be found on the HEVs in peripheral lymph nodes and has potent chemoattractant activity for T cells. The major attractants for B cells are CXCL12 and CXCL13.

Although many chemokines are present in different organs in the body, some selectivity in their expression can guide tissue-selective leukocyte trafficking. For example, CCL25 attracts CCR9-positive lymphocytes to the small intestine, and CCL17 and CCL22 attract CCR4-bearing lymphocytes to the skin. Chemokines can form heteromeric complexes to activate chemokine receptors. For example, CXCL13 can enhance triggering of CCR7 by CCL19 and CCL21. Conversely, formation of chemokine complexes may protect a chemokine from degradation.[18]

Integrins and Their Immunoglobulin Superfamily Ligands

Integrins are a large family of heterodimeric molecules consisting of an α and a β chain.[2] Traditionally, they are thought to mediate firm adhesion between leukocytes and endothelial cells. However, in specific low-shear conditions, they can also participate in rolling. For leukocyte trafficking, the most important integrins are $\alpha_4\beta_7$, leukocyte function–associated antigen-1 (LFA-1), and $\alpha_4\beta_1$.

$\alpha_4\beta_7$ is a principal homing receptor for lymphocyte trafficking to mucosa-associated lymphatic tissues. It binds to MAdCAM-1 on HEVs in organized mucosal lymphatic tissues, such as Peyer's patch and the appendix, and to flat-walled venules in the lamina propria. α4-integrin can also pair with a β_1 chain to form $\alpha4\beta_1$ dimers, which are utilized by lymphocytes primarily in inflammatory conditions. It binds to vascular cell adhesion molecule-1 (VCAM-1) on the endothelium and has been proven central in mediating lymphocyte trafficking into the brain in multiple sclerosis.

LFA-1 (CD11a/CD18) is a member of the group of leukocyte integrins that contain a unique α chain (CD11 a, b, c, and d) but share a common β chain (β_2/CD18). LFA-1 is present on practically all leukocyte subsets. It interacts with its counterreceptors, intercellular adhesion molecules (ICAM-1 and ICAM-2), on the endothelial cell surface (see Fig. 11.5). ICAM-1 is upregulated at sites of inflammation, whereas ICAM-2 is constitutively present on the vascular endothelium. To be functional, LFA-1 must be activated. Activation of LFA-1 is thought to be primarily a product of chemokine signaling. Alternative activation pathways include triggering through glycosyl-phosphatidylinositol (GPI)–linked molecules and CD44, as well as through other costimulatory lymphocyte surface molecules.[20] LFA-1–dependent pathways display no significant organ specificity in their function.

Mac-1 (CD11b/CD18) is also involved in leukocyte migration, although its contribution is overshadowed by LFA-1. Like LFA-1, Mac-1 uses ICAM-1 and ICAM-2 as its ligands (see Fig. 11.5). Both VCAM-1 and ICAM-1 also play an important role in initiating transmigration, as they mediate the signals to endothelial cells that are needed to change their shape and other properties in order to allow leukocyte entrance.[20]

Other Homing-Associated Molecules

Several other molecules belonging to various molecular families also participate in the adhesion cascade. CD44 is a multifunctional proteoglycan found on a large variety of different cell types.[21] Using endothelial hyaluronan as its ligand, it mediates lymphocyte rolling. It can form bimolecular complexes with $\alpha_4\beta_1$ that strengthen leukocyte–endothelial cell interaction. *In vivo* inhibition studies using function-blocking antibodies indicate that CD44 plays an important role in directing lymphocyte trafficking to sites of inflammation (*e.g.*, skin and joints).

CD31, a member of the immunoglobulin superfamily, is found on many subsets of lymphocytes as well as in the continuous endothelium of all vessel types. It is expressed primarily at the intercellular junctions and is involved in a stimulus-specific manner in transmigration, especially through the endothelial basement membrane. Other molecules involved in the transmigration process are CD99 and junctional adhesion molecules A and C, which are expressed on both leukocytes and endothelial cells. These molecules interact sequentially in a homotypic fashion during diapedesis. Endothelial JAM-A can also use LFA-1, JAM-C can utilize Mac-1, and JAM-B can use $\alpha_4\beta_1$ as leukocyte ligands.

The role of ecto-enzymes in the adhesion cascade has been recently recognized.[22,23] Vascular adhesion protein-1 (VAP-1), CD73, and CD38 have well-established roles in leukocyte trafficking. Because of their enzymatic properties, they can rapidly modify adhesive interactions and modulate the microenvironment. VAP-1 is a homodimeric sialoglycoprotein, which, under conditions of inflammation, is rapidly translocated onto the endothelial cell surface. It mediates early phases of leukocyte interaction with the endothelium, as well as transmigration. Besides its adhesive function, it also possesses an amine oxidase activity that can produce potent immunomodulators, such as hydrogen peroxide (H_2O_2) and aldehyde as end products.

CD73 is present both on a subpopulation of lymphocytes and on the endothelium. It is an ecto-nucleotidase. The main product of its enzymatic activity in dephosphorylation of adenosine monophosphate (AMP) is adenosine, which is highly anti-inflammatory in nature. Endothelial CD73 may also have a counterreceptor on the lymphocyte surface because lymphocyte binding to the endothelium inhibits the enzymatic activity of CD73. This facilitates the extravasation process of the lymphocyte. The importance of CD73 in dampening inflammation is clearly seen in mice deficient in CD73, as they are highly susceptible to inflammatory injuries as a result of leaky vasculature.

CD38, an adenosine diphosphate (ADP)–ribosyl cyclase, is expressed on most lymphoid cells and can use CD31 as its endothelial cell ligand. It regulates calcium fluxes and the sensitivity of leukocytes to chemokine signals via its enzymatic activity.

INTRAORGAN LYMPHOCYTE LOCALIZATION

Following its extravasation from the blood vessels, a lymphocyte needs to interact with several matrix molecules, such as fibronectin, laminin, and collagens. Adhesive interactions between a lymphocyte and the extracellular matrix molecules are largely mediated by β_1 integrins, although lymphocytes can also use CD44 to interact with fibronectin and collagens. The directional movement and the final localization of a lymphocyte within the tissues are controlled by chemokines.[24] Modern two-photon imaging has provided detailed information about the kinetics and directionality of lymphocyte movement in living tissues.[7]

Besides directing the entry of T cells into tissues, CCL21 and CCL19 also determine the final destination of the T lymphocyte within lymph nodes, the spleen, and Peyer's patches. CCL19 and CCL21 produced by stromal cells in the T cell areas within lymphoid tissues guide the T lymphocyte into the interfollicular space. In an analogous manner, CXCL13, the B cell–attracting chemokine-1, is produced by a subset of follicular DCs found in the secondary lymphoid organs. It attracts B cells possessing CXCR5 to the light zone of the follicles. CXCL12, in contrast, guides CXCR4 positive B cells to the dark zone (Fig. 11.6).

The correct localization of a lymphocyte and an APC within the lymph node is a prerequisite for an optimal immune response.

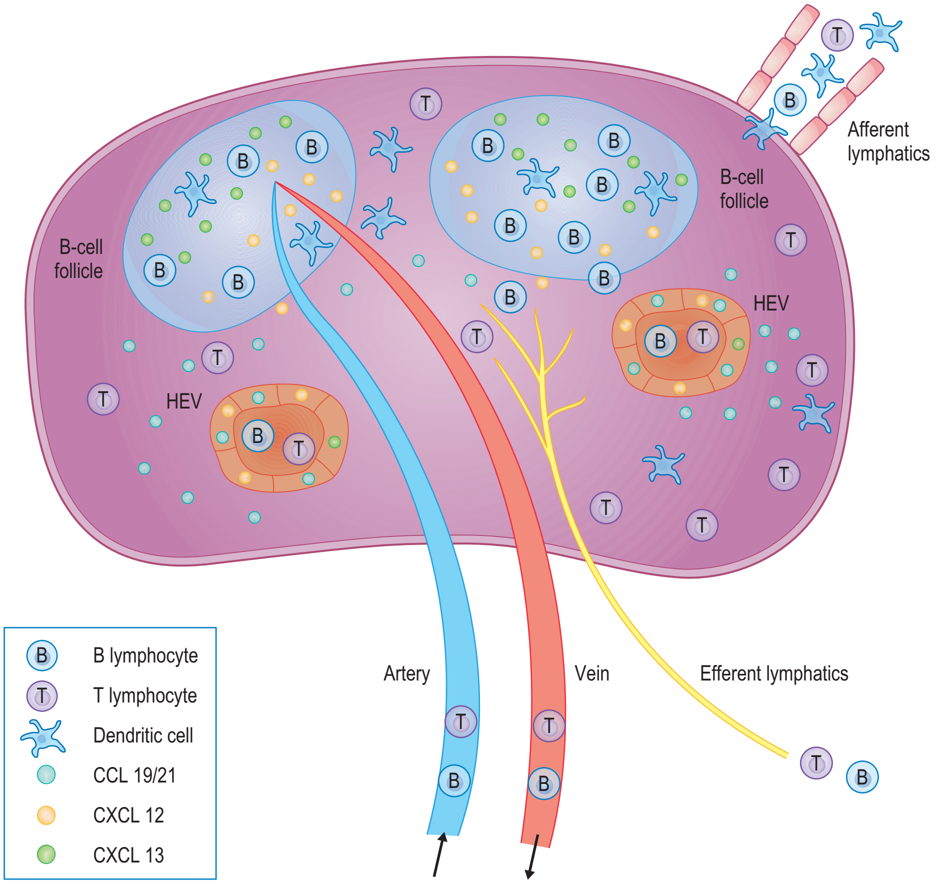


FIG 11.6 The Role of Chemokines in Entrance and Intraorgan Localization of Lymphocytes. CCL19 and CCL21 are involved in T-lymphocyte entrance into the lymph node via the high endothelial venules (HEVs). They also guide migration of T cells to the interfollicular areas within the node. In contrast, CXCL12 and CXCL13 attract B lymphocytes on the vessel wall and direct their movement to the follicles.

For example, B-cell collaboration with T cells is ensured by the upregulation of CXCR5 expression on a subpopulation of T cells and the upregulation of CCR7 on certain B cells. This promotes the movement of these cells to the B- and T-cell zone boundary.

The importance of CCL21 and CCL19 in lymphocyte trafficking is demonstrated by a spontaneous mutant mouse strain, plt, which has reduced expression of these chemokines. In these mice, both lymphocyte entry via HEVs and the organization of T-cell areas within lymph nodes are defective. This phenotype is recapitulated in CCR7-knockout mice. CCL21/CCL19 and CXCL12/CXCL13 are currently the best-known determinants of lymphocyte localization within lymphoid tissues, although several other chemokines are needed for optimal encounters among various cell populations to create a crisp immune response. In addition to chemokines, other attractants (*e.g.,* oxysterols) guide lymphocytes within the nodes. For example, Epstein-Barr virus induced genes in B cells (EBI2) have a strong migratory response to 7α,25-dihydroxycholesterol, promoting B-cell positioning to the lymphoid follicles.[25]

CELL TRAFFICKING WITHIN LYMPHATICS

Although leukocyte trafficking within the lymphatics is an essential part of the recirculation process and the immune response in general, the molecular mechanisms that regulate this tightly controlled and cell type–selective migration remain poorly understood. It is known that CCL21 on the lymphatic endothelium attracts CCR7-positive lymphocytes into the afferent lymphatics and hence to the draining lymph nodes.[26] Adhesion molecules, such as the macrophage mannose receptor and the common lymphatic endothelial and vascular endothelial receptor-1 (Clever-1, also known as "Stabilin-1" and "FEEL-1"), also mediate lymphocyte migration into the lymphatics. Lymphocytes adhere to lymphatic vessels via these molecules in *in vitro* assays, and inhibition and/or lack of these molecules impair lymphocyte trafficking within the lymphatics *in vivo.* Both macrophage mannose receptor and Clever-1 are present on the afferent and the efferent lymphatics and, thus, are able to control both incoming and outgoing lymphocyte traffic in lymph nodes.[27,28]

Sphingosine 1 phosphate receptor 1 (S1P1), which is expressed in many different cell types, also participates in the entry of lymphocytes into the afferent lymphatics and the exit of lymphocytes from lymph nodes into the efferent lymphatics. Although its absence leads to defects in vascular development and embryonic lethality, studies of conditional knockout mice have shown that lymphocytes lacking S1P1 can enter lymph nodes normally. However, they are not able to exit from lymph nodes into the efferent lymphatics. Additional studies indicate that S1P1 may control lymphocyte trafficking by regulating endothelial permeability and/or suppressing lymphocyte chemotaxis in S1P gradients and be a useful target for drug development in many diseases.[29]

CLINICAL IMPLICATIONS

Leukocyte trafficking plays a pivotal role in the pathogenesis of all infectious and inflammatory diseases.[16] Trafficking is essential for mounting a proper immune response against an invading microbe. However, in many other cases, inappropriate leukocyte migration also causes tissue destruction. Here, we have chosen a few representative examples to illustrate some of the general principles.

TABLE 11.1 Clinical Features of Immunodeficient Patients With Abnormalities in Homing-Associated Molecules

	LAD I	LAD II	LAD III
Clinical Manifestation			
Recurrent bacterial infections	+++	+	+++
Neutrophilia	Yes	Yes	Yes
Developmental abnormalities	No	Yes	No
Laboratory Findings			
CD18 expression	Marked decrease or absence	Normal	Normal
sLex expression	Normal	Absent	Normal
Neutrophil rolling	Normal	Marked decrease	Normal
Neutrophil firm adherence	Marked decrease	Mild decrease	Marked decrease
Primary detect	CD18	FUCT1	KINDLIN3

Immunodeficiencies

Various genetic defects in leukocyte trafficking have been identified during the past two decades among a multitude of other immunodeficiencies.[30] Leukocytes from patients with Wiskott-Aldrich syndrome (WAS) show reduced migration to the chemokines CCL2, CCL3, and CXCL12 as a result of a mutation in the gene encoding an intracellular WAS protein that is responsible for proper cytoskeletal organization of the hematopoietic cells. Mutations of CXCR4 in patients with WHIM (warts, hypogammaglobulinemia, infections, and myelokathexis) lead to an enhanced chemotactic response to CXCL12 expressed on the bone marrow endothelium and leukocyte accumulation in bone marrow. A single case of inherited dysfunction of E-selectin has also been reported. The patient had recurrent infections but did not demonstrate neutrophilia. She could not synthesize E-selectin on the endothelium, although her serum level of soluble E-selectin had increased.[31]

Three different forms of adhesion defect in integrins lead to poor or total lack of homing of certain leukocyte populations.[32] Their typical features are listed in Table 11.1. These diseases are extremely rare, but they offer a perfect recapitulation in human of the basic principles of leukocyte trafficking that was first deciphered in animal models.

LAD I

Patients with leukocyte adhesion defect-1 (LAD-1) have defects in the synthesis, pairing, or expression of β_2 integrins.[33] The severe form leads to a drastic reduction or absence of CD11a/CD18 (LFA-1), CD11b/CD18 (Mac-1), CD11c/CD18, and CD11d/CD18. Although patients with LAD-1 exhibit neutrophilia, they suffer from severe recurrent bacterial infections as a result of defects in leukocyte extravasation into sites of inflammation. These patients usually die young. Patients suffering from a milder form of LAD-1 (<10% of normal β_2 integrins) demonstrate impairment in leukocyte migration and suffer from frequent infections as well, but they usually survive until adulthood. In these patients, lymphocyte migration to inflammatory sites is close to normal, probably because lymphocytes can utilize the VLA-4–VCAM-1 pathway to compensate for the lack of β_2 integrins. LAD-1 variants with modestly reduced β_2 integrin expression have been associated with mutations that create heterodimers but do not bind ligands. In all types of LAD I, granulocytes typically show normal rolling, but failure to arrest on endothelial cells.

LAD II

Patients with LAD II suffer from several abnormalities, including recurrent infections, neutrophilia, impaired pus formation, mental retardation, growth defects, and deficiency in H blood group antigen.[33] The cause of the disease is impaired transport of guanosine diphosphate (GDP)–fucose from the cytoplasm to the Golgi lumen as a result of mutations in a GDP–fucose transporter. The defect leads to impaired fucose modification of selectin ligands, most notably that of PSGL-1. Consequently, these patents demonstrate a marked decrease in leukocyte rolling under flow conditions. Interestingly, some patients respond to treatment with oral fucose.

LAD III

The third LAD syndrome involves defective inside-out signaling of β_1, β_2, and β_3 integrins.[33] In the leukocytes of these patients, activation of multiple leukocyte integrins, including LFA-1 and VLA-4, by chemokine-triggered, G protein–coupled signals is genetically defective. Concurrent defects in platelet integrins lead to a bleeding diathesis. LAD III is caused by mutated Kindlin3, which is a major cytoplasmic activator of integrins.

Autoimmune or Inflammatory Diseases

Multiple Sclerosis

Migration of lymphocytes through the blood–brain barrier into the central nervous system (CNS) is a key pathogenic event that occurs during the development of multiple sclerosis (MS) and its widely used animal model, experimental allergic encephalomyelitis (EAE)[34] (Chapter 66). Under normal conditions, lymphocytes only patrol outside of the CNS parenchyma in the leptomeningeal and perivascular compartments. Naïve lymphocytes inefficiently bind to endothelial cells in meningeal microvessels without a rolling step using α_4 integrin and then extravasate through the fenestrated choroid vessels using P-selectin and CCR6. Upon inflammation, meningeal and choroid vessels become more adhesive, and the endothelial cells within the CNS parenchyma start to express ICAM-1 and VCAM-1. These alterations allow effector lymphocytes, which have primarily been activated in the cervical lymph nodes that ultimately drain the CNS-derived antigens, to penetrate the blood–brain barrier and enter the CNS parenchyma.

In vivo animal studies have shown that the disease course can be dramatically altered by blocking the function of leukocyte α_4 integrin.[34] This treatment is able to prevent the disease and even reverse the paralytic disease and lymphocytic infiltrations in the brain. Antibodies against $\alpha_4\beta_1$ integrin show this remarkable efficacy even if started a month after onset of the clinical disease. Certain EAE models can also be blocked by targeting the α_4 integrin ligand VCAM-1 or by blocking the function of CD44. In contrast, inhibition of many other adhesion molecules, such as L-selectin, PSGL-1, and E- and P-selectins, fails to yield a consistent effect on EAE. It should also be noted that although $\alpha_4\beta_1$ integrin is very important for lymphocyte homing to brain, it is not a brain-specific destination. It is also involved in leukocyte trafficking to other organs, such as the pancreas and the gut.

Inflammatory Bowel Diseases

As described above, effector lymphocytes activated in the inductive sites (Peyer's patches and mesenteric lymph node) of the intestine selectively migrate to the effector sites of the gut (mainly the lamina propria).[35] The key tissue-specific adhesion code for the intestinal entry is $\alpha_4\beta_7$ integrin on lymphocytes and its endothelial counterreceptor MAdCAM-1. Although the canonical CCR9-CCL25 interaction is important for attracting T and B lymphocytes to the small intestine, the critical chemotactic receptors for the entry into the large intestine are GPR15 (on T cells except for Tregs) and CCR10 (on plasmablasts).

In humans, the homing of regulatory T cells to the colon is GPR15 independent. This creates pathways by which the trafficking of proinflammatory versus antiinflammatory lymphocyte subtypes in the large bowel can be altered. In inflammatory bowel disease (IBD), the expression of MAdCAM-1 is upregulated in the vessels of lamina propria, together with other vascular adhesins, such as VCAM-1 and ICAM-1. Increased expression of CCL20 in the IBD mucosa enhances the recruitment of both T and B cells to the inflamed gut through ligation of CCR6. Successful pharmacological and genetic targeting of these key gut-selective lymphocyte trafficking alleviates inflammation in multiple animal models of IBD.

ON THE HORIZON

Adhesion Molecules as Targets for Novel Diagnostic and Therapeutic Agents

- Use of adhesion molecules as targets for the diagnosis of immune deficiencies, inflammatory disorders, and cancers.
- Use of adhesion molecules as targets for imaging inflammatory responses.
- Use of adhesion modulating therapies and other manipulations of lymphocyte trafficking, including small molecules, to treat infections, immune deficiencies, autoimmune disorders, transplant rejection, ischemia-reperfusion injuries, and cancers.

Cancer

Immunoevasion is one of the main hallmarks of cancer, and insufficient lymphocyte trafficking to solid tumors contributes to tumor progression. In the neoangiogenic vessels of tumors, immature and disorganized endothelial cells often express insufficient amounts of adhesion and attraction molecules.[36] Moreover, tumor-derived factors often render the local endothelial cells unresponsive to inflammatory cytokines. Nevertheless, chronic inflammation in tumors can induce the formation of HEV-like vessels, which show augmented expression of trafficking molecules.

Lymphocyte migration can be harnessed to improve the outcome of cellular immunotherapy for cancer.[36] Infusion of activated or genetically modified effector T cells, and cytotoxic CD8 T cells in particular, would benefit from a more effective targeting of the cells into the tumor. Homing of CAR-T cells to the tumors could be enhanced by transduction with adhesion molecules and chemokine receptors, possibly with simultaneous induction of tumor vessel normalization. However, anti-adhesive therapies targeting inappropriate accumulation of Tregs in the tumor would also enhance anti-cancer immune responses.

THERAPEUTIC PRINCIPLES

- Proadhesive strategies have been mainly developed using gene therapy.
- Inappropriate inflammation associated with many diseases can be dampened by antiadhesive therapeutics.
- Function blocking monoclonal antibodies (mAbs) are effective antiadhesive molecules.
- Cell migration can also be blocked by ligand and receptor analogues, small-molecule inhibitors, and genetic means (*e.g.,* RNA interference).

Adhesion Molecules as Diagnostic Targets

Immunodeficiency Disorders

In classic LAD I deficiency, a definitive diagnosis can be established by flow-cytometric analysis of immunofluorescent-stained CD18 integrin on blood leukocytes.[32] In individuals without the deficiency, practically all peripheral blood lymphocytes are CD18 positive. In severe cases of LAD I deficiency, there is a complete absence of this adhesion molecule, whereas in moderate deficiency about 1–10% of control levels is seen. LAD II can be diagnosed with commercially available antibodies by documenting a lack of fucosylated adhesion molecules (*e.g.,* abnormal expression of sLeX on blood leukocytes). Diagnosis of LAD III requires activation of cells, the use of monoclonal antibodies (mAbs) that recognize specific activation epitopes, or functional assays.

Soluble Adhesion Molecules

Most adhesion molecules are found in soluble forms in body fluids. They can be generated by alternative splicing of messenger RNA (mRNA) with deletion of transmembrane anchors, by proteolytic cleavage of cell-bound proteins by various sheddases (proteases), or by cleavage of the GPI linkage. Either of the soluble forms can function as molecular sinks, competing for their specific ligand(s) with the membrane-bound forms. Alternatively, they can trigger biological responses by interacting with their ligand-bearing cells.

The availability of commercial kits for measuring the levels of soluble adhesion molecules has led to numerous reports describing an increase or a decrease of certain adhesion molecules in different diseases. In most cases, there appears to be little or no additional diagnostic or predictive value derived from the use of these tests compared with more traditional parameters of inflammatory activity. Therefore at present, the indications for measurements of soluble adhesion molecules, or chemokines, in inflammatory diseases and cancers remain to be defined.

Imaging

The use of neutrophil scans or radioactively labeled nonspecific molecules to localize inflammatory foci is not satisfactory in terms of timing, expense, biohazards, specificity, or sensitivity. Hence, there have been trials to radioactively or nonradioactively label mAbs that recognize endothelial adhesion molecules, infuse them intravenously, and monitor their accumulation by appropriate imaging devices. Inflammation-inducible molecules, such as E-selectin and MAdCAM-1, have been used as target antigens. In the case of radioactively labeled E-selectin antibodies, the utility of this approach has been verified in patients.

CLINICAL PEARLS

- Blocking of α_4 integrin with natalizumab ameliorates disease activity in multiple sclerosis (MS) and in Crohn disease.
- Blocking of α_4/β_7 with vedolizumab is an effective treatment for inflammatory bowel disease (IBD).

Therapeutic Applications of Adhesion Modulating Therapies

Pro- and anti-adhesive therapies have long been an obvious pharmaceutical goal in the field of leukocyte trafficking. Inflammation-promoting strategies would be beneficial for treating many immune deficiencies, persistent infections, or cancers. Proadhesive control of lymphocyte traffic would benefit specific areas, such as vaccine development and bone marrow cell transplantation. In practice, lymphocyte recirculation routes are already being empirically exploited, for example, by varying the anatomical site (skin vs. intestine) of vaccination to optimize the immunization response. Anti-adhesive therapy, on the other hand, can be seen as a form of treatment applicable to all disease categories that involve an inflammatory component. In addition, it could provide novel precision drugs individually tailored for treating organ-specific inflammatory disorders, which would be predicted to diminish the problems of generalized immunosuppression.

Antibodies and Small-Molecular Drugs

A number of specific antagonists of adhesion molecules have been developed.[37] Function-blocking mAbs against adhesion molecules and chemokines have often been the drug candidates used for proof-of-principle experiments. In parallel, recombinant ligand or receptor analogs have been developed. Ultimately, knowledge of the structure of the adhesion molecules has allowed the design of rational, small-molecular drugs, such as those affecting the conformational state of leukocyte integrins. Possibilities to modulate mRNA expression of adhesion molecules through antisense oligonucleotides and RNA interference have added another potential tool to the pharmaceutical armamentarium.

Adhesion-Modulating Drugs in Clinical Use

Although all these forms of therapy have been enormously successful in a panoply of animal models, transfer to the clinic has been difficult. However, a small number of very potent drugs targeting adhesion molecules have been introduced.

The first two selective adhesion molecule (SAM) inhibitors were natalizumab and efalizumab. The way for natalizumab, a humanized anti-α_4 integrin antibody, was paved by the excellent results of α_4–blocking in EAE (see above). In patients with relapsing MS, a monthly intravenous injection of natalizumab leads to a significant reduction in the numbers of new lesions, numbers of relapses, and the risk of sustained disability.[34] The beneficial effects of natalizumab treatment are evident also in patients who do not respond to interferon- (IFN-β) therapy. Because natalizumab also inhibits $\alpha_4\beta_7$, the gut homing receptor, natalizumab also alleviates inflammation in the gut, and it has been shown to be effective in patients with Crohn disease.[35]

Efalizumab, a humanized, function-blocking antibody directed against CD11a, was the second antiintegrin antibody to enter clinical application. It showed good efficacy in patients with plaque psoriasis, but it was withdrawn from the market because of an increased risk of an opportunistic infection (see below).

The most recent clinically approved selective modulator of an adhesion molecule is vedolizumab.[35] It is a humanized mAb that binds to a combinatorial epitope in α_4/β_7 integrin but does not interact with α_4/β_1 and α_E/β_7 integrins. This makes vedolizumab a remarkably selective inhibitor of lymphocyte homing to the intestine. It has been approved for the treatment of both ulcerative colitis and Crohn disease. An mAb, etrolizumab, which targets the β_7 integrin (thus blocking both α_4/β_7 and α_E/β_7, but not α_4/β_1), which may induce remission in ulcerative colitis, and an anti-MAdCAM-1 antibody that appears to be effective in ulcerative colitis, but not in Crohn disease, are in early-phase clinical trials.[38]

The most feared adverse effect of natalizumab and efalizumab treatment is the possible reactivation of polyoma John Cunningham (JC) virus, which leads to the development of potentially fatal progressive multifocal leukoencephalopathy (PML).[39] Up to 1% of natalizumab-treated patients develop PML in the high-risk groups, and more than 700 cases of PML have been recorded among natalizumab-treated patients. However, the benefits of this therapy in MS still outweigh the risks. With good risk stratification and vigilance, the drug has maintained its status as a very effective biological treatment option for MS. Efalizumab, in contrast, was voluntarily withdrawn from the market in 2009 upon emergence of PML in the treated patients. Theoretically, α_4/β_7 blocking should not interfere with lymphocyte surveillance functions in the CNS, and in line with this, no PML cases have been reported among vedolizumab-treated patients. Mild adverse effects with SAM therapy include an expected increase in susceptibility to infections and injection site reactions. However, apart from the risk of PML with natalizumab, both natalizumab and vedolizumab seem to have a good overall safety profile.

Lymphocyte migration can also be targeted therapeutically by modulating lymphocyte exit from the lymph nodes.[40] Fingolimod/FTY720 causes S1P receptor internalization and degradation. Consequently, it is a functional antagonist of SIP receptors 1, 3, 4, and 5. It inhibits lymphocyte egress from the lymphoid organs. It may selectively retain CCR7-positive T cells, including central memory T cells, which may be particularly important for the pathogenesis of brain inflammation, while sparing CCR7-negative effector memory cells. It has been approved by the US Food and Drug Administration (FDA) as the first-line treatment for relapsing MS, and it is the first orally active disease-modifying drug for this immunological disease. More selective S1P receptor 1 antagonists, such as ozanimod, have shown preliminary efficacy in IBD. Although used in certain clinical indications, targeting of chemokine/chemokine receptors for manipulating lymphocyte migration has not yet proceeded to clinical applications.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Gowans JL, Knight EJ. The route of recirculation of lymphocytes in the rat. Proc R Soc Lond B Biol Sci 1964;159:257–82.
2. Girard JP, Moussion C, Förster R. HEVs, lymphatics and homeostatic immune cell trafficking in lymph nodes. Nat Rev Immunol 2012;12:762–73.
3. Mazo IB, Massberg S, von Andrian UH. Hematopoietic stem and progenitor cell trafficking. Trends Immunol 2011;32:493–503.
4. Zhang SL, Bhandoola A. Trafficking to the thymus. Curr Top Microbiol Immunol 2014;373:87–111.
5. Chang JE, Turley SJ. Stromal infrastructure of the lymph node and coordination of immunity. Trends Immunol 2015;36:30–9.
6. Fu H, Ward EJ, Marelli-Berg FM. Mechanisms of T cell organotropism. Cell Mol Life Sci 2016;73(16):3009–33.
7. Germain RN, Robey EA, Cahalan MD. A decade of imaging cellular motility and interaction dynamics in the immune system. Science 2012;336:1676–81.
8. Sigmundsdottir H, Butcher EC. Environmental cues, dendritic cells and the programming of tissue-selective lymphocyte trafficking. Nat Immunol 2008;9:981–7.
9. Fan X, Rudensky AY. Hallmarks of tissue-resident lymphocytes. Cell 2016;164:1198–211.
10. Arnon TI, Cyster JG. Blood, sphingosine-1-phosphate and lymphocyte migration dynamics in the spleen. Curr Top Microbiol Immunol 2014;378:107–28.
11. Nourshargh S, Alon R. Leukocyte migration into inflamed tissues. Immunity 2014;41:694–707.
12. Qi H, Kastenmüller W, Germain RN. Spatiotemporal basis of innate and adaptive immunity in secondary lymphoid tissue. Annu Rev Cell Dev Biol 2014;30:141–67.
13. Headland SE, Norling LV. The resolution of inflammation: principles and challenges. Semin Immunol 2015;27:149–60.
14. Buckley CD, Barone F, Nayar S, et al. Stromal cells in chronic inflammation and tertiary lymphoid organ formation. Annu Rev Immunol 2015;33:715–45.
15. Kupper TS, Fuhlbrigge RC. Immune surveillance in the skin: mechanisms and clinical consequences. Nat Rev Immunol 2004;4:211–22.
16. Vestweber D. How leukocytes cross the vascular endothelium. Nat Rev Immunol 2015;15:692–704.
17. McEver RP. Selectins: initiators of leucocyte adhesion and signalling at the vascular wall. Cardiovasc Res 2015;107:331–9.
18. Proudfoot AE, Uguccioni M. Modulation of chemokine responses: synergy and cooperativity. Front Immunol 2016;7:183.
19. Sallusto F, Baggiolini M. Chemokines and leukocyte traffic. Nat Immunol 2008;9:949–52.
20. Kinashi T. Overview of integrin signaling in the immune system. Methods Mol Biol 2012;757:261–78.
21. Jordan AR, Racine RR, Hennig MJ, et al. The role of CD44 in disease pathophysiology and targeted treatment. Front Immunol 2015;6:182.
22. Salmi M, Jalkanen S. Cell-surface enzymes in control of leukocyte trafficking. Nat Rev Immunol 2005;5:760–71.
23. Salmi M, Jalkanen S. Ectoenzymes in leukocyte migration and their therapeutic potential. Semin Immunopathol 2014;36:163–76.
24. Braun A, Worbs T, Moschovakis GL, et al. Afferent lymph-derived T cells and DCs use different chemokine receptor CCR7-dependent routes for entry into the lymph node and intranodal migration. Nat Immunol 2011;12:879–87.
25. Cyster JG, Dang EV, Reboldi A, et al. 25-Hydroxycholesterols in innate and adaptive immunity. Nat Rev Immunol 2014;14:731–43.
26. Debes GF, Arnold CN, Young AJ, et al. Chemokine receptor CCR7 required for T lymphocyte exit from peripheral tissues. Nat Immunol 2005;6:889–94.
27. Salmi M, Karikoski M, Elima K, et al. CD44 binds to macrophage mannose receptor on lymphatic endothelium and supports lymphocyte migration via afferent lymphatics. Circ Res 2013;112:1577–82.
28. Karikoski M, Marttila-Ichihara F, Elima K, et al. Clever-1/stabilin-1 controls cancer growth and metastasis. Clin Cancer Res 2014;20:6452–64.
29. Gonzalez-Cabrera PJ, Brown S, Studer SM, et al. S1P signaling: new therapies and opportunities. F1000Prime Rep 2014;6:109.
30. Notarangelo LD, Badolato R. Leukocyte trafficking in primary immunodeficiencies. J Leukoc Biol 2009;85:335–43.
31. DeLisser HM, Christofidou-Solomidou M, Sun J, et al. Loss of endothelial surface expression of E-selectin in a patient with recurrent infections. Blood 1999;94:884–94.
32. Etzioni A. Defects in the leukocyte adhesion cascade. Clin Rev Allergy Immunol 2010;38:54–60.

33. Hanna S, Etzioni A. Leukocyte adhesion deficiencies. Ann N Y Acad Sci 2012;1250:50–5.
34. Steinman L. Immunology of relapse and remission in multiple sclerosis. Annu Rev Immunol 2014;32:257–81.
35. Habtezion A, Nguyen LP, Hadeiba H, et al. Leukocyte trafficking to the small intestine and colon. Gastroenterology 2016;150: 340–54.
36. Peske JD, Woods AB, Engelhard VH. Control of CD8 T-cell infiltration into tumors by vasculature and microenvironment. Adv Cancer Res 2015;128:263–307.
37. Mackay CR. Moving targets: cell migration inhibitors as new anti-inflammatory therapies. Nat Immunol 2008;9:988–98.
38. Rivera-Nieves J. Strategies that target leukocyte traffic in inflammatory bowel diseases: recent developments. Curr Opin Gastroenterol 2015;31:441–8.
39. Calabrese LH, Molloy E, Berger J. Sorting out the risks in progressive multifocal leukoencephalopathy. Nat Rev Rheumatol 2015;11:119–23.
40. Brinkmann V, Billich A, Baumruker T, et al. Fingolimod (FTY720): discovery and development of an oral drug to treat multiple sclerosis. Nat Rev Drug Discov 2010;9:883–97.

MULTIPLE-CHOICE QUESTIONS

1. In the peripheral lymph nodes:
   A. All lymphocytes enter the tissue through high endothelial venules
   B. Activated B lymphocytes exit the tissue by migrating to the subcapsular sinus and afferent lymphatics
   C. Chemokines immobilized on the high endothelial venules are crucial for triggering firm adhesion of bloodborne lymphocytes
   D. Effector T lymphocytes are imprinted to home selectively to gut

2. During the multistep adhesion, cascade lymphocytes:
   A. Use integrins mainly for tethering
   B. Enter the spleen from the red pulp via capillary endothelium
   C. Solely transmigrate through the endothelial cell cytoplasm by a transcytotic process
   D. Mainly utilize selectins for the rolling step

3. Which one of the following therapies modulating lymphocyte traffic is not currently used in clinics?
   A. Functional antagonists of sphingosine 1 phosphate receptors in reducing lymphocyte traffic to brain in multiple sclerosis
   B. Agonists of L-selectin in enhancing lymphocyte homing to spleen in sepsis
   C. Function-blocking antibody against α in alleviating multiple sclerosis
   D. Function-blocking antibody against $\alpha_4\beta_7$ integrin in reducing intestinal homing in inflammatory bowel disease

12

T-Cell Activation and Tolerance

Erik J. Peterson, Jonathan S. Maltzman

Activation of T lymphocytes during an immune response triggers a series of programmed gene regulation, proliferation, differentiation, and effector functions. These T-cell functions coordinate with other leukocytes to permit the immune system to react against foreign antigens without initiating self-reactivity or autoimmunity. Each of these functions is fully dependent on environmental cues that are recognized by cell surface receptors and are then translated through biochemical alterations within the cell. This chapter discusses signal transduction through one of the most studied of these receptors, the antigen-specific T-cell receptor (TCR) complex. It addresses the mechanisms whereby signals propagated through the TCR combine with those from costimulatory receptors to produce either productive activation or immune tolerance. It also discusses how abnormal TCR signaling, and imbalanced signaling through costimulatory or coinhibitor molecules, can contribute to T-cell dysfunction and disease (Table 12.1). Targeting these molecular pathways has resulted in several clinically relevant drugs currently used to treat autoimmunity, transplant rejection, and cancer.

CLINICAL RELEVANCE

- Dysfunction or deficiency of T-cell signaling proteins (induced or spontaneous) has been causally linked to several disease states in animal or human models. Molecules wherein mutations may lead to immune deficiency or dysregulation include:
  - CD45
  - LCK
  - ZAP-70
  - SLP-76
  - LAT
  - IL-2R γ chain
  - Mst1
- Molecules wherein mutations may lead to lymphocyte hyperproliferation include:
  - CTLA-4
  - SHP-1
  - CD95/CD95 ligand
  - SAP
  - CBL/CBL-b
  - ZAP-70
  - LYP
  - DGK

THE T-CELL ANTIGEN RECEPTOR COMPLEX

The TCR complex consists of a ligand-binding TCR α/β or γ/δ heterodimer (Chapter 4) in association with the CD3/ζ chain complex, which provides transmembrane signal transduction capability.[1] Specificity of the TCR for antigen resides exclusively within the highly polymorphic, clonotypic, ligand-binding α/β or γ/δ heterodimers. Although many of the biochemical events leading to α/β and γ/δ T-cell activation are similar, α/β T cells exhibit a broader spectrum of antigen reactivity and are thought to participate in a wider range of specific immune responses. This chapter focuses on α/β T cells.

The α/β TCR specifically recognizes short peptide ligands bound to major histocompatibility complex (MHC) antigens (Chapter 5) on the surface of antigen-presenting cells (APCs) (Chapter 6). Coreceptor molecules expressed on subsets of α/β T cells determine whether the TCR recognizes class I or class II MHC. CD4 T cells are stimulated by processed exogenous antigen presented by class II MHC molecules on the surface of professional APCs. CD8 T cells respond to peptides synthesized by APCs and presented by class I molecules. CD4 and CD8 associate with MHC class II and class I molecules, respectively, to stabilize the tripartite interaction between the TCR, antigen, and MHC, which increases the effectiveness of TCR engagement.

Although the α/β chains of the TCR contain all of the information necessary for antigen/MHC binding, these proteins are not sufficient to initiate the intracellular biochemical events that signal antigen recognition. Instead, signal transduction is accomplished by noncovalently associated CD3 and TCRζ polypeptides, which include several pairs of transmembrane hetero- or homodimers (Fig. 12.1). Each CD3 and ζ chain derives signaling capacity from the presence of one or more cytoplasmic regions known as *immunoreceptor tyrosine-based activation motifs* (ITAMs).[2]

KEY CONCEPTS

Protein Tyrosine Kinase Activation

T-cell receptor engagement activates several families of protein tyrosine kinases, which are required for propagation of second messenger-instigated intracellular signaling:

- Src family: LCK, FYN
- Syk family: ZAP-70
- Tec family: ITK, RLK

Activation of Protein Tyrosine Kinases by the TCR and the Role of the ITAMs

Among the earliest of the biochemical events following engagement of the TCR is the activation of LCK and FYN, two members of the SRC family of protein tyrosine kinases (PTKs).[2] Along with all SRC family PTKs, LCK and FYN share features critical

TABLE 12.1 Phenotypes Associated With Deficient Function of Selected T-Cell Signaling Molecules

		PHENOTYPE	
Molecule	**Affected Signaling Event**	**Mouse**	**Human**
TCR Signaling			
CD3 γ	TCR expression	$B^{+}T^{+}NK^{+}$ SCID	$B^{+}T^{+}NK^{+}$ SCID
CD3 ε	TCR expression	$B^{+}T^{-}NK^{+}$ SCID	$B^{+}T^{+/-}NK^{+}$ SCID
CD3 δ	TCR expression	$B^{+}T^{-}NK^{+}$ SCID	$B^{+}T^{-}NK^{+}$ SCID
CD3 ζ	TCR expression, TCR-mediated PTK activation	$B^{+}T^{+}NK^{+}$ SCID	$B^{+}T^{+}NK^{+}$ SCID
ZAP-70	TCR-mediated PTK activation	$B^{+}T^{+/-}NK^{+}$ SCID. TCRαβ T cells are absent, but TCRγδ T cells survive. Arthritis occurs in some inbred strains	$B^{+}T^{+/-}NK^{+}$ SCID. CD8 T cell lymphopenia. Overexpressed in some hematological malignancies
LCK	TCR-mediated PTK activation	$B^{+}T^{+/-}NK^{+}$ SCID. Impaired thymopoiesis and proliferation	$B^{+}T^{+}NK^{+}$ SCID. CD4 lymphopenia, absent CD28 expression on CD8 T cells, and hypogammaglobulinemia
CD45	Maintenance of SRC family PTK in "open" conformation	$B^{+}T^{+/-}NK^{+}$ SCID. Impaired thymopoiesis	$B^{+}T^{+/-}NK^{+}$ SCID. Impaired thymopoiesis, decreased cytotoxic T-cell responses, progressive hypogammaglobulinemia, genetic polymorphisms may correlate with increased prevalence of autoimmune disease
SAP	SHP-2 binding to SLAM	Increased susceptibility to lymphocytic choriomeningitis virus, reduced IgE production, NKT cell deficiency	X-linked lymphoproliferative disease (XLP) with B-cell hyperresponsiveness, NKT cell deficiency
WASP	Actin polymerization	Decreased T-cell proliferation and interleukin-2 (IL-2) production	Wiscott-Aldrich syndrome (immunodeficiency, atopic dermatitis, thrombocytopenia, bloody diarrhea)
CBL/CBL-b*	E3 ubiquitin ligase. Recruitment of CrKL/C3G inhibitory complex	Hyperproliferative T cells*	Proto-oncogene for leukemia
LAT	Coupling PTK activation to downstream signals	$B^{+}T^{-}NK^{+}$ SCID. Absolute block in thymopoiesis	
SLP-76	Coupling PTK activation to downstream signals	$B^{+}T^{-}NK^{+}$ SCID. Absolute block in thymopoiesis. Defect in vascular/lymphatic development	
ITK/RLK	Amplification of proximal PTK signals. Activation of PLC-γ1	Defective Th2 immune responses	
CTLA-4	Inhibition of CD28-mediated costimulation	Fatal lymphoproliferative disease with myocarditis, pancreatitis	Allelic variants associated with autoimmunity, including Hashimoto thyroiditis, Graves disease, and systemic lupus erythematosus
SHP-1	Downregulation of PTK activity	Autoimmunity, inflammatory lung disease. "Moth-eaten" mice	
LYP (Lymphoid phosphatase; *PTPn22* gene product)	Attenuation of LCK activity	Augmented TCR-stimulated IL-2 production and proliferation	Allelic variants are associated with increased risk of rheumatoid arthritis, systemic lupus erythematosus, type 1 diabetes mellitus
DGKζ	Downregulation of DAG-dependent Ras activation	Impaired T-cell anergy induction	
Mst1 (STK4)	Suppression of Akt, phosphorylation of FoxO transcription factor	Inefficient thymic egress; impaired positive selection; impaired Treg development and function	Recurrent infections; progressive loss of peripheral CD4 T cells; autoimmune hemolytic anemia
IL-2R Signaling			
γc	Coupling IL-2 binding to JAK activation	$B^{+}T^{-}NK^{-}$ SCID	$B^{+}T^{-}NK^{-}$ SCID, X-linked SCID
JAK3	Phosphorylation of STAT proteins	$B^{+}T^{-}NK^{-}$ SCID	$B^{+}T^{-}NK^{-}$ SCID

γc, common γ-chain (IL2Rγ); IgE, immunoglobulin E; NK, natural killer; SCID, severe combined immunodeficiency; TCR, T-cell receptor.

*CBL and CBL-b are closely related; CBL-b-deficient mice develop autoimmune features and more severe lymphoproliferative disease compared with mice lacking CBL.

for their function (Fig. 12.2). These include an amino-terminal myristoylation sequence that directs membrane localization, a SRC homology 3 (SH3) domain that permits associations with other proteins containing regions rich in proline residues, a SRC homology 2 (SH2) domain that dictates interactions with proteins phosphorylated on tyrosine residues, a catalytic region, and a carboxyl-terminal tyrosine residue. The precise mechanism whereby LCK and FYN are stimulated by the TCR is not clear, but both have been shown to associate physically with TCR CD3 components and/or coreceptors CD4 and CD8.

SRC family PTK enzymatic function is regulated, in part, by the state of tyrosine phosphorylation of the kinase. When the

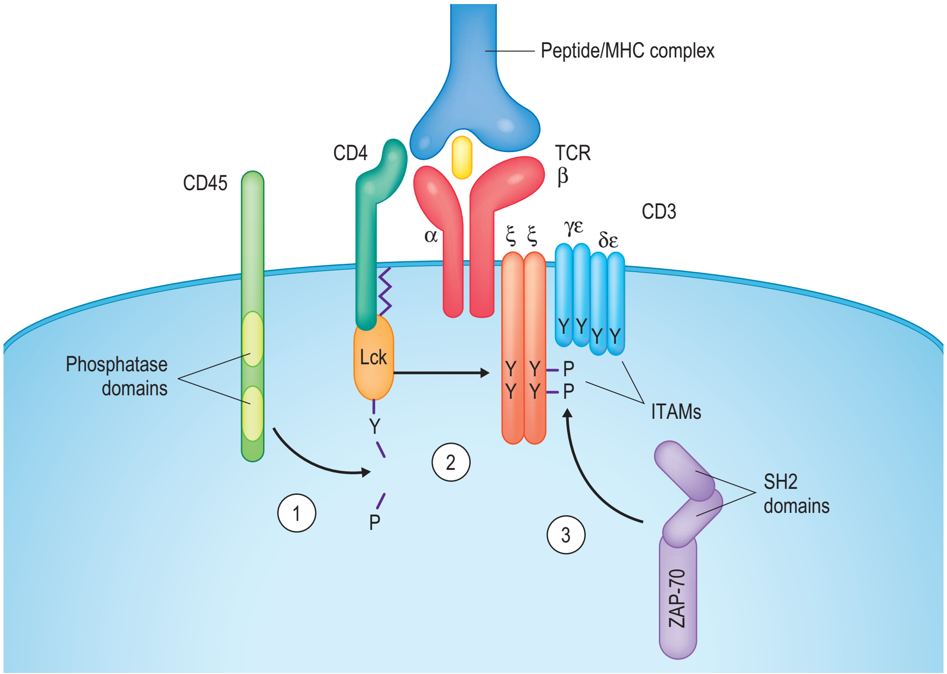


FIG 12.1 Biochemical Events in Early T-Cell Receptor (TCR) Signaling. The tyrosine phosphatase CD45 dephosphorylates the negative regulatory tyrosine residue on the CD4-associated protein tyrosine kinase (PTK) LCK, maintaining LCK in an activatable conformation (1). Engagement of the TCR α/β heterodimer and the CD4 (or CD8) coreceptors by major histocompatibility complex (MHC)–bound peptide antigen brings activated LCK into proximity with immunoreceptor tyrosine-based activation motif (ITAM)-bearing CD3 chains. LCK phosphorylates the CD3ζ chain within ITAMs (2). The phosphorylated CD3ζ-chain ITAMs interact with the tandem SH2 domains of the cytoplasmic PTK ZAP-70 (3), permitting activation of ZAP-70 and phosphorylation of downstream substrates.

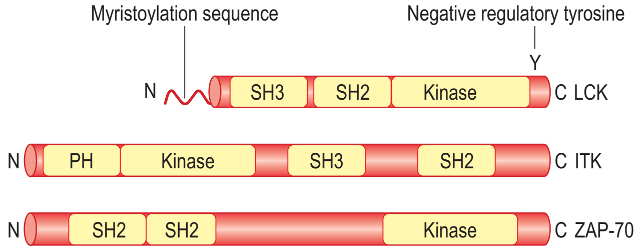


FIG 12.2 Domain Organization of T-Cell Receptor (TCR)-Stimulated Protein Tyrosine Kinases (PTKs). Comparative schematic representation of members of three families of PTKs required for T cell-activating signals. In addition to catalytic domains, LCK (SRC family), ITK (TEC family), and ZAP-70 (SYK family) each contain regions responsible for mediating protein–protein interactions, including SH3 and SH2 domains. SH3, Src homology 3; SH2, Src homology 2; PH, pleckstrin homology.

conserved carboxyl-terminal tyrosine residue is phosphorylated, SRC family PTKs adopt a "closed" conformation that is the product of an intramolecular interaction between that phosphotyrosine and the SH2 domain (Fig. 12.3). This intramolecular interaction inhibits the enzymatic activity of the PTK, limiting subsequent tyrosine phosphorylation–dependent signaling events. Phosphorylation of the carboxyl-terminal tyrosine (Y505 in LCK and Y527 in FYN) is dynamically regulated.[3] Phosphate is transferred to this residue by the cytoplasmic PTK CSK and is removed by the transmembrane protein tyrosine phosphatase CD45.

Phenotypes of CD45-deficient cells in mice and humans highlight the critical regulatory importance of the carboxyl-terminal tyrosine in SRC family PTKs.[4] TCR signal transduction in cell lines lacking CD45 is blocked at the most proximal step, and mice that have been genetically modified to lack CD45 expression exhibit profound defects in thymocyte development and subsequent T-cell activation. CD45 deficiency in humans results in a T^-, B^+, NK^+ severe combined immunodeficiency (SCID) (Chapter 35). These outcomes correlate with markedly impaired LCK enzymatic activity and hyperphosphorylation of Y505, the regulatory tyrosine.

Following TCR engagement and PTK activation, numerous cellular substrates become tyrosine-phosphorylated, including the CD3 and ζ-chain ITAMs (see Fig. 12.1). In resting T cells, key tyrosine residues within the ITAMs are embedded within the hydrophobic core of the plasma membrane lipid bilayer. Upon TCR triggering, conformational changes induced within the CD3 cytoplasmic tails result in enhanced tyrosine accessibility to the action of SRC family kinases.[2] ITAM phosphorylation creates a docking site for another cytosolic PTK, ζ-associated phosphoprotein of 70 kilodaltons (kDa) (ZAP-70). ZAP-70, a member of the SYK family PTKs, contains a catalytic domain that is located carboxyl-terminal to two tandem SH2 domains (see Fig. 12.2). The ZAP-70 SH2 domains have affinity for

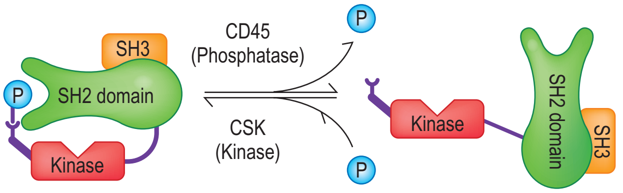


FIG 12.3 Model for Dynamic Regulation of LCK by Intramolecular Interaction Between an SH2 Domain and Phosphotyrosine. The transmembrane phosphatase CD45 dephosphorylates tyrosine 505 in the carboxyl-terminus of the SRC family protein tyrosine kinase (PTK) LCK. CD45 activity maintains LCK in an "open" conformation, permitting LCK kinase domain access to intracellular substrates. CSK activity opposes that of CD45; phosphorylation of tyrosine 505 results in an intramolecular interaction between the SH2 domain and phosphotyrosine. Inhibition of LCK kinase activity correlates with the "closed" conformation *(left).*

phosphotyrosine present within ITAMs. Thus inducible phosphorylation of the CD3 and ζ-chain ITAMs results in the formation of docking sites that mediate recruitment of ZAP-70. Upon recruitment to TCR, ZAP-70 enzymatic activity is increased as a result of phosphorylation by LCK as well as autophosphorylation. The net result of these phosphorylations is conversion of the TCR from an enzymatically inactive ligand-binding complex to a potent PTK.

LCK, FYN, and ZAP-70 are critically important for both thymocyte development and mature T-cell activation. Jurkat T cells (a human leukemic cell line) deficient in expression of either LCK or ZAP-70 cannot be activated via the TCR. Mice deficient in Zap-70 or Lck, but not in Fyn, exhibit a significant yet incomplete block in early T-cell development. The pre-TCR (a complex present on immature thymocytes that includes signaling components thought to be similar to the TCR on mature T cells) appears to require Src and Syk family PTKs to transduce signals. ZAP-70 deficiency and abnormal LCK function in humans create a T^-, B^+, NK^+ SCID.[5]

In addition to SRC and SYK family PTKs, TCR engagement results in the activation of a third family of cytosolic PTKs, the Tec family, which includes TEC, ITK, and RLK.[6] Tec PTKs contain SH2, SH3, and catalytic domains, as well as pleckstrin homology (PH) domains that mediate interactions with membrane-localized phospholipids (see Fig. 12.2). PH domains permit recruitment of Tec family kinases to the plasma membrane, where they can phosphorylate important substrates. Activation of Tec family members is dependent on the prior activation of SRC and SYK family PTKs. ITK positively regulates antigen receptor signaling through recruitment and activation of the lipid modulator PLCγ to a "signalosome" nucleated by the cytoplasmic adaptor protein SLP-76 (see below). The importance of normal ITK signaling was recently shown in humans, where homozygous mutation led to immunodeficiency resulting in death from Epstein-Barr virus (EBV)–associated lymphoproliferation.

Tec family kinases are thought to modulate TCR signal strength. Mice deficient in Tec kinases display variable partial defects in thymocyte development and peripheral T-cell maturation. T cells deficient in Tec PTKs display alterations in actin polymerization and cytoskeletal polarization, which promote recruitment of adhesion-related proteins, including β_2 integrins and signaling mediators, such as Vav, to membrane regions important for propagating activation signals. The importance of normal ITK signaling was recently shown in humans, where homozygous mutation led to immunodeficiency resulting in death from EBV-associated lymphoproliferation.

KEY CONCEPTS

TCR Signaling Pathways

T-cell receptor engagement leads to the initiation of signaling cascades, including:
- PLCγ1 activation
  - Calcium flux
  - DAG formation
- RAS/MAPK
- NF- κB
- PI3K

Second Messenger Cascades Downstream of the TCR-Stimulated PTKs

TCR engagement incurs numerous biochemical changes that are dependent on the activation of ZAP-70 and its association with the CD3 and ζ-chain ITAMs. One intermediate event in TCR signaling is activation of the membrane-associated enzyme phospholipase Cγ1 (PLCγ1) (Fig. 12.4).[7] PLCγ1 is phosphorylated by multiple TCR-dependent PTKs, including both ZAP-70 and members of the Tec family. TCR-stimulated tyrosine phosphorylation alone is not sufficient to activate PLCγ1; relocalization of the enzyme into adaptor-protein nucleated complexes probably plays a critical role.

Activated PLCγ1 catalyzes the hydrolysis of phosphatidylinositol-4,5-bisphosphate (PIP2), a minor plasma membrane phospholipid. PIP_2 hydrolysis gives rise to two second messengers, the soluble sugar inositol 1,4,5-trisphosphate (IP_3) and the lipid diacylglycerol (DAG). IP_3 binding to a cognate receptor on endoplasmic reticulum (ER) results in the release of calcium from this organelle. Consequent IP_3-mediated reductions in calcium ER concentrations are sensed by STIM1, an EF-hand domain–containing protein localized in the ER membrane.[8] In turn, STIM1 aggregation activates store-operated calcium entry channels, including the transmembrane protein ORAI. Marked calcium influx from the extracellular milieu attends calcium channel activation. Intracellular calcium serves as a second messenger that activates calcineurin, a serine/threonine phosphatase. Substrates of calcineurin include members of the nuclear factor of activated T cells (NFATs) family. In resting T cells, NFAT proteins are phosphorylated and reside in the cytosol. Activated calcineurin dephosphorylates NFAT, allowing it to translocate to the nucleus and bind the response elements in the promoters of genes important for T-cell activation.

Knowledge of the calcineurin pathway has been exploited clinically. The compounds cyclosporine and tacrolimus are mainstays in the prevention of human solid-organ transplant rejection (Chapter 81) and are also used, albeit not commonly, to treat T cell–driven autoimmune diseases (Chapter 87). Both cyclosporine and tacrolimus interfere with calcineurin-dependent T-cell activation, but via distinct mechanisms.

The other major product of PIP_2 hydrolysis, DAG, functions as second messenger to a parallel cascade of TCR signaling intermediates, including protein kinase D (PKD), Ras guanyl nucleotide releasing proteins (RasGRPs), and members of the protein kinase C (PKC) family of serine/threonine kinases.[9] PKD cooperates with PKC-dependent signals to activate high-affinity

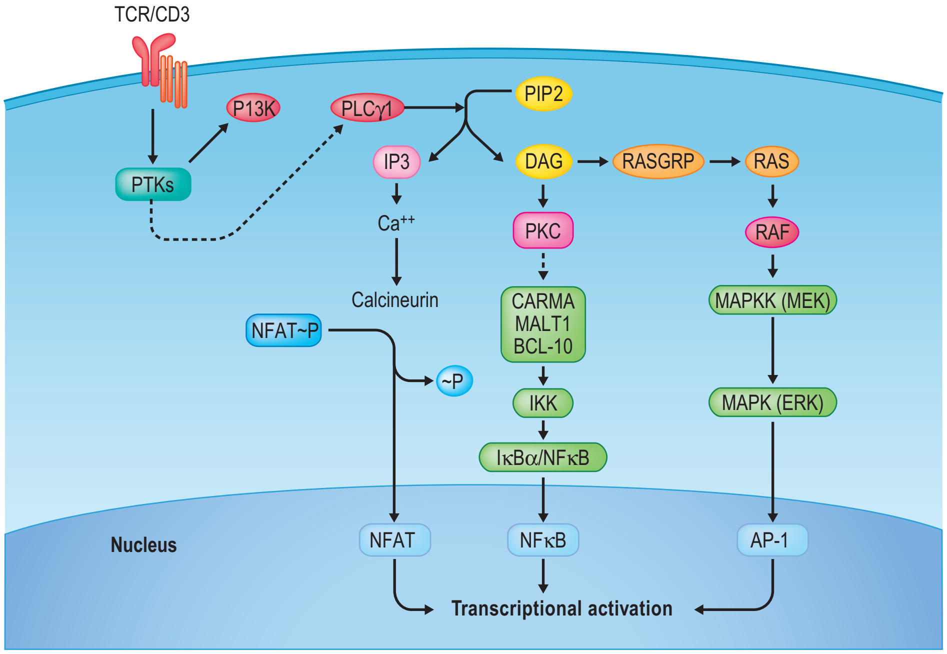


FIG 12.4 Signaling Pathways Activated by T-Cell Receptor (TCR) Engagement. TCR ligation results in activation of protein tyrosine kinases (PTKs), such as LCK and ZAP-70. Phospholipase Cγ1 (PLCγ1) becomes phosphorylated and activated by PTKs, including ITK. Hydrolysis of phosphatidyl inositol bisphosphate (PIP_2) by PLCγ1 releases diacylglycerol (DAG) and inositol trisphosphate (IP_3). IP_3 stimulates an increase in intracellular calcium concentration, which activates the phosphatase calcineurin. Calcineurin dephosphorylates NFAT, thereby signaling NFAT translocation to the nucleus. The formation of DAG leads to activation of RAS-GRP1 GEF activity and RAS activation. Active RAS binds and stimulates the kinase RAF1, initiating a cascade of serine/threonine kinases (MAPK cascade) leading to phosphorylation and nuclear translocation of the ERK kinases. DAG formation also results in activation of the CARMA/BCL-1/MALT1 complex leading to phosphorylation of IκB kinase (IKK). Active IKK phosphorylates IκB-α, leading to IκB-α degradation and release of NF-κB to the nucleus. TCR-activated PI3K catalyzes formation of PIP3 from membrane-associated PIP_2; the phosphatase PTEN antagonizes PIP_3 formation. PIP_3 binds to the Akt lipid-binding pleckstrin homology domain, a required element in Akt activation. Active Akt both promotes the PKC-CARMA-NFκB pathway and blocks FoxO-dependent transcriptional regulation. Transcription factors NFAT, NFκB, and those activated by the MAPK pathway cooperate to upregulate transcription of genes, such as *IL-2,* critical for T-cell activation.

binding capacity and clustering of integrins, a family of molecules that mediate TCR signal-augmented cell binding to adhesion molecules on APCs.[10] RasGRP mediates activation of the Ras cascade. Deletion of RasGRP1 leads to severe defects in thymocyte development in mice, highlighting the importance of the RasGRP pathway in TCR signaling. PKC is essential for full activation of the Ras/ERK and NF-κB cascades (see below) that are required to mount a TCR-stimulated gene transcriptional program and drive cellular activation.

In T lymphocytes, RasGRP1 is one of two distinct guanine exchange factors (GEFs) that mediate exchange of guanosine triphosphate (GTP) for guanosine diphosphate (GDP) bound to RAS.[11] The other, parallel-acting factor, Son of Sevenless (SOS), is recruited to the plasma membrane through its interaction with the GRB2 adaptor protein. Activation via the GRB2-SOS pathway acts as a digital “on–off” switch, whereas RasGRP1-mediated activation can be envisioned as an analog rheostat, leading to varying degrees of downstream signal, depending on the strength of upstream stimulus.[11] RAS activity is tightly regulated. RAS activity may be curbed through the hydrolysis of RAS-bound GTP to GDP, either through the intrinsic guanosine triphosphatase (GTPase) activity of RAS or through the recruitment of GTPase-activating proteins (GAPs) to the active RAS molecule.

Active RAS recruits the serine-threonine kinase RAF to the plasma membrane, initiating a phosphorylation cascade. Active RAF, in turn, phosphorylates MEK, which phosphorylates extracellular signal-regulated kinase (ERK), which undergoes nuclear translocation. Upon entry into the nucleus, ERK can phosphorylate and activate several transcription factors that are critical for TCR-induced transactivation of cytokines and other activation genes. Studies in cell lines and in genetically altered mice attest to the central importance of RAS activation for T-cell function. Jurkat T cells expressing inhibitory mutant forms of RAS fail to produce interleukin-2 (IL-2) following TCR engagement. In contrast, Jurkat cells expressing an activated form of RAS produce IL-2 much more readily than do wild-type cells. Similarly, mice transgenic for activating Ras mutants show alterations in thymocyte development and demonstrate a partially stimulated state in the absence of antigen binding.

TCR stimulation also triggers a cascade headed by phosphatidylinositol 3′-hydroxyl kinase (PI3K) (see Fig. 12.4).[12] PI3Ks are composed of two noncovalently bound subunits; the p85 regulatory subunit activates the kinase activity of the p110 catalytic subunit. PI3K phosphorylates phosphoinositides, which play an important role in the regulation of several downstream serine/threonine kinases, including protein kinase B (PKB) (Akt). One central function of Akt is promotion of the NFκB pathway downstream of PLCγ1- and DAG-driven PKC. Although TCR engagement alone can stimulate some degree of PI3K function, full activity of the lipid kinase requires costimulation of the T cell through receptors, such as CD28 (see below). The outcomes of PI3K and Akt signaling play essential diverse roles in T-cell development, differentiation, and effector function, most of which maximize effector T-cell activation responses. These include suppressing the induction of Foxp3-expressing regulatory T cells (Tregs); suppressing expression of proapoptotic molecules, including Bim or Bad, and of cell cycle inhibitors; and promoting T-cell survival via metabolic processes, including glucose uptake and glycolysis. The importance of this pathway is suggested by the efficacy of PI3K inhibitors in animal models of T cell–driven autoimmunity and in current clinical trials in human hematologic malignancies.[13]

TCR cross-linking also leads to activation of the serine-threonine kinase PKC-θ. PKC activation ultimately engenders nuclear translocation of members of the nuclear factor-κB (NFκB) family of transcription factors (see Fig. 12.4).[14] In the basal state, NFκB family members are sequestered in the T-cell cytoplasm through interaction with inhibitors of NFκB (IκB). TCR stimulation results in the formation of a multimolecular activating complex composed of CARMA1, BCL-10, and MALT1 (CBM). The assembly of the complex lies upstream of signals leading to activation of IκB kinases (IKKs), which phosphorylate and degrade IκB. Transient IκB degradation frees NFκB for translocation to the nucleus. Defects in T-cell activation and survival result from deficiencies in CBM proteins.

INTEGRATION OF SECOND-MESSENGER PATHWAYS BY ADAPTOR PROTEINS

Identification and characterization of adapter proteins has provided insight into how intracellular signaling pathways are initiated by TCR cross-linking.[15,16] Adaptors lack enzymatic or transcriptional regulatory activity. Instead, they possess modular domains responsible for subcellular relocalization and intermolecular interactions. Both constitutive and induced intermolecular interactions mediated by adaptor molecules can promote TCR signal transduction.

Adaptor proteins commonly contain modular domains that exhibit affinity for phosphorylated tyrosine residues (Fig. 12.5A). Such regions include the SH2 and phosphotyrosine-binding (PTB) domains, which recognize phosphorylated tyrosine residues within particular sequence contexts. PTB domains obtain their specificity based on residues amino-terminal to the key phosphotyrosine, whereas SH2 domains recognize sequence motifs carboxyl-terminal to phosphotyrosine. Other adaptor domains include SH3 modules, which bind proline-rich regions, WW regions responsible for interactions with proline/tyrosine or proline/leucine motifs, and PH domains that have specificity for phospholipids.

Several hematopoietic-specific adaptors play essential roles in T-cell development, in coordinating the signals necessary for mature T-cell activation, and in the process of terminating T-cell responses. Two examples are linker of activated T cells (LAT) and SH2 domain-containing leukocyte phosphoprotein of 76 kDa (SLP-76). Both LAT and SLP-76 were identified during efforts to characterize the substrates of the PTKs stimulated by TCR engagement.

LAT is a 36-kDa integral membrane protein that contains numerous tyrosine residues within its cytoplasmic tail (see Fig. 12.5A).[16] Tyrosines exist within specific sequence motifs needed to bind the SH2 domains of other T-cell signaling molecules. In stimulated T cells, LAT inducibly associates with the SH2 domains of GRB2, GADS (Grb2 family member), PLCγ1, and the regulatory (p85) subunit of PI3K. It is likely that these induced intermolecular interactions are critical for communicating TCR engagement to downstream second-messenger cascades. The importance of LAT for T-cell activation was suggested by the signaling phenotype of LAT-deficient variants of the Jurkat T-cell leukemic line. Engagement of the TCR on these cells fails to result in signaling events downstream of ZAP-70 activation. Furthermore, Lat plays an essential role in T-cell development, as Lat-deficient mice have significantly decreased numbers of thymocytes. Thymocytes that remain in $Lat^{-/-}$ mice are arrested at an early stage of development and cannot give rise to any peripheral T cells.

Using similar approaches, it was shown that the adapter SLP-76 is absolutely required for both T-cell development and signaling via the mature TCR.[7] SLP-76 contains an amino-terminal region with tyrosine phosphorylation sites, a central proline-rich domain, and a carboxyl-terminal SH2 domain. Unlike LAT, SLP-76 does not possess a transmembrane domain. By means of its proline-rich region, SLP-76 constitutively associates with the adaptor GADS. Through its SH2 domain, SLP-76 can inducibly interact with other tyrosine-phosphorylated adaptors, such as HPK-1 and ADAP. Following tyrosine phosphorylation, SLP-76 inducibly binds other SH2-domain-containing proteins (*e.g.,* VAV, an exchange factor for the RAC GTP-binding protein; NCK, an adaptor protein; and the Tec family kinase ITK). In mutant T-cell lines lacking SLP-76, the PLCγ1 and RAS/MAPK signaling cascades are not activated by TCR ligation, despite normal TCR-stimulated activation of Zap-70. Mice deficient in SLP-76 show a complete block in early thymocyte development. Each of the SLP-76 domains appears critical for its function, as the overexpression of mutant variants of SLP-76 is unable to restore TCR signaling efficiency fully in either Jurkat T cells or SLP-76-deficient mice.

One model for how SLP-76 and LAT functionally interact holds that each contributes multiple molecular interactions to a larger “signalosome” complex at the plasma membrane (see Fig. 12.5B).[16] Following TCR engagement, the two adaptors associate with each other, bridged by GADS and PLCγ1. Formation of this complex brings Itk into proximity of PLCγ1, resulting in its phosphorylation and activation and leading to the generation of IP3 and DAG, as described above.

Other adapter proteins play important roles in restraining TCR-mediated signal transduction events, frequently by dictating subcellular localization of regulatory enzymes. For example, in the absence of TCR stimulation, the transmembrane adaptor PAG binds to the cytoplasmic PTK CSK, bringing it to the plasma membrane.[17] Membrane-localized CSK phosphorylates the regulatory tyrosine of LCK, thus keeping LCK inactive in resting T cells.

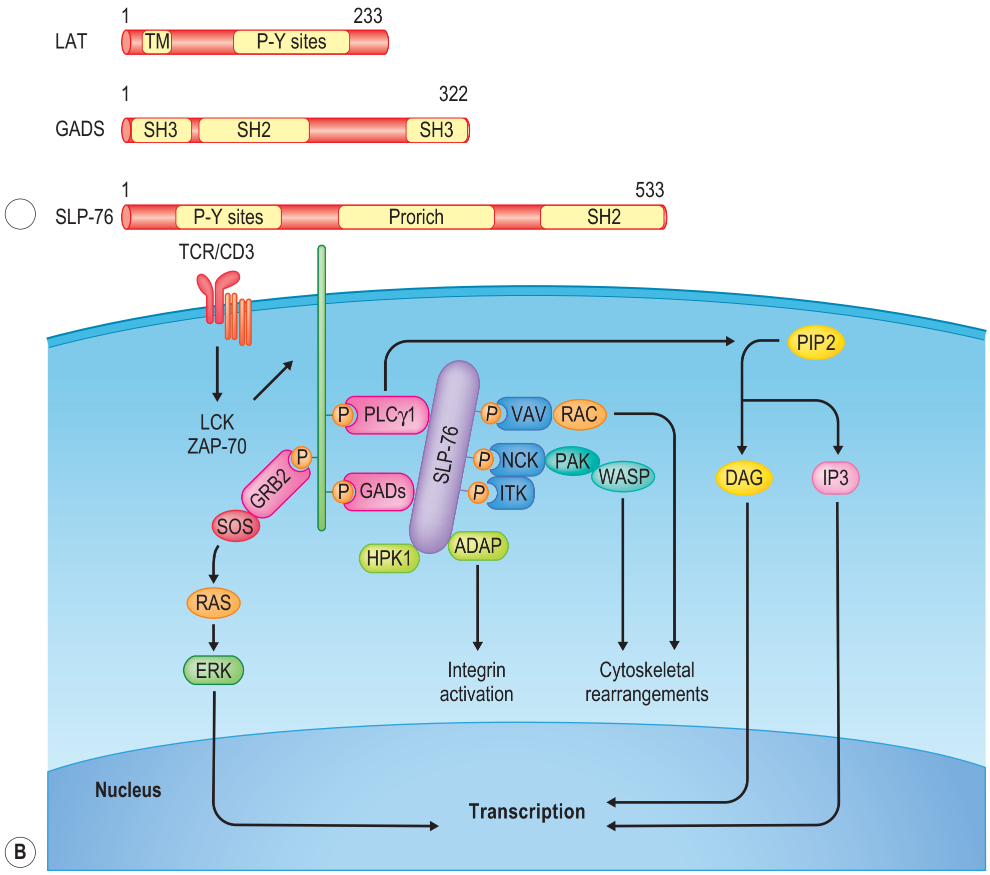


FIG 12.5 Model for Adaptor Protein-Mediated Coupling of the T-Cell Receptor (TCR) to Phospholipase Cγ1 (PLCγ1) Activation. (A) Structural schematics of three adaptors implicated in plasma membrane proximal biochemical events. SH3 domains mediate association with proline-rich regions; SH2 domains associate with phosphorylated tyrosine residues. (B) LAT and SLP-76 are among the substrates of the TCR-activated protein tyrosine kinases (PTKs). When tyrosine residues within the LAT cytoplasmic tail are phosphorylated, GADS binds to LAT through the GADS SH2 domain. Recruitment of GADS results in relocalization of SLP-76, as the proline-rich region of SLP-76 mediates constitutive association with the SH3 domain of GADS. Tyrosine-phosphorylated SLP-76, in turn, becomes associated with ITK via the ITK SH2 domain. ITK is thus brought into proximity with membrane-localized substrates, including PLCγ1. Activation of PLCγ1 leads to hydrolysis of phosphatidylinositol 4,5-bisphosphate (PIP_2) and activation of transcription factors, such as nuclear factor-κB (NFκB), activating protein 1 (AP-1), and nuclear factor of activated T cells (NFAT). SH2 domain containing leukocyte phosphoprotein of 76 kDa (SLP-76) also recruits several other signaling molecules, such as VAV, NCK, HPK1, and ADAP, thereby regulating changes in the actin cytoskeleton and adhesion. Phosphorylation of linker of activated T cells (LAT) also leads to recruitment of Grb2/SOS and an additional pathway for RAS activation. TM, transmembrane domain; P-Y, sites for phosphorylation of tyrosine; pro-rich, proline-rich regions.

Coreceptors Transduce Signals That Are Integrated With TCR Signals

T-cell proliferation and the initiation of effector function require that the T cell must receive signals in addition to the TCR via other cell surface receptors.[18] This requirement for multiple signals allows the T cell to be extremely sensitive to TCR binding while protecting against the inappropriate activation of potentially dangerous effector cells. Because T cells respond to antigens presented on APCs, stimulation under physiological conditions involves the potential engagement of multiple coreceptors on the T cell by cognate ligands on the APCs. Some coreceptors may function to increase the avidity of T cells for interacting APCs. However, many coreceptors exhibit intrinsic signal-transducing capacity. Some signal independently of the TCR; others intersect with TCR-driven signaling machinery. Additionally, coreceptors may function as recruiters of cytoplasmic signaling molecules, including adaptor proteins, as described above.

The most intensively studied coreceptors are CD4 and CD8 (Chapter 4). CD4 or CD8 expression on peripheral T cells define subsets that respond to MHC class II- or class I-bound peptide antigens, respectively (Chapter 6). Either CD4 or CD8 can contribute to enhanced TCR signal strength because they each associate with LCK.[19] This constitutive interaction, which occurs via specific residues within the CD4 and CD8 cytoplasmic domains, localizes a key effector enzyme to the TCR complex.

Counterbalanced Costimulatory and Coinhibitory Signals Determine T-Cell Response Thresholds

As stated earlier, productive T-cell stimulation results from a composite of TCR cross-linking and costimulatory signals. Non-TCR activating signals are transduced via a functional family of T-cell coreceptor proteins termed *costimulatory molecules.* An additional array of coreceptors termed *coinhibitors* serve to dampen or restrict potential for TCR-induced cellular activation and balance the stimulation-promoting effects of costimulatory molecules. Although functionally divergent, costimulatory and coinhibitory molecules share structural features, such as immunoglobulin (Ig)–like extracellular domains. Key costimulatory receptors possessing Ig-like domains include CD28 and ICOS (inducible costimulator); coinhibitory counterparts include CTLA-4 (cytotoxic T lymphocyte antigen-4), PD-1 (programmed death 1), and BTLA (B- and T-lymphocyte attenuator).[20] Increased understanding of both activation-promoting and activation-restraining coreceptors has recently been exploited to substantial effect in management of human autoimmune and malignant diseases.

The best-characterized T-cell costimulatory molecule is CD28, a constitutively expressed homodimeric transmembrane glycoprotein.[18,20] CD28 binds to two ligands expressed on APCs: B7.1 (CD80) and B7.2 (CD86). Ligation of CD28 in isolation has little effect on T-cell activation; however, when CD28 is engaged along with the TCR, many TCR signals are augmented. Indeed, concomitant CD28 and TCR engagement is required for activation of naïve T cells.

CD28 ligation engages several signal transduction pathways implicated in T-cell activation. CD28 contains no intrinsic enzymatic activity, but tyrosine residues within its cytoplasmic tail become inducibly phosphorylated during T-cell activation. These phosphorylated tyrosines recruit several signal-transducing molecules possessing SH2 domains, including GRB2 and the regulatory p85 subunit of PI3K. TCR- and CD28-dependent coactivation of PI3K leads to transactivation of the prosurvival genes *BCL-2* and *BCL-XL.* Via GRB2 association, CD28 engages an NF-κB-dependent survival program dependent on PKCθ and the GEF VAV1. CD28-driven activation of the GTP-binding protein RAS-related C3 botulinum toxin substrate 1 (RAC1) is crucial for the protein kinase activity of c-JUN N-terminal kinase (JNK), which, in turn, plays an important role in CD28-dependent T-cell cytokine production and apoptosis resistance.

Costimulation with CD28 agonists dramatically augments TCR-driven IL-2 production, both by increasing transcription of the *IL-2* gene and by stabilizing its messenger RNA (mRNA).[21,22] Notably, if the TCR is engaged in the absence of CD28 costimulation, activation does not occur. Rather, isolated TCR engagement leads either to T-cell death by apoptosis (Chapter 13) or to a state of unresponsiveness known as *anergy* (see below). CD28 is required not only for T-cell priming during infection but is also important for protective secondary T-cell responses during microbial challenge.[23] Significant nonredundant roles for CD28 have also recently been described in the Treg subset of CD4 T cells. In animals lacking CD28 specifically in Treg, splenomegaly, lung inflammation, and accumulation of activated CD4 T cells occur.[24] These findings suggest that the survival- and activation-promoting function of CD28 costimulation is important for balanced homeostasis of Treg and effector CD4 T-cell subsets. On balance, inhibition of CD28 ligation by CD80/CD86 results in decreased immune responses. Blockade of CD28 costimulation using CTLA4-Ig fusion proteins (abatacept; belatacept) can significantly ameliorate synovial inflammation in rheumatoid arthritis and restrain lymphocyte-driven rejection responses in transplanted kidneys, respectively.

Another Ig domain-containing costimulatory molecule, ICOS is induced on the T-cell surface following combined TCR and CD28 stimulation.[18] ICOS interaction with APC-expressed ICOS-L is required for development of T-follicular helper (Tfh) cells, a subset of CD4 cells required for germinal center formation and B-cell antibody class switching. In murine models of rheumatoid arthritis and multiple sclerosis, ICOS antibody blockade results in reduced Tfh and germinal center formation, associated with suppression of autoimmune responses. ICOS deficiency associates with human common variable immune deficiency (CVID) (Chapter 34), further suggesting that ICOS plays an important role promoting T-cell dependent humoral immunity.

The CTLA4 coinhibitory molecule shares at least two features with CD28: membership in the Ig domain–containing superfamily; and high-affinity binding capacity for APC-expressed ligands CD80 and CD86. Unlike CD28, CTLA4 expression is inducible in conventional CD4 T cells. However, CTLA4 is present constitutively on Treg cells. Experimental data support both T cell–intrinsic and cell-extrinsic mechanisms by which CTLA4 can limit immune responses. Ligation of CTLA4 induces a biochemical cascade, including the activation of PP2A and SHP-2 phosphatases.[24] In conventional T cells, CTLA4 functions to antagonize the TCR “stop signal,” whereby T cells are induced to pause and maintain lengthy physical contact with antigen-bearing dendritic cells (DCs).[18] Moreover, using a CTLA4-dependent process, Treg are thought to limit access of conventional T cells to CD80/CD86 by downregulating these ligands on APCs. Monoclonal, antagonist anti-CTLA4 antibodies have been employed in patients with cancer an attempt to release the CTLA4-driven “negative checkpoint” on T-cell function thought to be impairing natural antitumor T-cell responses. A CTLA4 antagonist was approved by the US Food and Drug Administration (FDA) in 2011; its use has resulted in sometimes dramatic tumor regression and durable remissions in patients with advanced metastatic melanoma and other malignancies.

Like CTLA4, the coinhibitor PD-1 is an Ig domain-containing surface receptor negative regulator of T-cell activation, likely through attenuation of TCR-driven PI3K and Akt. Ligation of PD-1 by cognate ligand PDL-1 (also known as B7-H1) results in recruitment of phosphatases, such as SHP2, to motifs within the PD-1 cytoplasmic tail. Phosphatase-associated PD-1 interacts with TCR microclusters, resulting in dephosphorylation of apical TCR signaling mediators, including ZAP-70 and CD3ζ, thus dampening TCR signal strength. At a cellular level, PD-1 enforces T-cell unresponsiveness by inhibiting TCR-mediated “stop signals” that promote stable, prolonged contact between T cells and APCs. Disruption of PD-1 in mice results in autoantibody formation and glomerulonephritis. In chronic viral infections, CD8 T cells display a hyporesponsive “exhausted” phenotype associated with elevated surface expression of PD-1.[26]

Antibody blockade of PD-1/PDL-1 interactions results in restoration of effector functions in exhausted T cells, suggesting that PD-1 manipulation could benefit patients with viral infections. However, the therapeutic potential of PD-1 has been most dramatically realized in the field of tumor immunotherapy.[27] In many malignancies, both CD4 and CD8 T cells accumulate within the tumor stroma. Tumor-infiltrating lymphocytes (TILs) are characterized by high expression of PD-1 and other coinhibitory molecules; tumor-expressed ligands for these molecules are thought to enforce a state of reduced cytolytic or cytokine-secreting functions. Monoclonal anti-PD1 and anti-PDL1 antagonist antibodies have recently been employed in an attempt to reinvigorate TIL antitumor capacity. PD1/PDL1 blockade has sometimes resulted in dramatic tumor regressions and responses, in responses of diverse tumors ranging from metastatic melanoma to non–small cell lung carcinoma, to lymphoma. Other T-cell coinhibitors, including Lag3 and BTLA, remain under active investigation as potential therapeutic targets.[27]

A second large class of costimulatory molecules includes members of the tumor necrosis factor receptor (TNFR) family.[18] Costimulatory TNFR family members include OX40 (CD134), 4-1BB (CD137), herpesvirus entry mediator (HVEM), CD30, and glucocorticoid-induced TNF receptor (GITR). Cytoplasmic domains within these type I transmembrane proteins contain sequences that recruit a family of adaptor molecules known as TNF receptor–associated factors (TRAFs). Antigen receptor costimulation functions for both CD4 and CD8 T cells have been described for each member of this family. With regard to positive regulation of T-cell signaling, a well-characterized member of this group is OX40.[28] OX40 is upregulated on activated CD4 T cells after CD3/CD28 stimulation. Trimerization of OX40 induced by engagement of APC-expressed OX40L leads to recruitment of TRAF2, TRAF3, and TRAF5. TRAF complexes control activation of a survival-enhancing NF-κB pathway. OX40 deficiency leads to defects in CD4 T cell proliferation, reduced survival of effector memory T cells, and impaired formation of effective responses to secondary T-cell stimulation with antigen. Antibody-mediated blockade of OX40-OX40L results in reduced induction of experimental autoimmune encephalitis (EAE) and collagen-induced arthritis.

Knowledge of basic signaling components utilized by TCR and costimulatory molecules has been exploited in the generation of chimeric antigen receptors (CARs). These engineered molecules possess an antigen-specific extracellular region (often a single-chain variable fragment), a transmembrane domain, and cytosolic signaling machinery from CD3 ζ-chain fused to a region from a costimulatory molecule (*e.g.,* CD28, OX40, 4-1BB) cytosolic region. Cross-linking of these CARs thus may provide both signal 1 (TCR) and signal 2 (costimulation). T cells modified to express CD19-specific CARs have recently been successful in treating B-cell malignancies and are being developed for use in other malignancies.

Spatial and Temporal Distribution of TCR Signaling Proteins

Recent methodological advances in live cell imaging and engineering of lipid bilayers that mimic physiological antigen-presenting surfaces have permitted exciting insights into the temporal–spatial relationships between signaling molecules. Following TCR engagement, a highly organized interface termed the *immunological synapse* (IS) forms between the T cell and the APC (Chapter 6). Within the IS, a number of transmembrane and cytoplasmic proteins are coordinately polarized and form a structure called the *supramolecular activation cluster* (SMAC).[29] The SMAC is composed of concentric rings with a central region (cSMAC) enriched in TCR and CD3/ζ. The more peripheral ring (pSMAC) is enriched for the integrin leukocyte function–associated antigen-1 (LFA-1). Inhibitory signaling molecules, such as the transmembrane phosphatase CD45, are excluded from the SMAC. These observations initially suggested that formation of the SMAC is required for signal initiation. The subsequent finding of a paucity of active signaling molecules located in the SMAC has led to alternative models in which the SMAC is involved in directed cytokine secretion or in signal termination.

Efficient formation of the T cell IS is heavily dependent on nucleation-promoting factors, such as Wiskott-Aldrich syndrome protein (WASP), which promotes reorganization of the actin cytoskeleton.[29] Complexes composed of TCR, active kinases, and adaptor molecules, termed *microclusters,* form at the immune synapse within seconds following TCR engagement.[30] TCR microclusters are dynamically regulated in space and time. Within 2 to 3 minutes of formation, microclusters migrate along polymerized actin from the periphery to the center of the IS, where dephosphorylation and dissociation of the components occur. Sustained TCR signaling depends on constant re-formation of microclusters containing SLP-76 and ZAP-70 at the periphery of the central zone. Costimulatory signals generated by CD28 or the integrin very late antigen-4 (VLA-4) can alter the dynamics of microcluster formation and movement and enhance T-cell activation. The functions of either the SMAC or the microclusters in TCR signal initiation and termination remain incompletely understood.

KEY CONCEPTS

Mechanisms of Tolerance: Central and Peripheral

- Central
  - Clonal deletion/AIRE
- Peripheral
  - Immune privilege
  - Anergy
  - Regulation

TOLERANCE

Tolerance is an inherent property of the immune system that governs the ability to respond against foreign antigens (nonself) without attacking the host (self). In normal hosts, tolerance protects against autoimmune tissue injury. The concept of immune tolerance predated understanding of the cellular and molecular bases for the phenomenon. Owen's experiments in cattle showed that a shared blood supply during development led to lifelong immune tolerance.[31] Billingham et al. expanded upon the concept of tolerance establishment in 1953 by showing that *in utero* inoculation with foreign tissue resulted in tolerance to foreign skin grafts later in life.[32] Layered and complementary mechanisms of immune system tolerance have developed to calibrate immune responsiveness with maintenance of maximum capacity for protective activation. The remainder of this chapter discusses critical roles played by T cells in the establishment and maintenance of tolerance.

Establishment of immune tolerance to self may occur either during T-lymphocyte development or after maturation in tissues.[33] *Central tolerance* refers to active deletion that is induced during development of T cells in the thymus (Chapter 8). Despite the presence of mechanisms that promote expression of a full repertoire of self-peptides (see below) in the thymus, establishment of central tolerance alone is inadequate to prevent tissue-damaging autoimmune responses. In healthy vertebrate hosts, a small percentage of the T cells that normally develop and emigrate from the thymus carry autoreactive TCR. Control of such "escaped" autoreactive cells in secondary lymphoid organs and tissues is achieved through a series of extrathymic processes; these are referred to in aggregate as *peripheral tolerance.*

Central Tolerance/Clonal Deletion

Techniques that allow "fate mapping" of clonal populations during T-cell development have opened windows into the molecular mechanisms of the thymic process of clonal deletion.[34,35] Monoclonal antibodies specific for Vβ17 TCR variable regions allow tracking of developing thymocytes that react with endogenous retrovirus-derived "self" peptides. Such peptides form the dominant antigens for Vβ17-bearing thymocytes. In the normal murine thymus, immature Vβ17-expressing $CD4^+CD8^+$ (double-positive [DP]) thymocytes are present, but the fraction of Vβ17-expressing cells in the mature CD4 single-positive (CD4SP) or CD8SP compartments declines precipitously. The paucity of Vβ17-expressing mature thymocytes thus suggests that clonal deletion can occur at the transition from DP to SP.

TCR transgenic mice expressing a single specificity TCR potentiated additional investigation into central tolerance. Mice bearing transgenic TCRs reactive with a Y chromosome–encoded antigen (H-Y) are among the oldest models of clonal deletion. Massive deletion of developing thymocytes results in small thymi in self-antigen–expressing (male) mice; only a small number of DP thymocytes avoid the apoptotic cell death induced by TCR engagement. In contrast, transgenic T cells develop normally in littermate female mice that lack H-Y antigen. Additional TCR transgenic models of central tolerance have shown that autoreactive thymocyte deletion can occur before, during, or after the DP stage.

Deletion for self-reactivity implies that the extent of autoreactivity of the developing thymocytes is systematically calibrated. A currently favored model is one that employs strength of signal as determined by a combination of TCR affinity and costimulation to assess the extent of self-reactivity.[36] Signals that exceed an intensity threshold lead to clonal deletion of cells bearing that TCR. This model is supported by data from animals in which the signaling machinery has been genetically altered to increase or decrease TCR signal intensity. For example, an increase in the number of TCR-associated ITAMs (presumably leading to increases in downstream signals) enhances clonal deletion in a TCR transgenic system.[37]

For clonal deletion to establish comprehensive self-tolerance, T cells must come into contact with all potential self-antigens during thymocyte maturation. It is easy to recognize how deletion of cells reactive to MHC and other widely expressed protein products can occur in the thymus. However, there are also self-antigens whose expression is restricted to a specific tissue or developmental time point.

The complex mechanisms whereby clonal deletion of developing thymocytes reactive with tissue-restricted antigens (TRAs)—expressed only in the pancreas or testes, for example—occurs are beginning to be unraveled. Some evidence supports the possibility that all TRAs are transported to the thymus by APCs, such as migratory DCs.[36] Another, non–mutually exclusive model holds that subsets of thymus-resident APCs may "ectopically" express TRAs. The leading candidate for controlling TRA expression within medullary thymic epithelial cells (mTECs) is the transcription factor AIRE (autoimmune regulator).[38] AIRE expression ectopically drives mTEC to express peptides from open reading frames representing TRA. Self-reactive developing T cells are thus exposed to TRAs in the thymus and are deleted. AIRE-deficient humans and mice develop autoimmune polyendocrinopathy–candidiasis–ectodermal dystrophy (APS-1 or APECED in humans; Chapter 35), emphasizing the role of central tolerance in preventing T cell–mediated autoimmune disease.

Isolated TRA presentation by mTEC is unlikely to fully account for successful clonal deletion, given the relative rarity of mTEC in thymic stroma and the abundance of autoreactive thymocytes. One potential mechanism for enhancing exposure of autoreactive cells to deleting TRA-driven signals is antigen presentation by additional thymic stromal cells. A recent observation that intercellular transfer of TRA between mTEC and thymic-resident or migrant DCs readily occurs supports this model.[36]

Peripheral Mechanisms of Tolerance

Clonal deletion, like all biological processes, is imperfect. Given the potential for autoimmune tissue damage should central tolerance mechanisms fail (as exemplified in those humans and mice carrying AIRE mutations), there must be "backup" mechanisms to control autoreactivity of T cells that manage to escape the thymus. These mechanisms include immune privilege, anergy, and regulation.

Immune Privilege

Medawar first described the concept of "immune privilege" more than 50 years ago. Typically described areas of immune privilege include the anterior chamber of the eye, the brain, and the fetus in pregnant females.[39] The eye and the brain are critical organs for basic survival functions and yet have a limited capacity for regeneration. Thus uncontrolled immune responses in these organs could have a detrimental effect on survival. A fetus expresses MHC derived from both parents; thus the mother's immune system must develop tolerance of the paternal antigen–bearing fetus to prevent pregnancy loss.

Immune-privileged tissues evade or suppress immune effector functions through multiple mechanisms. Cells of the eye, brain, and fetal villous trophoblast display low level or absent surface expression of classic MHC class Ia protein. This feature likely protects them from cytotoxic T-lymphocyte (CTL)–mediated lysis (Chapter 17). Ocular cells express proapoptotic cell surface molecules, such as CD95 ligand (CD95L) and TRAIL (TNF-related apoptosis inducing ligand) (Chapter 13). These TNF family members may contribute to apoptosis of infiltrating CD95-bearing T cells and other inflammatory cells bearing death receptors that serve as cognate ligands for TRAIL. In mice, the presence of CD95L on ocular cells is critical for the acceptance of corneal allografts. Soluble factors elaborated by ocular DCs likely contribute to immune privilege. These include the cytokines transforming growth factor-β (TGF-β) and IL-10, which may lead to induction and/or recruitment of Tregs, and migration inhibitory factor (MIF) that suppresses natural killer (NK) cell–dependent cytolytic capacity. In addition, ocular DCs may produce high levels of indoleamine oxidase, an enzyme that supports Treg differentiation.

T-Cell Anergy

Cellular proliferation and/or potentiation of T-cell effector function are not inevitable consequences of TCR engagement. Under some conditions, TCR ligation results in *anergy,* a cellular fate characterized by reduced proliferation and cytokine production in response to subsequent TCR engagement.[40] T-cell anergy can ensue either when the TCR is engaged without concomitant costimulation or when the ligand for the TCR is not of sufficient affinity to initiate the full spectrum of biochemical second messengers. Anergy is observed in T cells stimulated with metabolically inactivated APCs incapable of providing costimulation. Conversely, CD28 costimulation prevents the induction of anergy. Treatment of cultured T cells with IL-2 can overcome the anergic state *in vitro.* In addition to TCR and costimulatory signals, recent work has shown that environmental cues, such as nutrient and energy store availability, and the products of Tregs (described below) also control the anergy/activation fate choice.

T-cell anergy can be induced by activation of the calcineurin/NFAT pathway without concomitant increases in the RAS/ERK pathway–dependent activating protein-1 (AP-1) transcription factor activity (Fig. 12.6A). Relatively unopposed NFAT activity can be induced experimentally either by treatment with calcium ionophore or by stimulation through the TCR while blocking CD28 costimulation. Concurrent treatment of TCR-stimulated cells with protein synthesis inhibitors or with NFAT pathway inhibitors, such as cyclosporine, abrogates the development of later unresponsiveness, supporting the notion that anergy induction and maintenance depend on the transcription and translation of anergy-associated factors.[41]

Inhibition of Ras pathway function in anergized T cells correlates with biochemical and gene regulatory events (see Fig. 12.6B).[42] First, preferential FYN kinase–dependent activation of c-CBL within anergized cells leads to recruitment of the RAF kinase to the nucleotide exchange factor RAP1. The RAF-RAP1 association prevents RAF recruitment to RAS and thus restricts ERK pathway activation. Second, upregulation of NFAT-dependent transcription factors EGR2 and EGR3 results in transactivation of proteins implicated in the restraint of T-cell activation, including GRAIL (gene related to anergy in lymphocytes), CBL-b, and ITCH.[41] E3 ubiquitin ligase activity associated with a number of the latter factors is responsible for ubiquitin-mediated proteolysis of TCR signal–promoting molecules, such as PKCθ and RAS-GRP. Third, posttranslational modification of the positive TCR signaling mediator LAT is also observed in anergy.[43] As described above, LAT is normally palmitoylated and its membrane localization restricted to detergent-insoluble lipid rafts. In previously anergized T cells, however, LAT palmitoylation and phosphorylation are both decreased. Fourth, increased expression of members of the diacylglycerol kinase family (DGKα and DGKζ) is observed in T-cell anergy. Augmented DGK expression reflects augmented capacity for phosphorylation-dependent conversion of DAG, the key lipid mediator upstream of RAS signaling, into an inert metabolite phosphatidic acid (PA).[44]

Cellular sensing of adequate nutrient and energy stores required for optimal differentiation and proliferation regulates the anergy/activation T-cell fate decision. The importance of amino acid– and energy-sensing pathways was suggested by observations that anergy is induced in T cells activated in the presence of antagonists to leucine or glucose despite the presence of robust combined antigen receptor and costimulatory pathway signaling.[40] Several lines of evidence suggest that establishment of "metabolic anergy" after integration of signaling inputs from antigen recognition (TCR), immune (CD28, IL-2 receptor), and metabolic (*e.g.,* GLUT1) receptors and sensors is governed by mechanistic (formerly mammalian) target of rapamycin (mTOR).[45] TCR and CD28 ligation induce T-cell anergy rather than activation when the stimuli are given in the presence of rapamycin, a selective mTOR inhibitor, suggesting that optimal T-cell activation and avoidance of anergy requires activation of mTOR. mTOR is activated by signals that communicate abundant nutrients (*e.g.,* leucine-stimulated RAG proteins). Conversely, mTOR is inhibited by activated AMPK, an enzyme that responds to low energy stores reflected by increased ratios of adenosine monophosphate (AMP) to adenosine triphosphate (ATP). Through mediators, such as AKT, mTOR activation functions to promote cell cycle entry and prevent transcriptional activation of the anergy factors GRAIL and CBL-b.

KEY CONCEPTS

Immunosuppressive Drugs That Affect T-Cell Signaling

- Cyclosporine and tacrolimus inhibit T-cell receptor (TCR)–generated signals:
  - Anti-TCR antibodies block TCR signals
  - CTLA-4Ig blocks CD28 signals
  - Rapamycin inhibits mechanistic (formerly mammalian) target of rapamycin (mTOR) activation
- Inhibitory receptor blockade enhances antitumor function of suppressed/exhausted T cells
  - Anti-CTLA4 antagonist antibodies permit CD28-CD80/86 interactions
  - Anti-PD1/anti-PDL1 antibodies prevent suppressive interactions between tumor-expressed PDL1 and PD1 on tumor-infiltrating T cells

Regulation

Subsets of T lymphocytes can enforce tolerance through active regulation of autoreactive immune responses. Tregs can suppress effector functions of other immune cells of both myeloid and lymphoid lineages (Chapter 18). The most extensively studied—and arguably the most important for tolerance induction—Treg-sensitive immune functions are proliferation and cytokine production by naïve and memory conventional T cells. Tregs inhibit these processes through both cell contact–dependent and soluble molecule secretion mechanisms.[46] By regulating the activation and proliferation of antigen-specific effectors, Tregs promote tolerance to self and suppress autoimmunity *in vivo.*

Identification of Tregs was initially described on the basis of the correlation between high CD25 expression and the potent suppressive activity of a subset of CD4 T cells. Most Tregs also express GITR, CD103, CTLA-4, lymphocyte activation gene-3 (*LAG-3*), and low levels of CD45RB. Although no single surface marker that specifies Tregs exists, CD4 T-cell expression of the transcription factor Foxp3 correlates tightly with suppressive capacity and strongly suggests a unique differentiation pathway for Tregs.[47] The absence of Foxp3, occurring either via spontaneous mutation (exemplified by the scurfy mouse, which develop fatal autoimmune disease) or through targeted disruption of the gene, leads to the complete loss of T cells with regulatory activity. Conditional deletion of Foxp3 in peripheral T cells results in loss of the suppressive phenotype. Conversely, overexpression of Foxp3 by transgenesis or retroviral methods leads to an excess of T cells with regulatory activity. Together, these findings suggest that Foxp3 is both necessary and sufficient for Treg-suppressive functions. The bulk of Foxp3-positive Tregs develop in the thymus. Dynamic Foxp3 expression and development of suppressive capacity can be observed in naïve CD4 T cells after exposure to TGFβ or retinoic acid; recent work strongly suggests that inducible Treg may have nonredundant functions in suppressing chronic inflammation.[48]

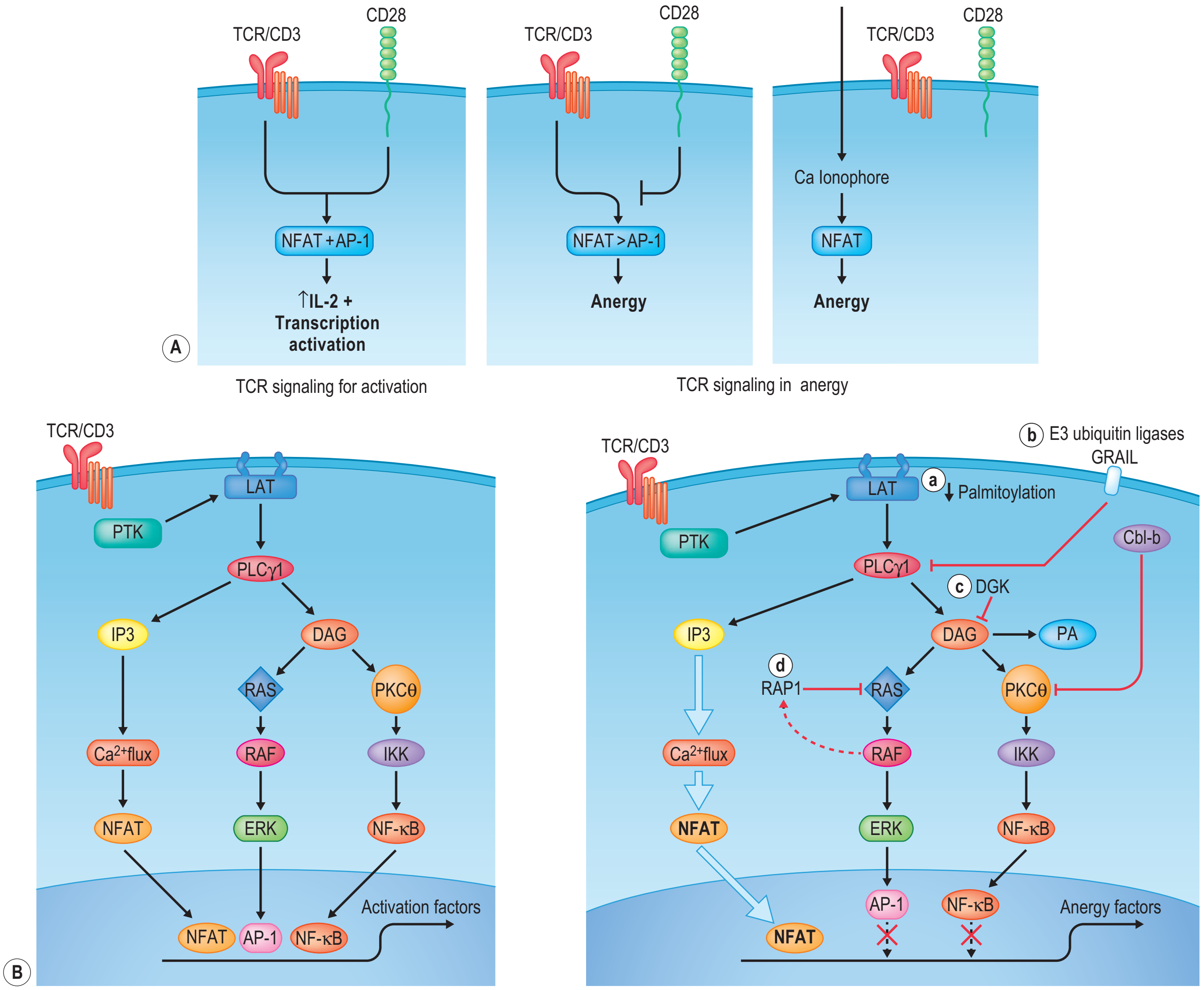

A
TCR/CD3
CD28
NFAT + AP-1
↑IL-2 +
Transcription
activation
TCR signaling for activation
NFAT > AP-1
Anergy
Ca Ionophore
NFAT
Anergy
TCR signaling in anergy
B
TCR/CD3
LAT
PTK
PLCγ1
IP3
DAG
RAS
PKCθ
Ca2+flux
RAF
IKK
NFAT
ERK
NF-κB
Activation factors
NFAT
AP-1
NF-κB
b E3 ubiquitin ligases
GRAIL
a Palmitoylation
Cbl-b
c DGK
PA
d
RAP1
Anergy factors
AP-1
NF-κB
NFAT

FIG 12.6 **T-Cell Anergy Induction and Maintenance Correlates With Differential Activation of T-Cell Receptor (TCR)–Dependent Second-Messenger Signaling Cascades.** (A) Stimulation of T cells by cross-linking of both TCR/CD3 and CD28 lead to upregulation of both the nuclear factor of activated T cells (NFAT) and activating protein-1 (AP-1) transcription factors, leading to increased transcription of the interleukin-2 (*IL-2*) gene and activation. An imbalance of activated NFAT and AP-1 by blockade of CD28 signals *(middle panel)* or calcium ionophore *(right panel)* leads to an anergic phenotype. (B) TCR signaling required for full T-cell activation features Calcium flux-, RAS-, and PKC-dependent biochemical events leading to cooperative transcriptional regulation by NFAT, AP-1, and NF-κB transcription factors *(left panel).* In anergized cells, TCR-dependent signaling is differentially impaired (DAG-dependent events more so than calcium-dependent events) through multiple mechanisms: (a) decreased palmitoylation of LAT results in diminished recruitment to the immunological synapse; (b) upregulated GRAIL and CBL-b degrade positive signaling regulators PLCγ1 and PKCθ; (c) diacylglycerol kinases (DGKs) convert DAG (PKC and RAS activator) to phosphatidic acid (PA); (d) active RAP1 recruits RAF, thus preventing RAS-mediated signaling to ERK. Anergy mechanisms and mediators are highlighted in red.

In humans, mutations in FOXP3 account for a majority of cases of immune dysfunction/polyendocrinopathy/enteropathy/X-linked (IPEX) syndrome. The signs and symptoms of FOXP3 deficiency are similar in mice and humans. Affected human males develop an autoimmune syndrome consisting of lymphoproliferation, thyroiditis, insulin-dependent diabetes mellitus, enteropathy, and other immune disorders.

Tregs can inhibit conventional T-cell responses through secretion of suppressive cytokines, induction of T-cell apoptosis, or repression of APC function.[49] Key Treg-secreted cytokines include IL-10, TGFβ, and IL-35; each of these molecules has the capacity to induce cell cycle arrest. Tregs, via high expression of CD25 (IL-2Rα), may compete with neighboring T effector cells for limited supplies of IL-2. The resulting growth factor deprivation can kill proliferating T cells in a B cell lymphoma 2–interacting (Bcl2-interacting) mediator (BIM)-dependent manner. Tregs preferentially express Galectin-1, a β-galactoside-binding protein that ligates CD45 and other lymphocyte surface molecules and that may suppress lymphocyte activation through induction of cell cycle arrest. Treg surface proteins that work to effectively decrease efficiency of antigen presentation by DCs include CTLA-4, which interferes with CD28-dependent T-cell costimulation; LAG3, which binds to MHC class II molecules and prevents DC maturation; and Nrp1, which mediates prolonged Treg/DC interactions that may restrict access of effector T cells to DCs.

SUMMARY AND FUTURE DIRECTIONS

ON THE HORIZON

- Better understanding of T-cell intracellular signaling pathways will lead to improved immunomodulatory therapies
- Promising targets for manipulation of T-cell signaling include:
  - SYK inhibition
  - AKT inhibition
  - Tec kinase inhibition in cancer
- Targets for relief of negative checkpoints on T-cell activation for immunoenhancement in cancer or infection include:
  - LAG3, TIGIT, TIM3, BTLA, VISTA

Antigen receptor signaling controls key T-cell developmental checkpoints as well as the fate and behavior decisions of mature T cells. Membrane-proximal TCR signaling depends on tyrosine phosphorylation events that are coupled to “second messenger” cascades through adaptor proteins. TCR affinity for ligand is the major determinant for life-or-death decisions by immature thymocytes that result in establishment of a broadly self-tolerant mature T-cell repertoire. In peripheral T cells, auxiliary signals from costimulatory molecules, coreceptors, growth factor receptors, and metabolic sensors all shape the transcriptional regulation programs that govern cell function. T cells proliferate and produce cytokines when members of the CD28 or TNFR families of receptors signal in tandem with the TCR. Immunoreceptor and costimulatory signals are also critical in maintaining immune tolerance through clonal deletion, anergy, and peripheral suppression by Tregs. Coinhibitory signals that restrain T-cell antitumor functions may be antagonized for immunotherapy of cancer. More detailed understanding of T-cell intracellular signaling pathways will lead to more specific, targeted immunomodulatory therapies.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Weiss A, Littman DR. Signal transduction by lymphocyte antigen receptors. Cell 1994;76:263–74.
2. Guy CS, Vignali DA. Organization of proximal signal initiation at the TCR:CD3 complex. Immunol Rev 2009;232:7–21.
3. Stanford SM, Rapini N, Bottini N. Regulation of TCR signalling by tyrosine phosphatases: from immune homeostasis to autoimmunity. Immunology 2012;137:1–19.
4. Rhee I, Veillette A. Protein tyrosine phosphatases in lymphocyte activation and autoimmunity. Nat Immunol 2012;13:439–47.
5. Notarangelo LD. Immunodeficiency and immune dysregulation associated with proximal defects of T cell receptor signaling. Curr Opin Immunol 2014;31:97–101.
6. Gomez-Rodriguez J, Kraus ZJ, Schwartzberg PL. Tec family kinases Itk and Rlk / Txk in T lymphocytes: cross-regulation of cytokine production and T cell fates. FEBS J 2011;278:1980–9.
7. Smith-Garvin JE, Koretzky GA, Jordan MS. T cell activation. Annu Rev Immunol 2009;27:591–619.
8. Hogan PG, Lewis RS, Rao A. Molecular basis of calcium signaling in lymphocytes: STIM and ORAI. Annu Rev Immunol 2010;28:491–533.
9. Krishna S, Zhong XP. Regulation of lipid signaling by diacylglycerol kinases during T cell development and function. Front Immunol 2013;4:178.
10. Burbach BJ, Medeiros RB, Mueller KL, et al. T cell receptor signaling to integrins. Immunol Rev 2007;218:65–81.
11. Jun JE, Rubio I, Roose JP. Regulation of ras exchange factors and cellular localization of ras activation by lipid messengers in T cells. Front Immunol 2013;4:239.
12. So T, Croft M. Regulation of PI-3-Kinase and Akt signaling in T lymphocytes and other cells by TNFR family molecules. Front Immunol 2013;4:139.
13. Stark AK, Sriskantharajah S, Hessel EM, et al. PI3K inhibitors in inflammation, autoimmunity and cancer. Curr Opin Pharmacol 2015;23:82–91.
14. Shi JH, Sun SC. TCR signaling to NF-kappaB and mTORC1: expanding roles of the CARMA1 complex. Mol Immunol 2015;68:546–57.
15. Jordan MS, Singer AL, Koretzky GA. Adaptors as central mediators of signal transduction in immune cells. Nat Immunol 2003;4:110–16.
16. Balagopalan L, Kortum RL, Coussens NP, et al. The linker for activation of T cells (LAT) signaling hub: from signaling complexes to microclusters. J Biol Chem 2015;290:26422–9.
17. Simeoni L, Lindquist JA, Smida M, et al. Control of lymphocyte development and activation by negative regulatory transmembrane adapter proteins. Immunol Rev 2008;224:215–28.
18. Chen L, Flies DB. Molecular mechanisms of T cell co-stimulation and co-inhibition. Nat Rev Immunol 2013;13:227–42.
19. Artyomov MN, Lis M, Devadas S, et al. CD4 and CD8 binding to MHC molecules primarily acts to enhance Lck delivery. Proc Natl Acad Sci USA 2010;107:16916–21.
20. Baumeister SH, Freeman GJ, Dranoff G, et al. Coinhibitory pathways in immunotherapy for cancer. Annu Rev Immunol 2016;34:539–73.
21. Sharpe AH. Mechanisms of costimulation. Immunol Rev 2009;229: 5–11.
22. Rudd CE, Taylor A, Schneider H. CD28 and CTLA-4 coreceptor expression and signal transduction. Immunol Rev 2009;229:12–26.
23. Linterman MA, Denton AE, Divekar DP, et al. CD28 expression is required after T cell priming for helper T cell responses and protective immunity to infection. Elife 2014;3. doi:10.7554/eLife.03180.
24. Zhang R, Huynh A, Whitcher G, et al. An obligate cell-intrinsic function for CD28 in Tregs. J Clin Invest 2013;123:580–93.
25. Qureshi OS, Zheng Y, Nakamura K, et al. Trans-endocytosis of CD80 and CD86: a molecular basis for the cell-extrinsic function of CTLA-4. Science 2011;332:600–3.
26. Wherry EJ, Kurachi M. Molecular and cellular insights into T cell exhaustion. Nat Rev Immunol 2015;15:486–99.
27. Turnis ME, Andrews LP, Vignali DA. Inhibitory receptors as targets for cancer immunotherapy. Eur J Immunol 2015;45:1892–905.
28. Croft M. Control of immunity by the TNFR-related molecule OX40 (CD134). Annu Rev Immunol 2010;28:57–78.
29. Le Floc'h A, Huse M. Molecular mechanisms and functional implications of polarized actin remodeling at the T cell immunological synapse. Cell Mol Life Sci 2015;72:537–56.
30. Yokosuka T, Saito T. Dynamic regulation of T cell costimulation through TCR-CD28 microclusters. Immunol Rev 2009;229:27–40.
31. Owen RD. Immunogenetic consequences of vascular anastomoses between bovine twins. Science 1945;102:400–1.

32. Billingham RE, Brent L, Medawar PB. Actively acquired tolerance of foreign cells. Nature 1953;172:603–6.
33. Bluestone JA, Bour-Jordan H, Cheng M, et al. T cells in the control of organ-specific autoimmunity. J Clin Invest 2015;125:2250–60.
34. von Boehmer H, Kisielow P. Negative selection of the T cell repertoire: where and when does it occur? Immunol Rev 2006;209:284–9.
35. Gallegos AM, Bevan MJ. Central tolerance: good but imperfect. Immunol Rev 2006;209:290–6.
36. Klein L, Kyewski B, Allen PM, et al. Positive and negative selection of the T cell repertoire: what thymocytes see (and don't see). Nat Rev Immunol 2014;14:377–91.
37. Love PE, Hayes SM. ITAM-mediated signaling by the T cell antigen receptor. Cold Spring Harb Perspect Biol 2010;2:a002485.
38. Taniguchi RT, Anderson MS. The role of Aire in clonal selection. Immunol Cell Biol 2011;89:40–4.
39. Forrester JV, Xu H, Lambe T, et al. Immune privilege or privileged immunity? Mucosal Immunol 2008;1:372–81.
40. Chappert P, Schwartz RH. Induction of T cell anergy: integration of environmental cues and infectious tolerance. Curr Opin Immunol 2010;22:552–9.
41. Mueller DL. Mechanisms maintaining peripheral tolerance. Nat Immunol 2010;11:21–7.
42. Valdor R, Macian F. Induction and stability of the anergic phenotype in T cells. Semin Immunol 2013;25:313–20.
43. Hundt M, Tabata H, Jeon MS, et al. Impaired activation and localization of LAT in anergic T cells as a consequence of a selective palmitoylation defect. Immunity 2006;24:513–22.
44. Zhong XP, Guo R, Zhou H, et al. Diacylglycerol kinases in immune cell function and self-tolerance. Immunol Rev 2008;224:249–64.
45. Powell JD, Delgoffe GM. The mammalian target of rapamycin: linking T cell differentiation, function, and metabolism. Immunity 2010;33:301–11.
46. Chaudhry A, Rudensky AY. Control of inflammation by integration of environmental cues by regulatory T cells. J Clin Invest 2013;123:939–44.
47. Josefowicz SZ, Lu LF, Rudensky AY. Regulatory T cells: mechanisms of differentiation and function. Annu Rev Immunol 2012;30:531–64.
48. Schmitt EG, Williams CB. Generation and function of induced regulatory T cells. Front Immunol 2013;4:152.
49. Shevach EM. Mechanisms of Foxp3$^+$ T regulatory cell-mediated suppression. Immunity 2009;30:636–45.

MULTIPLE-CHOICE QUESTIONS

1. Individual mutations in T-cell signaling proteins can lead to immune deficiency, dysregulation, or lymphocyte proliferation. Among the following molecules, which one is primarily associated with lymphocyte hyperproliferation and an increased risk of autoimmunity?
   A. ZAP-70
   B. SLP-76
   C. CD45
   D. CD95

2. Which of the following are mechanisms of central tolerance
   A. Immune privilege
   B. Treg regulation
   C. Clonal deletion/AIRE
   D. Anergy

3. In the clinic, monoclonal antibodies directed against T cell signaling coinhibitory molecules are proving to be efficacious against cancers. Which of the following is a coinhibitory molecule that where blockade has proven to be helpful in tumor immunotherapy?
   A. CTLA-4
   B. CD28
   C. OX40
   D. CD3

Regulated Necrosis and Its Immunogenicity

Wulf Tonnus, Andreas Linkermann

For more than a century and a half, we have known that the process of regulated cell death (RCD) is an integral part of life. The first insights into the mechanisms underlying RCD came from the identification of apoptotic pathways in 1972.[1] For almost 40 years, the apoptotic process of RCD was viewed as the opposite of necrosis, which was considered accidental (and therefore unregulated). In 2008, however, it was realized that necrosis could be antagonized under certain conditions. This led to the appreciation that necrosis could also be a genetically determined, regulated process.

The current challenge is to elucidate the various genes and pathways that take part in regulated physiological and pathophysiological cell death and to better understand why there are multiple pathways. It has become clear from these studies that the various pathways create differences in the immunogenicity of the byproducts of necrosis. As tissue injury and inflammation are tightly linked, these discoveries fuel the need to better understand how necrosis influences pathologies, including autoimmune diseases, transplant rejection, and cancer. This chapter focuses on the most clearly identified regulated necrosis (RN) pathways. Our "selection" of the RN pathways discussed is based on the current appreciation of the importance of these pathways, either physiologically or pathophysiologically, to differences in the immunogenicity of these RN pathways and their role in diseases (Fig. 13.1).

CELL DEATH AND DAMAGE-ASSOCIATED MOLECULAR PATTERNS—THE CONCEPT OF NECROINFLAMMATION

Pattern recognition receptors (PRRs) bind a wide range of damage-associated molecular patterns (DAMPs). These include exogenous stimuli such as lipopolysaccharides (LPS) as well as an array of intracellular content.

It is now understood that necrosis triggers inflammation and that inflammation can, in turn, lead to further RN. This observation gave rise to the concept of necroinflammation. Necroinflammation can create an autoamplificatory feedback loop that results in DAMP release between organs. For example, ischemic or traumatic tissue in the lung can initiate a positive feedback loop that leads to acute respiratory distress syndrome (ARDS).[2,3] Recent observations have indicated that the DAMPs released by this type of feedback after renal transplantation can lead to RN in the lung as well.[4]

For years, immunology focused on the issue of self versus nonself recognition. However, within the past two decades, the danger/injury model has become more and more prominent. Briefly, this model proposes that cellular stress activates the innate immune system, thus leading to an inflammatory response that can, under certain circumstances, then induce a specific adaptive response.

KEY CONCEPTS

Necroinflammation—An Autoamplificatory Feed-Forward Loop

- Induced by cells dying by necrosis (*e.g.*, in hypoxia)
- Associated proinflammatory damage-associated molecular pattern (DAMP) release
- Immune cell infiltration
- Regulated necrosis (RN) in parenchymal and endothelial cells induced by immune cells
- RN in inflammatory cells (*e.g.*, macrophages undergoing pyroptosis)
- More necrotically dying cells
- DAMP transfer to remote organs
- Remote organ injury

Key to this is the concept are DAMPs (Chapter 3). DAMPs include a wide array of different stimuli, including exogenous stimuli (*e.g.*, LPS) and intracellular contents (*e.g.*, uremic acid and high-mobility group box-1 [HMGB1]).[5,6] These are sensed by either "classic" PRRs or "nonclassic" receptors found on the surface or within cells of innate immunity. When cells die by RN, these DAMPs become accessible to the immune system in the extracellular space. They are sensed especially by cells of the innate immune system. Dendritic cells (DCs) (Chapter 6) are activated by directly sensing DAMPs via surface receptors (*e.g.*, calreticulin [CALR]–CD91, adenosine triphosphate [ATP]–P2X7R, or HMGB1–Toll-like receptor 4 [TLR4]). Sensing of DAMPs by monocytes leads to activation of inflammasomes, which then promote expression of mature inflammatory cytokines (*e.g.*, interleukin-1β [IL-1β] and IL-18) (Chapter 9). Natural killer (NK) cells (Chapter 17) also directly interact with injured cells. However both monocytes and NK cells give rise to inflammation and thus support the maturation of immature DCs.

The infiltration of innate immune cells into the inflamed tissues and their subsequent inflammatory response is thought to account for a significant part of the overall damage to the organ, above and beyond initial necrosis itself. Mechanisms that mediate this deterioration may include the edema that arises following capillary leakage, induction of cell death by innate immune cells with the effect of decreased organ function, and dysregulated partial oxygen pressure within the area of inflammation/necrosis. Also, mature DCs interact with CD8 T cells (Chapter 17) via major histocompatibility complex (MHC) class I interaction and CD4 T cells via MHC class II interaction (Chapter 5), thus shifting the immune response to an antigen specific response, as cytotoxic T cells and active B cells now specifically attack stressed cells (Fig. 13.2). However, in addition to these DAMPs, other intracellular content has the potential to be sensed as a target as well, which could lead to the development of autoimmunity.

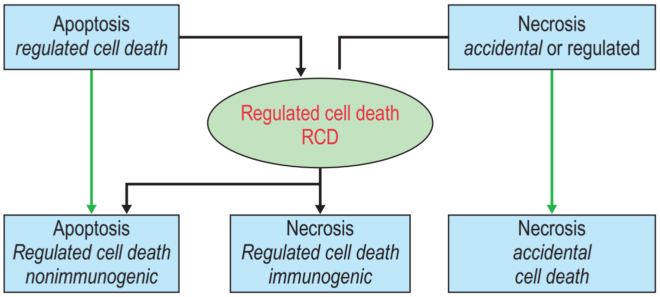

Apoptosis
regulated cell death
Necrosis
accidental or regulated
Regulated cell death
RCD
Apoptosis
Regulated cell death
nonimmunogenic
Necrosis
Regulated cell death
immunogenic
Necrosis
accidental
cell death

FIG 13.1 New Concept of Regulated Cell Death. When apoptosis was first identified, two models of regulated cell death were recognized: apoptosis and necrosis. More recently, however, cell death is divided into regulated cell death (including apoptosis and regulated necrosis) and accidental cell death (instant necrosis). A second parameter classifies cell death as either immunogenic or nonimmunogenic.

When challenged with injurious conditions, such as ischemia–reperfusion in solid-organ transplantations, options for cellular fate are directly succumbing via accidental cell death (ACD), restoring cellular homeostasis, or dying by RCD. ACD happens passively under certain conditions, such as extreme heat, and occurs nearly immediately in an uncontrolled fashion. Because of the direct loss of membrane integrity, huge amounts of DAMPs are released. This leads to massive recruitment of innate immunity and local inflammation. As this kind of stimulus is not given in ischemia–reperfusion, this is not further discussed here. However, if these cells do not explode directly, they balance on the edge. The issue then becomes whether balance will be restored (*e.g.*, by means of autophagy or unfolded protein response following endoplasmic reticulum [ER] stress) or, if the damage is too great, will balance be lost and RCD be induced?

Within the cells that succumb to cell death, the necroinflammatory loop starts with perturbation of intracellular homeostasis (class V DAMPs), which triggers the heat shock response, a system of critical importance for correct protein folding. Secreted or surface-exposed heat shock proteins (HSPs) can be sensed by either classic (*e.g.*, TLR2/4) or nonclassic (*e.g.*, CD91) receptors on DCs. In parallel, ER stress (which is often result of reactive oxygen species [ROS] generation) leads to secretion of calreticulin (an ER chaperone referred to as "CALR"), which then acts extracellularly as a class I DAMP by binding to CD91. This leads inactive dendritic cells (iDCs) to be activated. For full activation, however, an inflammatory milieu is required.

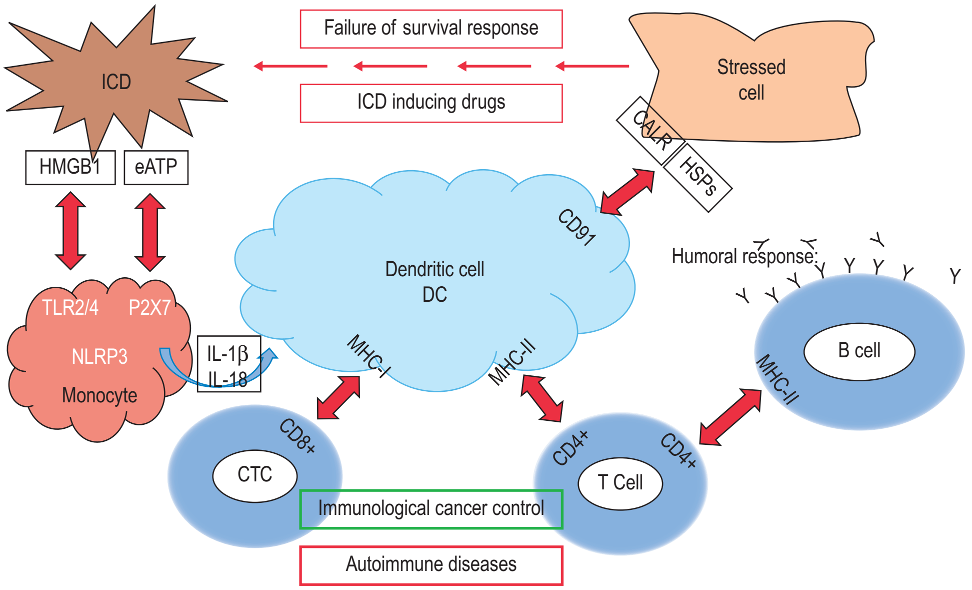

ICD
Failure of survival response
ICD inducing drugs
Stressed cell
CALR
HSPs
HMGB1
eATP
CD91
Dendritic cell
DC
Humoral response
TLR2/4
P2X7
NLRP3
Monocyte
IL-1β
IL-18
MHC-I
MHC-II
B cell
MHC-II
CD8+
CTC
CD4+
T Cell
CD4+
Immunological cancer control
Autoimmune diseases

FIG 13.2 Central Role of Dendritic Cells (DCs). Innate immunity plays a crucial role in necroinflammation. When cells are stressed, they release damage-associated molecular patterns (DAMPs), such as calreticulin (CALR) and heat shock proteins (HSPs). These DAMPs are sensed by scavenger receptors, such as CD91. This leads to partial activation of inactive DCs. If those cells fail to restore metabolic balance, they succumb to immunogenic cell death (ICD) and release additional DAMPs, which activate inflammasomes, such as LRR and PYD domains–containing protein 3 (NLRP3) in monocytes. Activation of the inflammasomes leads to the release of interleukins, which help DCs to gain full activation. These fully activated DCs use major histocompatibility complex (MHC) class I and class II antigen presentation to stimulate cytotoxic CD8 T cells (CTCs) and CD4 T cells, which interact with B cells to trigger a humoral response. Activation of the immune response by necroinflammation can thus promote the control of cancer by immunogenic means. However, the same mechanisms can promote the development of autoimmune diseases.

An inflammatory milieu is generated mainly by monocytes. These become active as the survival response fails and the formerly stressed cell succumbs to RN. Ferroptosis is the typical mode of RN within this process. The ferroptotic cell can "leech" redox equivalents, such as nicotinamide adenine dinucleotide phosphate (NADPH), from neighbored cells (see below), which subsequently undergo secondary necrosis. This secondary necrosis seems to involve necroptosis, but the mechanism of necroptosis induction in this case remains to be elucidated. RN releases high amounts of DAMPs, so it can be named immunogenic cell death (ICD) in this context as well.

Adenosine triphosphate (ATP) as a DAMP can reach the extracellular space (eATP). When the plasma membrane is intact, the autophagy machinery seems to be necessary for ATP export. However, as soon as the membrane ruptures upon RN, export is obviously no longer required. eATP represents a class II DAMP as it activates the purinergic-receptor, ligand-gated ion channel P2X7R. This results in potassium influx, which is sensed by the NACHT, LRR and PYD domains-containing protein 3 (NLRP3) inflammasome and leads to IL-1β and IL-18 maturation. eATP is thus extremely immunogenic.

Upon membrane rupture, HMGB-1 also reaches the extracellular space and is therefore the prototype of a RN-DAMP. iHMGB-1 acts via binding to TLR4 of monocytes and activates the NLRP3 inflammasome, as well.

The DAMPs from the stressed cells and the inflammatory cytokines fully activate iDCs, which then create a *per-se* cytotoxic inflammatory environment and prime naïve CD4 and CD8 T cells, thereby inducing an antigen-specific response. This response is the basis for certain conditions, such as chemotherapy in cancer. To stably control (or wipe out) a cancer, a specific response against tumor epitopes is required, so induction of RN should be beneficial. In contrast, in some settings, such as solid-organ transplantation, an antigen-specific response induced by DAMPs can give rise to antibody-mediated rejection, which typically starts after cortisone tapering. In this setting, RN is detrimental. The activation of an antigen-specific response results in additional cells being attacked. As they succumb, these dying cells replenishing the DAMP pool, further attracting cells of innate immunity and further promoting DC maturation and T-cell priming. Thus tissue injury amplifies while the loop closes (Fig. 13.3).

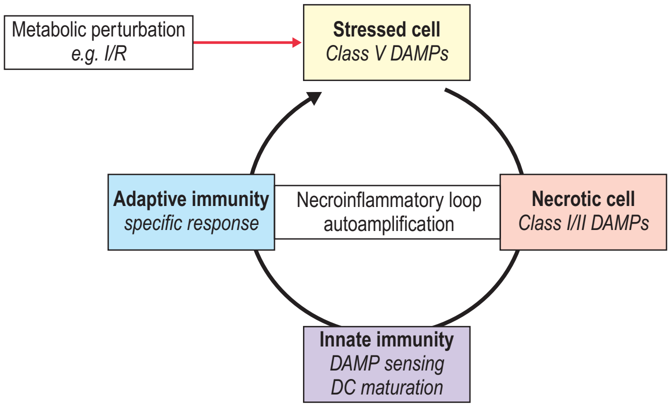


FIG 13.3 **The Autoinflammatory Loop.** The necroinflammatory loop is triggered by metabolic perturbations (*e.g.,* ischemia–reperfusion [I/R]). The so-stressed cells release class V danger-associated molecular patterns (DAMPs). In the next step, now-necrotic cells release class I/II DAMPs, which can be sensed by cells of innate immunity, such as dendritic cells (DCs), which are then stimulated to mature. Via mechanisms described in Fig. 13.2, adaptive immunity is empowered to unleash a specific response to (neo-)antigens. This leads to further stressed and dying cells, which, in turn, can themselves then release DAMPs and thus create an autoinflammatory loop.

REGULATED CELL DEATH REGULATES ITS IMMUNOGENICITY IN AN ACTIVE MANNER

Apoptosis is the prototype of nonimmunogenic cell death. Avoidance of immunogenicity is achieved by active processing and covering of DAMPs during the apoptosis program. For example, DNA is fragmented, cell organelles are consumed, and proteins degraded. Moreover, everything is packed in blebbing membranes. It is a matter of debate, however, if apoptosis should be considered somehow immunogenic as well, as some DAMPs (*e.g.,* HMGB-1) are still released and phosphatidylserines, which are usually located in the inner leaflet of the plasma membrane, are flipped outside.

Phosphatidylserine surface expression serves as an "eat me" signal for macrophages and other phagocytes, which can then act to remove these cells in an immunologically silent manner. Thus apoptotic cells recruit cells of innate immunity, but in a manner that is programmed to not induce further inflammation (as long as the uptake works properly, see "LAP" below). Thus the debate about the inflammatory potential of apoptosis might be more or less of an academic nature. The highly complex program of apoptosis appears to have evolved to prevent necroinflammation and is therefore favored in physiological settings that result in regular cell turnover (Fig. 13.4).

Under metabolic pressure, such as ischemia–reperfusion injury, however, tissue damage occurs primarily by RN, and thus the necroinflammatory loop *per se* tends to refuel itself. Therefore mechanisms are required to keep this in check. One such mechanism is the active production of IL-33 by cells undergoing necroptosis (see below), which acts to limit the immunogenicity to a certain microenvironment,[7] as IL-33 stabilizes regulatory T cells (Tregs) through ST2 receptors.[8] Cells undergoing pyroptosis (see below) actively produce and secrete IL-1β and IL-18 upon their demise. Both of these ILs are highly proinflammatory. As pyroptosis is typical for cells of innate immunity, this might be an alert function of this first-line defense. In conclusion, even in RN, cells regulate their inflammatory potential in both proinflammatory and antiinflammatory ways. When balanced, this generates a beneficial environment for both defense and regeneration.

Necroinflammation can be opposed by the uptake of extracellular debris by monocytes. This mechanism is referred to as LC3-associated phagocytosis (LAP).[9] Dysfunction of proteins of this noncanonical autophagy pathway (*e.g.,* ATG5, ATG7, ATG16L, or Rubicon) cause chronic inflammation as a result of inadequate processing of extracellular disposal. This also extends to apoptotic debris, further highlighting the importance of correct clearance.

In humans, ATG16L mutations are markedly associated with the development of autoimmunity.[10] Usually, IL-10 is secreted by monocytes engulfing dying cells to dampen the immune response. However, LAP-deficient monocytes actively produce proinflammatory IL-1β and IL-6 instead. Orchestrated with other elevated proinflammatory cytokines (IL-17, IL-18, IL-23),

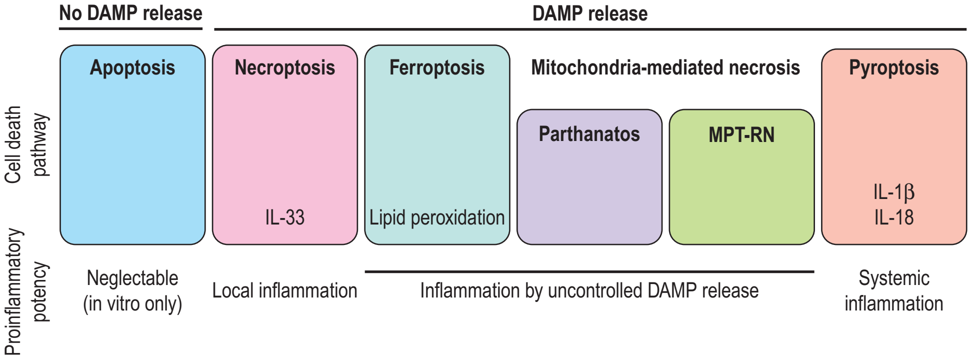


FIG 13.4 **Hypothetical Model of the Immunogenicity of Regulated Necrosis (RN) Pathways.** Apoptosis is nonimmunogenic, whereas loss of plasma membrane integrity results in the release of damage-associated molecular patterns (DAMPs)[5,6] into the extracellular space in all other pathways. During necroptosis, antiinflammatory cytokines, such as interleukin-33 (IL-33), are actively released, which triggers regulatory T cell (Treg) recruitment and thereby may limit the immunogenic response to a distinct microenvironment. In contrast, during ferroptosis or mitochondrial necrosis, no active production of cytokines or immunomodulatory factors has been described, and the uncontrolled DAMP release accounts for the immunogenic potency of these pathways. Pyroptosis, the most immunogenic cell death known so far, encompasses active production of proinflammatory cytokines, such as IL-1β and IL-18, resulting in a long-lasting, systemic inflammatory response. The difference in immunogenicity may explain the evolutionary preservation of several pathways of RN.

a low immunogenic stimulus can thus induce a strong immune response. In the presence of this stimulus, common self epitopes are classified as DAMPs by the immune system, which explains why mice deficient in LAP develop a lupus-like autoinflammatory disease. As a result, functional LAP may be required for prevention of memory B-cell priming during solid-organ transplantation as well.

SIGNALING PATHWAYS OF REGULATED CELL DEATH

RCD is an umbrella term for any genetically determined signaling pathway that results in cellular demise.[11] RCD thus includes both nonimmunogenic apoptosis and immunogenic RN. Fig. 13.5 provides an overview of the most important pathways of RCD, including caspase-dependent and caspase-independent forms of RCD.

CASPASE-DEPENDENT CELL DEATH

Apoptosis

Apoptosis is a complex program for nonimmunogenic cellular demise that is regulated by caspases. During the first hours of apoptosis, cells maintain their plasma membrane integrity and therefore, by definition, are not necrotic. The apoptotic program may be activated via the intrinsic (mitochondrial) or the extrinsic (death receptor) pathways.

The intrinsic pathway is controlled by different members of the Bcl-2 protein family, such as Bcl-2 or BAX, and therefore reacts to intracellular changes (*e.g.*, DNA damage) by forming transient pores, a process referred to as mitochondrial outer membrane permeabilization (MOMP). The intrinsic pathway reacts to internal stimuli, whereas the default activation of the extrinsic pathway occurs following an outside-in signaling, typically mediated by death receptors, such as Fas (also known as CD95 or Apo1) or the tumor necrosis factor receptor 1 (TNFR1). These death receptors, upon trimerization/hexamerization, recruit downstream molecules via death domains (DDs), such as TRADD or FADD, to form a receptor-associated platform, the death-inducing signaling complex (DISC). If nuclear factor-κB (NF-κB)–signaling (the canonical response to tumor necrosis factor receptor 1 [TNFR1] activation) is inhibited and/or receptor-interacting protein kinase 1 (RIPK1) polyubiquitination is lost, this complex is capable of activating initiator caspases (*e.g.*, caspase-8 [CASP-8] or caspase-10 [CASP-10]). Together with RIPK1, RIPK3, and FADD, a CASP-8–cFLIP heterodimer forms a cellular signaling platform named a ripoptosome that usually prevents necroptotic signaling (see below) by cleaving RIPK1, RIPK3, and Cylindromatosis (CYLD) while not proteolytically activating downstream effector caspases, such as CASP-3, CASP-6, and CASP-7. Should CASP-8 be activated, however, it will form homodimers to activate named downstream caspases, which induce apoptosis.

RIPK1 polyubiquitination represents a major checkpoint for a cellular decision of either survival or RCD. For this reason, it is tightly regulated. Inhibitors of apoptosis 1 and 2 (cIAP1/2) and the linear ubiquitination complex (LUBAC) add ubiquitin chains, whereas OTULIN, CYLD, and A20 remove them. Polyubiquitination is required for NEMO-dependent canonical NF-κB signaling and mitogen-activated protein kinase (MAPK) activation. Both provide a survival signal. In contrast, RCD is licensed on loss of polyubiquitination. Remarkably, both pro-survival NF-κB signaling and necroptosis lead to local inflammation, whereas apoptosis is noninflammatory.

The physiological role of apoptosis has been questioned following the observation that CASP-8 deficient mice (which die *in utero*) can be rescued on a RIPK3-deficient background to become viable and fertile. However, CASP-8 is of critical importance only for the extrinsic pathway of apoptosis, and BAX-/BAK-deficient mice, which cannot undergo intrinsic

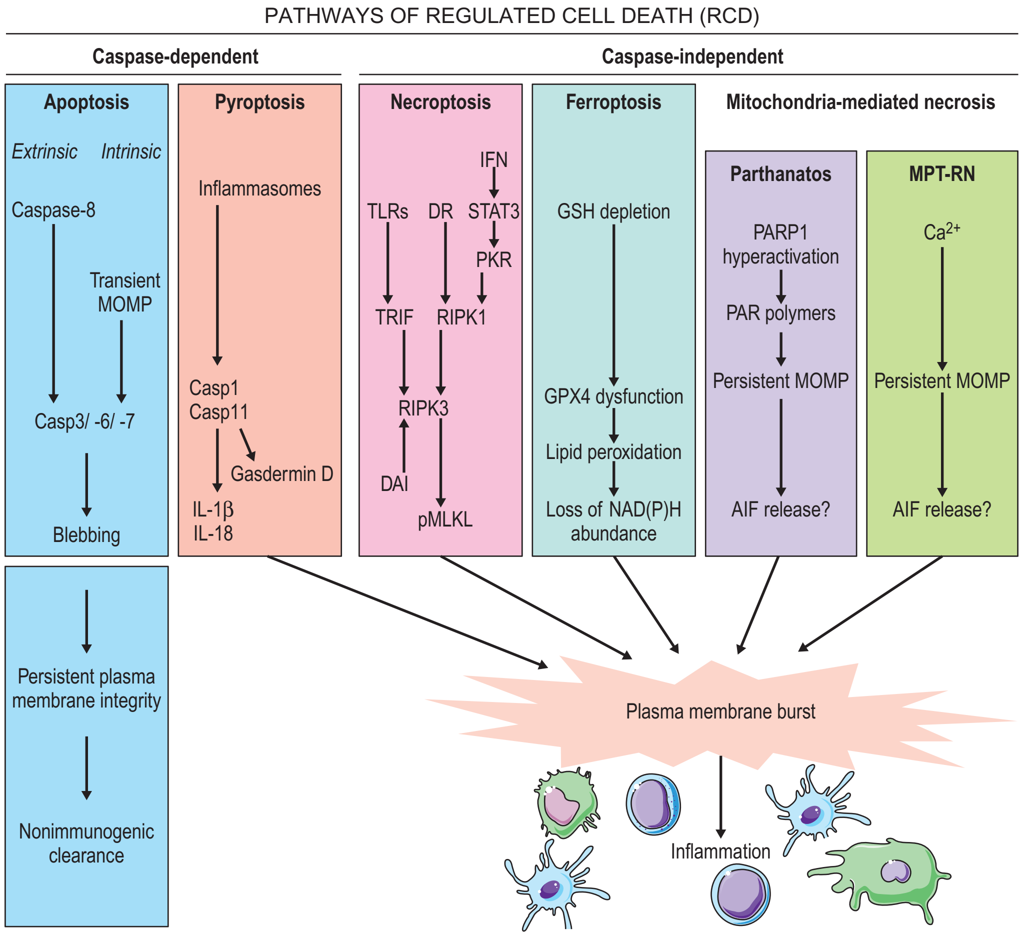

PATHWAYS OF REGULATED CELL DEATH (RCD)
Caspase-dependent
Caspase-independent
Apoptosis
Extrinsic
Intrinsic
Caspase-8
Transient MOMP
Casp3/ -6/ -7
Blebbing
Persistent plasma membrane integrity
Nonimmunogenic clearance
Pyroptosis
Inflammasomes
Casp1
Casp11
Gasdermin D
IL-1β
IL-18
Necroptosis
IFN
TLRs
DR
STAT3
PKR
TRIF
RIPK1
RIPK3
DAI
pMLKL
Ferroptosis
GSH depletion
GPX4 dysfunction
Lipid peroxidation
Loss of NAD(P)H abundance
Mitochondria-mediated necrosis
Parthanatos
PARP1 hyperactivation
PAR polymers
Persistent MOMP
AIF release?
MPT-RN
Ca2+
Persistent MOMP
AIF release?
Plasma membrane burst
Inflammation

FIG 13.5 **Pathways of Regulated Cell Death (RCD)—An Overview.** In contrast to nonimmunogenic, caspase-mediated classical apoptosis, which can be mediated by either the intrinsic (mitochondria-mediated) or the extrinsic (death receptor–mediated) pathway of apoptosis, all other known forms of RCD are characterized by loss of plasma membrane integrity and therefore are categorized as regulated necrosis (RN). Apoptosis and pyroptosis are triggered by proteases (caspases). Necroptosis is regulated by kinases, most importantly the RIPK3-mediated phosphorylation of MLKL. Ferroptosis is predominantly mediated by lipid peroxidation. Finally, mitochondrial permeability transition- regulated necrosis (MPT-RN) and parthanatos represent mitochondrial necrosis. AIF, apoptosis-inducing factor; CASP, caspase; DAI, DNA-dependent activator of interferon regulatory factors; DR, death receptor; GSH, glutathione; GPX4, glutathione peroxidase 4; IFN, interferon; IL, interleukin; MLKL, mixed-lineage kinase domain-like protein; MOMP, mitochondrial outer membrane permeabilization; PARP1, poly(ADP-ribose)-polymerase 1; PKR, protein kinase R; RIPK, receptor interacting protein kinase; STAT3, signal transducer and activator of transcription 3; TLR, Toll-like receptor; TRIF, toll–interleukin-1 receptor [TIR]–domain–containing adapter-inducing interferon-β.

apoptosis, are not viable. Therefore apoptosis appears to be important for normal development and, particularly in the case of CASP8, for the inhibition of necroptosis.

Other than hereditary autoimmune syndromes (*e.g.*, autoimmune lymphoproliferative syndrome [ALPS]), no clear role for apoptosis in diseases has been unequivocally reported, and inhibition of caspases in certain diseases has either deteriorated or not provided any benefits to disease outcomes.

Pyroptosis

Pyroptosis (from “pyro” = fever/inflammation and “ptosis” = to fall) is a form of regulated cell death induced by inflammatory stimuli, transduced via inflammatory caspases and (probably) executed by gasdermins. Inflammatory caspases include CASP-1, CASP-4/-5 (human), and CASP-11 (murine). Unlike apoptosis, caspase activation in this context leads to a necrotic phenotype.

An extraordinarily potent inflammatory stimulus for pyroptosis is LPS, a typical component of bacterial membranes. Intracellularly, LPS is directly bound by pro-CASP-11, which thereby becomes activated and cleaves gasdermin D. Proteins of the gasdermin family are composed of a self-inhibitory C-terminal and a death-inducing N-terminal fragment. Upon caspase-mediated cleavage of gasdermin D, the N-terminal fragment loses its self-inhibiting C-terminal fragment and therefore becomes active. Other gasdermins have not been shown to be cleaved by caspases but do induce pyroptosis upon (artificial) cleavage as well. Their physiological role remains unknown. The axis LPS–CASP-11–gasdermin D–pyroptosis is also referred to as “noncanonical”.

The canonical pathway is initiated by inflammasomes as NLRP3 (Chapter 3), which forms upon stimuli that include pathogen-associated molecular patterns (PAMPs), DAMPs, and absent in melanoma-2 (AIM2) as a response to cytosolic DNA. These inflammasomes recruit CASP-1 via adapter molecule ASC (apoptosis-associated speck-like protein containing a C-terminal caspase recruitment domain) and lead to gasdermin D cleavage and IL-1β/IL-18 maturation.

As both the noncanonical and canonical pathways lead to necrotic cell death, the question about its mechanism arose. It is clear that proinflammatory cytokines pro-IL-1β and pro-IL-18 are actively matured by caspases, resulting in the release of IL-1β and IL-18 from the cytosol into the extramembrane space upon necrosis execution. It is a matter of debate if, in parallel, another mechanism of release of IL-1β and IL-18 exists and functions while the plasma membrane remains intact. In this context, maturation is independent of gasdermins, but release is not. So how do N-terminal gasdermin fragments kill cells and release cytokines? There is evidence that these fragments form oligomers, which can bind to selected negatively charged lipids, such as phosphoinositide and cardiolipin. After docking to membranes, pores are formed. If these are directly formed by gasdermins or just mediated by them (or even artificial, see also "Necroptosis") remains a matter of debate. Phosphoinositides are restricted to the inner leaflet of mammalian membranes, whereas cardiolipin is restricted to mitochondrial and bacterial membranes. This might protect bystander cells and reduce numbers of intracellular bacteria before necrotic release.

In pyroptosis, proinflammatory factors are released. Indeed, active production of systemically effective inflammatory cytokines during RN has been exclusively described for pyroptosis, thus rendering this cell death modality the most inflammatory. A working model for the immunogenicity of various RCD pathways is presented in Fig. 13.4. Pyroptosis, as all other pathways or RN, therefore exhibits an example of ICD.

As expected of a cell death subroutine that causes systemic inflammation downstream of an inflammasomal trigger, pyroptosis is typically observed in macrophages during gram-negative infection and upon culture with bacterial or viral intracellular pathogens. Pyroptosis is thus thought to mediate the immunogenic destruction of colonized niches.

CASPASE-INDEPENDENT REGULATED NECROSIS

Necroptosis

Necroptosis is the best characterized mode of RN. It was discovered as a type of necrotic cell death in apoptosis-resistant cell lines. Classically, it is induced upon tumor necrosis factor-α (TNF-α) stimulation whilst apoptosis is blocked (*e.g.,* by a pan-caspase inhibitor, such as zVAD). In this context, TNFR1, Fas, and other death receptors (described at pathways of extrinsic apoptosis, see above) transduce this signal into the cell (Fig. 13.6).

RIPK1, a key checkpoint of cellular fate, contains a unique motif next to its DD referred to as RIP homotypic interacting motif (RHIM). This motif is contained in only four proteins within the whole mammalian proteome, and all these proteins are associated with the regulation of necroptosis. The necrosome, a higher order structure with a poly-RIPK3 backbone, is central to necroptosis execution and can, next to named death receptor–associated pathway, also be engaged by TRIF-binding TLR3 and TLR4, as well as by viral sensing protein DNA-dependent activator of interferon regulatory factors (DAIs).[12] Recently, DAI (also known as ZBP1) has been demonstrated to mediate *in utero* lethality of RIPK1-deficient mice by forcing RIPK3 oligomerization, which places inactive RIPK1 as an inhibitor of necroptosis. The RIPK1 kinase inhibitor necrostatin-1 (Nec-1) seems to stabilize this conformation and thereby inhibit necroptosis on death receptor activation by not releasing RIPK1s' inhibition to RIPK3 oligomerization.

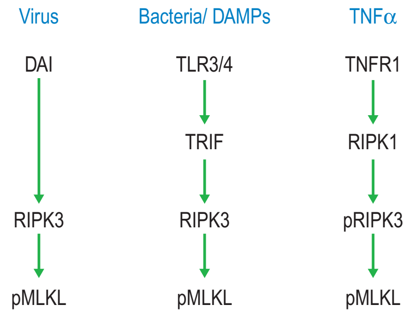


FIG 13.6 Necroptosis Induction. Different triggers are known to induce necroptosis, and they are sensed by specific adaptors. Viruses are sensed by DNA-dependent activator of interferon regulatory factors (DAIs), which subsequently triggers RIPK3 polymerization by replacing RIPK1 (receptor-interacting protein kinase 1). Bacterial pathogen-associated molecular patterns (PAMPs) and other DAMPs are sensed by Toll-like receptors (TLRs) 3 and 4, which trigger RIPK3 polymerization by RHIM-domain containing protein Toll–interleukin-1 receptor [TIR]–domain–containing adapter-inducing interferon-β (TRIF). The best-described pathway is activated upon tumor necrosis factor (TNF)-α stimulation of TNF receptor 1, which leads to RIPK1 phosphorylation. This phosphorylation licenses RIPK3 polymerization as well. As result of the active necrosome based on the poly-RIPK3-backbone, mixed linage kinase domain-like protein (MLKL) is phosphorylated and triggers membrane permeabilization and CXCL1/IL-33 transcription in the nucleus.

Downstream of the active necrosome, mixed lineage kinase-like domain protein (MLKL) becomes phosphorylated by RIPK3 (see Fig. 13.5). After phosphorylation of MLKL, this pseudokinase forms oligomers targeting the plasma membrane via its four-helical bundle (4HB) motif to mediate plasma membrane rupture and necrotic cell death by currently undefined means. Phosphorylated MLKL (pMLKL) is required, but is not sufficient to execute necroptosis. Whether pMLKL directly forms pores or mediates pore formation is unknown (as well as for gasdermins in pyroptosis). However, this is of great interest as pMLKL targets multiple intracellular membranes, translocates to the nucleus to induce CXCL1/IL-33, and is stably expressed in terminally differentiated cells, such as podocytes and endothelia, without killing these. Taken together, this might point to a still-unknown physiological role of pMLKL.

Mice deficient in RIPK3 or FADD and MLKL die following challenge with influenza A virus, which has been demonstrated to depend on DAIs' RHIM domain. Viruses also indirectly activate necroptosis by JAK-STAT–dependent protein kinase R upregulation.

Bacterial infection is sensed via TLR3 and TLR4. This also recruits RHIM domain–containing protein TRIF, which leads to necrosome formation. Current understanding favors the idea that necroptosis is an evolutionary conserved program to defend against viruses and certain bacteria. In line with this, some viruses express caspase-inhibitors, such as crmA (*e.g.*, cowpox virus), whereas viral protein M45 (*e.g.*, CMV) specifically targets necroptosis. M45 has been demonstrated to be a viral RHIM domain and thus suppresses DAI-induced RIPK3 oligomerization within the necrosome. Interestingly, cytomegalovirus (CMV) is a member of the herpesvirus family, which is famous for its persistence within the host. One might conclude that this virus adapted to the necroptotic trapdoor by this mechanism.

During necroptosis, some chemokines (Chapter 10) and cytokines (Chapter 9) are actively produced to be released apart from DAMPs. These include CXCL1 and IL-33, a stimulator of ST2 signaling on Tregs. This suggests that necroptosis, apart from being immunogenic through the release of DAMPs, also limits the inflammatory response to a certain surrounding of the damage by creating a microenvironment that may prevent necroptosis from causing a systemic inflammatory response syndrome (SIRS) and death. Necroptosis may thus be the least immunogenic RN pathway (see Fig. 13.4).

CLINICAL RELEVANCE

Selection of Clinically Relevant Conditions Associated With Necroptosis

- Acute liver failure
- Autoimmune disorders
- Acute respiratory distress syndrome (ARDS)
- Cancer (necrosis in the center of solid tumors)
- Chemotherapy
- Contrast-induced acute kidney injury (CIAKI)
- Myocardial infarction
- Sepsis
- Solid-organ transplantation
- Stroke
- Transplant rejection

Necroptosis critically contributes to diverse pathophysiological settings, such as ischemia–reperfusion injury in solid organ transplantations, myocardial infarction, stroke, and SIRS. RIPK3- and MLKL-deficient mice have been demonstrated to be protected from preclinical models of such diseases by several independent groups, so inhibitors of necroptosis (RIPK1 kinase inhibitors, RIPK3 kinase inhibitors, and MLKL inhibitors) have entered phase I and phase II clinical trials. No cell death–preventing therapy has been approved by the US Food and Drug Administration (FDA) as of the writing of this chapter. However, preclinical and first clinical data are very promising. Necroptosis inhibitors may soon become the first-in-class compounds to prevent RN.[13]

Ferroptosis

Ferroptosis is an important RN pathway in ischemic injury and cancer. Unlike extrinsic apoptosis and necroptosis, it is not initiated by specific receptors. In renal tubules, it mediates an event referred to as *synchronized regulated necrosis of an entire functional unit.* It thus provides a biochemical basis for the clinical observation of necrotic casts in the urine sediment of patients with acute kidney injury. Ferroptosis key molecules have been associated with renal clear cell carcinomas. Ferroptosis is critically mediated by the loss of NADPH (the major cellular redox equivalent) abundance as result of lipid peroxidation. This peroxidation is actively mediated by lipoxygenase ALOX5 and glutathione peroxidase 4 (GPX4) seems to be its major counterpart. Upon dysfunction of GPX4, ferroptosis occurs rapidly in all metabolically active cells by ROS generated in mitochondria and peroxisomes and at the plasma membrane.[2,14] Computed models of sepsis metabolomics also predict a benefit in improving intracellular cysteine levels, but these results still await experimental verification.

GPX4 requires glutathione (GSH) to function. Inhibition of the Glu/Cys-antiporter system X_c^- in the plasma membrane depletes intracellular cysteine required for GSH synthase. Inhibitors of antiporter system X_c^- are inducers of ferroptosis, such as the compound "erastin," which was found in a screen for lethal compounds against Ras-transformed tumor cells. Inhibitors of this process have also derived from screens, such as the first-in-class compound ferrostatin-1.[15] In fact, it is no wonder that ferroptosis inhibitors have been identified in cancer cell lines, as evidence accumulates that many cancers are highly sensitive to ferroptosis induction, especially lymphomas and clear cell renal cancer.[16] Interestingly, p53, which is among the most mutated or deleted proteins in diverse cancers, is also an inhibitor of system X_c^-. Thus p53 may mediate some of its antitumor properties by inducing ferroptosis.

As GPX4 is responsible for efficient repair of oxidized sphingolipids, GPX4 inhibition by GSH depletion, direct allosteric inhibitors (*e.g.*, RSL3), or genetic depletion leads to accumulation of peroxidized lipids. Defects in enzymes required for generation of polyunsaturated fatty acids, such as acyl-CoA synthetase long-chain family member 4 (ACSL4) or lysophosphatidylcholine acyltransferase 3 (LPCAT3), sensitize cells to undergo ferroptosis as well, further emphasizing the role of lipid peroxidation. Interestingly, although known as a redox equivalent, NADPH depletion has been identified as a downstream event of lipid peroxidation. However, this might explain the spread of RN from one cell to another in a functional unit, which is a unique feature of ferroptosis. NADPH drifts freely between neighboring cells through intercellular pores, such as gap junctions, that may give rise to a diffusion drag, leading to RCD in adjacent cells.[17,18]

Ferroptosis and synchronized RN are both active in the urogenital tract and synchronized organs, such as the brain and the heart. This has led to the hypothesis that the physiological role of ferroptosis may be preserved for obliteration processes similar to those of the paramesonephric (Müllerian) duct.

At the time of writing this chapter, no immunomodulatory role for cells that die by ferroptosis had been described, and the immunogenicity of ferroptosis is certainly high. In contrast to the more immunogenic pyroptosis and the less immunogenic necroptosis, ferroptosis may thus be a prototype cell death that leads to inflammation induced by DAMPs in the classic sense. However, it cannot be excluded that the lipid peroxidation and the high concentrations of ROS that are released into the extracellular space function as proinflammatory stimuli.

Mitochondrial Necrosis

Mitochondrial necrosis is largely a mystery. Clear data exist to demonstrate that widespread depletion of mitochondria does not affect necroptosis[19] and that RIPK3 depletion does not affect parthanatos.[20] Thus necroptosis does not require mitochondria, and mitochondrial cell death does not require RIPK3. However, almost any other RN pathway is, to some extent, affected by mitochondrial dynamics. Most models consider the release of the so-called apoptosis-inducing factor (AIF, clearly a misnomer) from mitochondria as a necrosis-inducing outcome. The detailed mechanisms about AIF-induced necrosis are unclear. Evidence from isolated mitochondria, as well as from immunofluorescence and electron microscopy, suggests a central role of swelling of this organelle and MOMP, as in apoptosis. Overlapping data point to the triggers of the loss of mitochondrial membrane integrity, most prominently the pathways of mitochondrial permeability transition (MPT)–induced regulated necrosis (MPT-RN) and parthanatos (see below).

Recently, the idea that "mitochondrial necrosis" is just an upstream event to "metabolic cell death," which certainly would involve ferroptosis and probably could be a trigger for necroptosis as well, started to emerge, but this requires further evaluation. Ferroptosis may be considered "mitochondrial" cell death, but the key regulators GPX4 and system Xc^- are clearly outside of mitochondria, so we do not classify ferroptosis in this section despite the role of mitochondria in this particular cell death subroutine. For this chapter, we separate MPT-RN from parthanatos. Although there are doubts about the correct classification of mitochondrial cell death, evidence from genetically modified animals leaves no doubt about the *in vivo* relevance.

MPT-RN

MPT-RN is the consequence of mitochondrial permeability transition, a highly effective shortcut between the mitochondrial matrix and the cytosol. MPT is mediated through a pore (the MPT pore [MPTP]), the composition of which has been a matter of debate for at least 2 decades. A current widely accepted model assumes a multiprotein complex physically or functionally involving proteins from the mitochondrial matrix, inner and outer mitochondrial membranes, the transmembrane space, and the cytosol. The pore is controlled by a cyclophilin named cyclophilin D (CYPD), a key modulator of MPTP opening and thus MPT-RN. Genetic absence of CYPD protects mice from diverse *in vivo* challenges, including stroke, myocardial infarction, and renal ischemia–reperfusion injury. As with other cyclophilins, the immunosuppressant cyclosporine (CsA) inhibits the opening of the MPTP and therefore prevents MPT-RN. This effect may well account for some of the immunosuppressive function of CsA. Other models, however, favor CYPD to interact with the c subunit of F_1F_0–ATPase complex.[21] Ca^{2+} is a well-known trigger of MPT-RN *in vitro,* but it remains largely unclear how Ca^{2+} triggers persistent MPT, which probably results in AIF release from mitochondria to the cytosol.

Parthanatos

Parthanatos is defined as cell death that occurs following so-called overactivation of the DNA repair enzyme poly [ADP-ribose] polymerase 1 (PARP1). PAR polymers are formed and translocate to the outer mitochondrial membrane by unknown targeting mechanisms. The default induction of PARP1 overactivation include a wide array of stimuli, ranging from DNA damage (*e.g.,* through irradiation) overreactive oxygen species stress to induction by toxins, such as methylnitronitrosoguanidine (MNNG). Comparable to MPT-RN, parthanatos results in AIF release. It has been suggested that AIF requires active PARP1 to transfer ATP-ribose groups from NAD^+ onto its targets. Interestingly, most of our knowledge of PARP1 originates from cancer research. Many tumors have been demonstrated to overactivate PARP1, and PARP inhibitors are in the FDA approval process for diverse cancers as a result of promising outcomes of phase III clinical trials. For *BRCA1/2* mutated ovarian and breast cancer, PARP-inhibitor olaparib has already been introduced in clinical practice.

KEY CONCEPTS

Prototype Inhibitors of Regulated Necrosis

- Inhibitors of necroptosis
  - Necrostatin-1 (Nec-1)
  - Nec-1s (Nec-1 stable)
  - Ponatinib
- Inhibitors of ferroptosis
  - Ferrostatin-1 (Fer-1)
  - 16-86
  - Liproxstatin-1
- Inhibitors of mitochondrial permeability transition (MPT)–induced regulated necrosis (MPT-RN)
  - Cyclosporine (CsA)
  - Sanglifehrin A (SfA)
- Inhibitors of parthanatos
  - Olaparib and many others
- Inhibitors of pyroptosis
  - zVAD-fmk (nonspecific caspase inhibitor)
  - Others to be developed

CONCLUDING REMARKS AND IMPLICATIONS FOR SOLID-ORGAN TRANSPLANTATIONS

Understanding the pathways of RN will allow screening for and/or design of specific inhibitors of the enzymes involved. Two major effects are expected from inhibiting RN. First, in clinically relevant conditions in which necrosis is the main determinant, such as stroke, myocardial infarction, sepsis, transplantation, acute liver failure, pancreatitis, and the center of solid tumors, it remains to be investigated how beneficial an antinecrosis therapy might be. Second, and probably of at least equal importance, necroinflammation deteriorates and amplifies the primary organ damage.

In transplantation, standard immunosuppression is tailored to preventing proliferation of immune cells, but it does not prevent the priming of memory B cells. To transplant an organ full of DAMPs after a long period of organ transfer may be the strongest possible stimulus for memory B cells, which can only expand upon tapering of the immunosuppression. That is exactly the time when antibody-mediated rejection (ABMR) becomes a major factor, lasting over years.[22] Importantly, this does not happen to a comparable extent in living donor transplantations despite full HLA mismatch and blood group incompatibilities. The upcoming years of RN and transplantation research should clarify how an anti-RN therapy may improve the outcome of transplants with a particular focus on ABMR (Table 13.1).

ON THE HORIZON

In Vivo *Interference With Regulated Necrosis*

- Receptor-interacting protein kinase 1 (RIPK1) inhibitors show promising results in animal studies/human phase I studies. Currently, they are the first-in-class compounds to prevent necrosis.
- Ferroptosis inhibitors have the potential to stop the necroinflammatory circle. First trials in humans are urgently awaited.
- As immune-checkpoint therapy becomes more and more the therapeutic standard for several cancers, questions will arise if induction of regulated necrosis (RN; to produce damage-associated molecular patterns [DAMPs]) and checkpoint blockade (to release autoimmune breaks) could work in a similar fashion.
- Most importantly, it is possible to interfere with necrosis!

TABLE 13.1 Classification of Damage-Associated Molecular Patterns (DAMPs)

Class	Definition	Example(s)
Ia	Promoting *DC maturation* as sensed by classical pattern recognition receptors (*e.g.*, TLR)	HMGB1 and nucleic acids
Ib	Promoting *DC activation* as recognized by nonclassical pattern recognition receptors (scavenger receptor CD91, purinergic receptor P2X7)	CALR and eATP
II	Second signal to NLRP3 inflammasome formation, giving rise to *monocyte activation*	Uric acid and eATP
III	Assisting *DC maturation* as exposed on stressed cells and sensed by NKG2D on natural killer cells	MICs
IV	Neoepitopes contributing to *DC activation*, as they are recognized by IgM antibodies, which activate the complement cascade	Oxidized phospholipids and actin cytoskeleton
V	Dyshomeostasis-associated molecular patterns sensed by NLRP3 receptor or UPR receptors of stressed cells, thereby contributing to an inflammatory milieu	Accumulation of unfolded proteins in the ER, ion perturbations and redox stress

An attempt to classify cell-death-associated damps.
DC, dendritic cell; HMGB1, high mobility group box-1; CALR, calreticulin; eATP, extracellular ATP; NK, natural killer; NKG2D, natural killer group 2 member D; MIC, MHC class I chain-related proteins; NLRP3, NOD-like receptor family, pyrin domain containing 3; ER, endoplasmic reticulum

ACKNOWLEDGEMENTS

This manuscript was supported by funding of the German Research Foundation, Cluster of Excellence 306 (EXC306) “Inflammation at Interfaces” to A.L. The authors acknowledge helpful discussions with Hans-Joachim Anders, Lorenzo Galluzzi, and Tom Vanden Berghe.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Kerr JF, Wyllie AH, Currie AR. Apoptosis: a basic biological phenomenon with wide-ranging implications in tissue kinetics. Br J Cancer 1972;26:239–57.
2. Linkermann A, Green DR. Necroptosis. N Engl J Med 2014;370:455–65. doi:10.1056/NEJMra1310050.
3. Mulay SR, Linkermann A, Anders HJ. Necroinflammation in Kidney Disease. J Am Soc Nephrol 2016;27:27–39.
4. Zhao H, et al. Necroptosis and parthanatos are involved in remote lung injury after receiving ischemic renal allografts in rats. Kidney Int 2015;87:738–48.
5. Land WG, Agostinis P, Gasser S, et al. DAMP - Induced Allograft and Tumor Rejection: The Circle is Closing. Am J Transplant 2016;16:3322–37.
6. Land WG, Agostinis P, Gasser S, et al. Transplantation and Damage-Associated Molecular Patterns (DAMPs). Am J Transplant 2016;16:3338–61.
7. Rickard JA, et al. RIPK1 regulates RIPK3-MLKL-driven systemic inflammation and emergency hematopoiesis. Cell 2014;157:1175–88. doi:10.1016/j.cell.2014.04.019.
8. Schiering C, et al. The alarmin IL-33 promotes regulatory T-cell function in the intestine. Nature 2014;513:564–8.
9. Martinez J, et al. Noncanonical autophagy inhibits the autoinflammatory, lupus-like response to dying cells. Nature 2016;533:115–19. doi:10.1038/nature17950.
10. Hampe J, et al. A genome-wide association scan of nonsynonymous SNPs identifies a susceptibility variant for Crohn disease in ATG16L1. Nat Genet 2007;39:207–11.
11. Vanden Berghe T, Linkermann A, Jouan-Lanhouet S, et al. Regulated necrosis: the expanding network of non-apoptotic cell death pathways. Nat Rev Mol Cell Biol 2014;15:135–47.
12. Linkermann A, et al. Synchronized renal tubular cell death involves ferroptosis. Proc Natl Acad Sci USA 2014;111:16836–41. doi:10.1073/pnas.1415518111.
13. Degterev A, Linkermann A. Generation of small molecules to interfere with regulated necrosis. Cell Mol Life Sci 2016;73:2251–67. doi:10.1007/s00018-016-2198-x.
14. Friedmann Angeli JP, et al. Inactivation of the ferroptosis regulator Gpx4 triggers acute renal failure in mice. Nat Cell Biol 2014;16:1180–91. doi:10.1038/ncb3064.
15. Dixon SJ, et al. Ferroptosis: an iron-dependent form of nonapoptotic cell death. Cell 2012;149:1060–72.
16. Yang WS, Stockwell BR. Ferroptosis: death by lipid peroxidation. Trends Cell Biol 2016;26:165–76. doi:10.1016/j.tcb.2015.10.014.
17. Shimada K, Hayano M, Pagano NC, et al. Cell-Line Selectivity Improves the Predictive Power of Pharmacogenomic Analyses and Helps Identify NADPH as Biomarker for Ferroptosis Sensitivity. Cell Chem Biol 2016;23:225–35. doi:10.1016/j.chembiol.2015.11.016.
18. Tonnus W, Linkermann A. “Death is my Heir”–Ferroptosis Connects Cancer Pharmacogenomics and Ischemia-Reperfusion Injury. Cell Chem Biol 2016;23:202–3. doi:10.1016/j.chembiol.2016.02.005.
19. Tait SW, et al. Widespread Mitochondrial Depletion via Mitophagy Does Not Compromise Necroptosis. Cell Rep 2013; 878–85.
20. Sosna J, et al. TNF-induced necroptosis and PARP-1-mediated necrosis represent distinct routes to programmed necrotic cell death. Cell Mol Life Sci 2014;71:331–48.
21. Bonora M, et al. Role of the c subunit of the FO ATP synthase in mitochondrial permeability transition. Cell Cycle 2013;12:674–83.
22. Sellares J, et al. Understanding the causes of kidney transplant failure: the dominant role of antibody-mediated rejection and nonadherence. Am J Transplant 2012;12:388–99.

MULTIPLE-CHOICE QUESTIONS

1. Which regulated cell death is the least immunogenic?
   A. Necroptosis
   B. Apoptosis
   C. Pyroptosis
   D. Ferroptosis
   E. Mitochondrial necrosis

2. Inflammasomes induce which pathway of regulated necrosis?
   A. Necroptosis
   B. Apoptosis
   C. Pyroptosis
   D. Ferroptosis
   E. Mitochondrial necrosis

3. Therapeutic interference with regulated necrosis (RN) is on the horizon, and a plethora of potential targets have arisen. Which RN pathway can be effectively inhibited *in vivo* by application of receptor interacting protein kinase 1 (RIPK1) inhibitors, the so-called “necrostatins”?
   A. Necroptosis
   B. Apoptosis
   C. Pyroptosis
   D. Ferroptosis
   E. Mitochondrial necrosis

The Microbiota in Immunity and Inflammation

Craig L. Maynard

Humans and other species of mammals are hosts to an array of microbial communities collectively referred to as the *microbiota.* The microbiota comprises prokaryotes (bacteria and archaea), viruses (bacteriophages as well as eukaryotic viruses), and eukarya or the meiofauna (mainly fungi and protozoa). As revealed by the Human Microbiome Project, specialized subcommunities colonize barrier surfaces of the digestive, respiratory, and urogenital tracts, and skin. Our immune system is believed to have evolved to ensure peaceful coexistence with these microorganisms that aid in immune homeostasis, pathogen resistance, and digestion in exchange for a nutrient-rich habitat. There is a perpetual cross-talk between the microbiota and the immune system throughout an individual's life. Specific modulation of this microbiota, particularly in infancy, has important long-term health consequences. Addition of specific symbionts to the microbiota can provide tangible health benefits. Reconstitution of a dysbiotic microbiota continues to be utilized or explored as therapy for inflammatory diseases.

OVERVIEW OF OUR NONMAMMALIAN "SELF"

Exposure to the microbiota and its products occurs via a gradual, ordered process. It begins *in utero* and accelerates during and after birth. The overall composition of an individual's microbiota is impacted by such phenomena as the method of birth delivery, diet, treatment with antibiotics, and environmental exposures. All of these modifying factors can have lasting effects on immune health.

Prokaryotes

The bacterial component of our microbiota is, by far, the most widely studied. This has been made possible, in part, by a fairly recent explosion of genomic and bioinformatics capabilities that has enabled taxonomic identification, and even enumeration, of the members of microbial communities without the need for bacterial culture. Two phyla—Bacteroidetes and Firmicutes—account for almost 90% of intestinal bacteria. The remainder comprises organisms belonging to the phyla Cyanobacteria, Proteobacteria, Actinobacteria, Fusobacteria, and Verrucomicrobia.[1] The bacteria that colonize humans belong mainly to three major phyla: Bacteroidete, Firmicutes, and Proteobacteria. It is estimated that the average human being plays host to approximately 100 trillion bacteria. Most of these cells inhabit the lumen of the gastrointestinal (GI) tract. The density of bacteria increases from ≈10^2–10^3 bacterial cells per mL in the stomach to 10^{12} cells/mL in the distal large intestine (Fig. 14.1). The sheer number of bacteria sets the stage for a very complex relationship between the mammalian immune system and the microscopic inhabitants of the various barrier surfaces.

Viruses

The viruses that inhabit mammalian hosts can be subdivided into bacteriophages, which infect prokaryotic cells; eukaryotic viruses, which infect host and other eukaryotic cells; and virus-derived genetic elements, which can incorporate into host chromosomes and result in the generation of infectious virus at a later date. It has been difficult to quantify the exact size of the *virome.* In the case of bacteriophages, it is generally accepted that they can exist in 10-fold greater numbers than prokaryotes. Bacteriophages can have profound effects on the structure and functions of intestinal prokaryotic communities via viral gene transfer of virulence factors and antibiotic-resistant genes between prokaryotic organisms or through predator–prey relationships. Indeed, the infectivity of some enteric viruses requires the microbiota.[2] Eukaryotic viruses include a vast array of viruses that permanently infect the host and can exist for decades in asymptomatic individuals. These viruses can persist locally or systemically. They can directly impact tissue-specific immunity, including in the GI tract.

Fungi

Collectively referred to as the *mycobiota,* fungal communities represent a considerably smaller proportion of the total microbes in the human body. Commensal fungi can be detected in the mouth, lungs, intestines, vagina, and skin (Fig. 14.2). Advances in our knowledge and understanding of the size and functions of the mycobiota have been hampered by relatively limited genomics and bioinformatics capabilities when compared with the study of bacteria or viruses. However, studies utilizing broad-spectrum antifungal agents have begun to highlight the possible roles played by fungi in protecting against disease processes, such as inflammatory bowel disease (IBD) (Chapter 75).

The microbes that inhabit the human body are often acceptably referred to as the *commensal microbiota,* which in the strictest sense refers to organisms that derive benefit from their host without negatively or positively affecting the host. Although this is true of some members of the microbiota, our relationship with other microbes is one of *mutualism,* whereby each organism performs unique and necessary functions that benefit the other. For example, certain bacteria find a home in the anaerobic environment of the cecum and proximal colon, where they receive a rich source of nutrients in the form of insoluble carbohydrates, which our own digestive enzymes are unable to process. By the process of anaerobic fermentation, these carbohydrates are broken

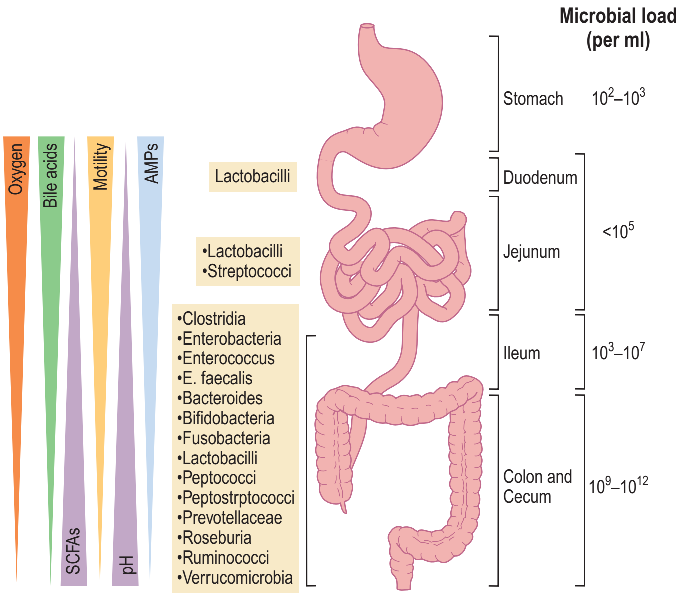


FIG 14.1 **Spatial Organization of Microbial Communities and Physiological Gradients Along the Mammalian Gastrointestinal (GI) Tract.** The numbers and types of bacterial communities, as well as physiological factors, vary along the length of the GI tract. It is well-appreciated that the oxygen levels, bile acid concentrations, intestinal motility, antimicrobial peptides (AMPs), and luminal pH in the proximal portion of the GI tract (stomach, duodenum, jejunum) play major roles in restricting the numbers and types of microorganisms. In general, aerobic and facultative anaerobic bacteria are found almost exclusively in the proximal portion of the GI tract. The hypoxic nature and more physiological pH of the distal small intestine (ileum) and colon, coupled with overall reductions of bile acids, AMPs, and gut motility, allows for unfettered growth of large numbers of obligate anaerobic bacteria. These oxygen-sensitive microbes are capable of producing large quantities of short-chain fatty acids (SCFAs; acetate, propionate, butyrate) from complex carbohydrates (fiber) to be used for important colonic and immunological processes. (From: Reinoso Webb C, Koboziev I, Furr KL, Grisham MB. Protective and pro-inflammatory roles of intestinal bacteria [Figure 2]. Pathophysiology 2016;23:67–80.)

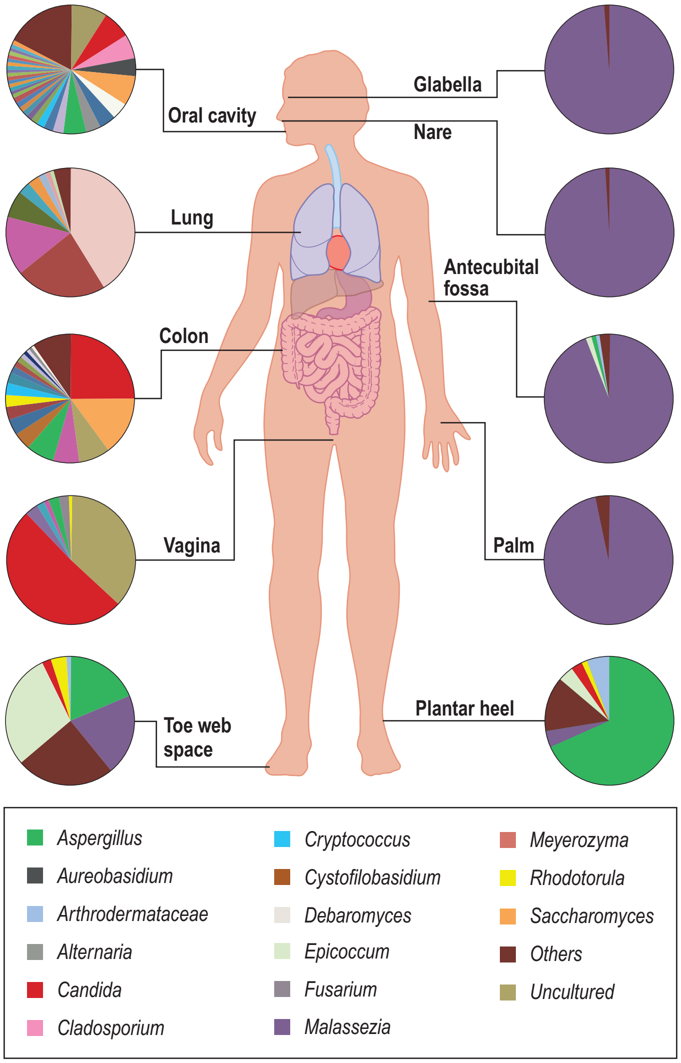


FIG 14.2 **The Human Mycobiota.** Complex populations of fungi have been found associated with the skin and all mucosal surfaces of the healthy human body. The pie charts indicate the relative proportions of fungal genera that are reported to be associated with the respective sites in representative fungal deep-sequencing studies. The fungal populations that are found on mucosal surfaces tend to be more diverse than those on the skin. The healthy lung probably reflects mostly environmental fungi, which are not included in the key. "Others" refers to sequences that represent <1% of the total recovered sequences at each site. "Uncultured" refers to sequences identified in the National Center for Biotechnology Information (NCBI) GenBank database as fungal, but of uncharacterized, origin. Data for pie charts were derived from studies of the fungal genera that are present in the oral cavity, lungs, colon, vagina, and skin. (From: Underhill DM, Iliev ID. The mycobiota: interactions between commensal fungi and the host immune system [Figure 1]. Nat Rev Immunol 2014; 14:405–416.)

down to generate short-chain fatty acids including butyrate, which not only is utilized preferentially by colonocytes but also impacts host immunity and metabolism. Obviously, parasitic or harmful microbes are not considered part of the commensal microbiota, although organisms that fit this definition can often coexist with the microbiota without driving overt disease under homeostatic conditions. More recently, the term *pathobiont* has been coined and is used to refer to any microbe that peacefully colonizes its host but can evoke severe inflammatory responses under specific genetic and/or environmental conditions.

IMMUNE PREPARATIONS FOR MICROBIAL COLONIZATION

Within the last few years, it has become obvious that the developing fetus may be exposed *in utero* to a microbiota or microbial products acquired via the placenta and/or maternal circulation.[3] The acquisition of microbial communities continues with the passage of the fetus through the birth canal and culminates in the first 2–5 years of life in humans. Despite the vast array of microbes that take up residence in the host, colonization is usually an ordered process that, in most cases, is well tolerated. This is because of a robust antimicrobial defense system that is initiated prenatally and fortified postnatally, simultaneous with rapid microbial colonization.

KEY CONCEPTS

Definitions

- **Microbiota**: A collective term for all the microscopic organisms that reside on or in the human body.
- **Microbiome**: The combined genomes of all the organisms that constitute the microbiota.
- **Mycobiota**: That subset of the microbiota that includes fungi alone.
- **Virome**: the collection of all viruses, including viruses integrated into the human genome, found in or on humans.
- **Dysbiosis**: A condition in which there is disequilibrium of the microbial communities that constitute the microbiota at a given body site.
- **Germ-free**: Experimental animals birthed and raised in a sterile environment, devoid of microbes.
- **Gnotobiotic** ("known life"): Describes animals in which the full complement of colonizing microbes is known.

Prenatal Development of the Immune System

In the fetal liver, some common lymphoid progenitor cells—ancestral to all lymphocytes—develop into a specialized subset of innate lymphoid cells (ILCs) referred to as *lymphoid tissue inducer (LTi) cells.*[4] As their name implies, LTi cells are essential for the development of all secondary lymphoid structures throughout the body. These structures will eventually become the sites for initiation of immune responses to the commensal microbiota, pathogenic invaders, and self antigens. In the developing fetus, LTi cells promote the development of mesenteric lymph node (MLN) and of the Peyer patches (PPs) (Chapters 2 and 20) in the distal ileum (Fig. 14.3). They also recruit B and T lymphocytes to these tissues and facilitate their organization into distinct B-cell follicles and T-cell zones, respectively. Throughout life, the MLN provides a so-called mucosal firewall that prevents systemic dissemination of gut bacteria.

Other local mechanisms are also initiated to limit host collateral damage to the neonatal intestine by early microbial encroachment. For example, Toll-like receptor 4 (TLR4), the receptor for lipopolysaccharide (LPS) that is derived from gram-negative bacteria, is highly expressed by intestinal epithelial cells (IECs) prior to birth, but its expression and signaling are rapidly downregulated following onset of colonization. In addition, a diverse array of lymphocytes collectively referred to as intraepithelial lymphocytes (IELs) can be found intercalated between the IECs. IELs display multiple features of activated cells and participate in the maintenance of epithelial barrier integrity by limiting bacterial translocation and promoting epithelial repair following injury.[5]

Passive Acquisition of Antimicrobial Immunity

The microbiota of the neonate is acquired from his or her mother, who has already established tolerogenic relationship with the same microbiota. The mother's "mucosal memory" also gets transmitted to her offspring. During transvaginal delivery, this mucosal memory is seeded with the mother's native microbiota. Mothers produce antibodies to bacteria-derived antigens. These bacteria-responsive antibodies enter the maternal circulation and ultimately get passed to the offspring in breast milk. Immunoglobulin A (IgA) is the major antibody isotype generated by mammary glands (Chapter 15). IgA inhibits bacterial translocation across the neonatal intestinal epithelium, thereby limiting collateral damage by an encroaching microbiota and providing passive immunity to pathogenic infection.[6] The offspring also acquires antibodies of the IgG isotype both *in utero* and via the breast milk. These antibodies serve to limit enteric infection in the newborn[7] and dampen neonatal mucosal T-cell and germinal center (GC) B-cell responses to commensal antigens.[8] Maternally acquired anticommensal antibodies can transfer bound microbial molecules to the offspring during gestation and via the breast milk. This transfer of microbial products contributes to the earliest education of the immune system and limits deleterious postnatal inflammatory responses.[3] Breast milk is also a rich source of immunosuppressive transforming growth factor-β (TGF-β) and interleukin-10 (IL-10) (Chapter 9), which also help promote tolerogenic responses to the microbiota.[9,10]

MICROBIOTA-DEPENDENT MATURATION OF THE INTESTINAL IMMUNE SYSTEM

Microbial colonization prompts a rapid organization of immune structures that are quickly seeded with immune cells. This process helps avoid overexuberant responses to the microbiota and sets the stage for continued tolerance of commensals and for defense against pathogens in the future. The intestinal epithelium limits direct encounter between luminal microbes and the immune cells in the underlying lamina propria not only by forming a physical barrier but also by the production and/or transport of immune and antimicrobial factors (Fig. 14.4). Thus the epithelium is an essential component of the elaborate network of checks and balances that synergize to help to keep pathogens at bay while limiting or preventing collateral damage induced by pathogens and commensals.

Gut-Associated Lymphoid Tissues

Although MLN and Peyer patches begin to develop before birth, complete maturation does not occur until after birth. Germ-free or "germ-reduced" mice display reduced size and cellularity and altered numbers and distribution of immune cells both in the gut and gut-associated lymphoid tissues (GALTs). Thus maturation of the mucosal immune system is contingent on the acquisition of microbiota. A third type of secondary lymphoid tissue—isolated lymphoid follicles (ILFs), which are also induced by LTi cells—is also completely dependent on colonization with the microbiota and thus only develops postnatally.

LTi cells cluster at the base of the crypts in structures referred to as *cryptopatches.* Stimulation of cryptopatches by peptidoglycan derived from gram-negative bacteria induces recruitment of B cells, thereby forming ILFs. The importance of ILFs in the direct control of bacterial growth is demonstrated by the fact that mice devoid of mature ILFs display an overrepresentation of gram-negative bacteria.[11]

Innate Lymphoid Cells

ILCs represent an early line of defense at mucosal surfaces. ILCs are classified as LTi or "LTi-like" (described above) cells or as "helper-like" cells. The progenitor of helper-like ILCs is distinct from that of LTi cells or natural killer (NK) cells. "Helper-like" ILCs are classified into three main subcategories based on their expression of particular sets of surface receptors, transcription factors, and secreted proteins. Unlike their adaptive, thymus-derived counterparts, ILCs do not express antigen receptors and are therefore activated by cytokine signals. This enables ILCs to launch an initial rapid response to microbial challenge and facilitate the development of an adaptive lymphocyte-mediated immune response. ILC1s, such as T-helper type 1 (Th1) cells, express the transcription factor Tbet and secrete interferon-γ

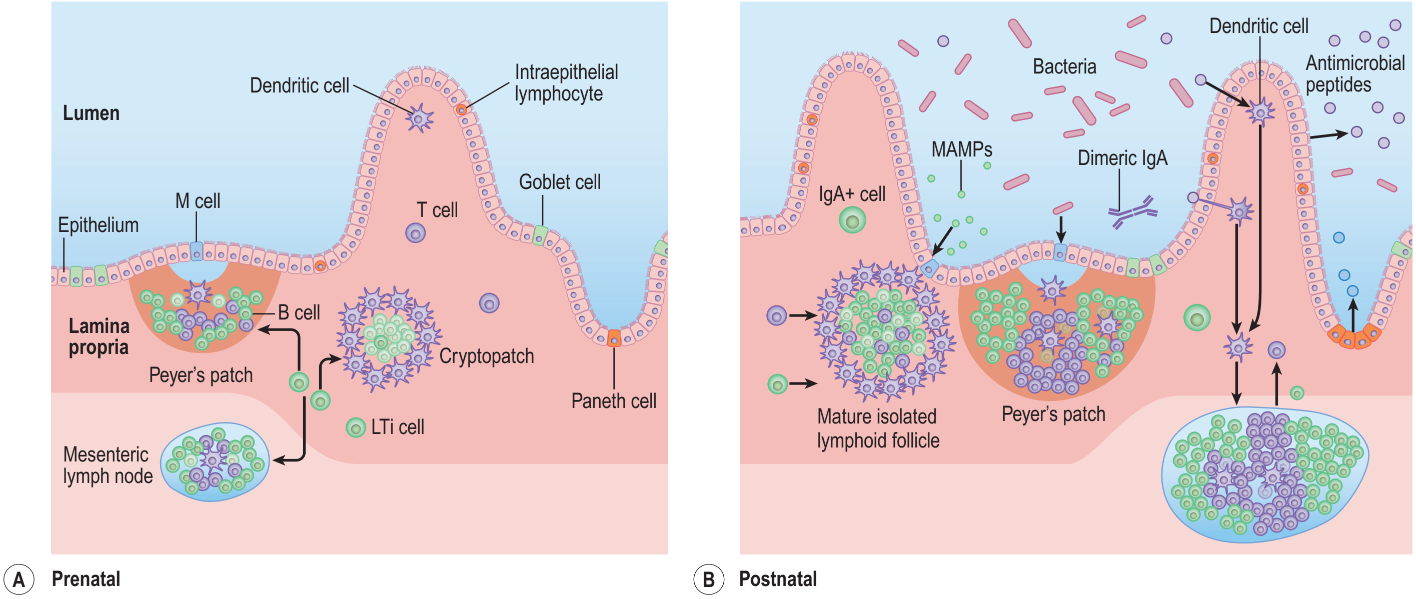


FIG 14.3 The Gut-Associated Lymphoid Tissue (GALT) Establishes Perinatal Host–Microbiota Mutualism in the Intestine. (A) Prenatally, secondary lymphoid tissues (Peyer patches and mesenteric lymph nodes) and cryptopatches develop by the spatiotemporal recruitment of lymphoid tissue inducer (LTi) cells to sites of the developing intestine and supporting neurovascular structures. This, in turn, stimulates the recruitment of dendritic cells (DCs), T cells, and B cells in preparation for the immune response to the microbiota. Intraepithelial lymphocytes (IELs) seed the epithelium before birth. (B) Postnatally, bacteria colonize the neonatal intestine immediately, initiating multiple events that affect the development or functional maturation of the mucosa and GALTs. *Shown from left to right:* Microbe-associated molecular patterns (MAMPs) sensed by pattern-recognition receptors on intestinal epithelial cells and DCs adjacent to cryptopatches stimulate the further recruitment of B cells and T cells, causing the cryptopatches to develop into mature isolated lymphoid follicles. The isolated lymphoid follicles release immunoglobulin A (IgA)–producing plasma cells, which are formed through T cell–dependent and independent interactions, into the lamina propria. Microbes also cross the epithelium and enter the Peyer patch through M cells, from which they are endocytosed by the DCs in the subepithelial dome. Antigen-loaded DCs in the Peyer patch interact with local lymphocytes to induce T-cell differentiation and T cell–dependent B-cell maturation in the germinal center (GC) to induce the development of IgA-producing plasma cells that home to the lamina propria, where they release dimeric IgA for transport into the intestinal lumen. DC-mediated luminal sampling of microbial products or transcytosis of bacteria across the epithelium results in antigen loading of the lamina propria DCs, which then migrate through the afferent lymphatics vessels (not shown) to a draining mesenteric lymph node to induce differentiation of effector T cells that traffic to the lamina propria. Shown on the far right, sensing of microbe-associated molecular patterns (MAMPs) stimulates the proliferation of intestinal epithelial cells in crypts, resulting in their increased depth and, in the small intestine, increased density of Paneth cells. This sensing also arms the intestinal epithelial cells for release of antimicrobial peptides. (From Maynard CL, Elson CO, Hatton RD, Weaver CT. Reciprocal interactions of the intestinal microbiota and immune system [Figure 1]. Nature 2012;489:231–241.)

(IFN-γ) and tumor necrosis factor-α (TNF-α) in response to intracellular bacteria. ILC2s, similar to Th2 cells, express GATA3 and secrete IL-13 and IL-5. ILC3s are arguably the most complex subset of ILCs and can be subdivided into "LTi-like" ILC3s, which populate the intestinal lymphoid follicles, and the more plastic NKp46$^+$ ILC3s, which are distributed throughout the intestinal lamina propria. Both subsets express retinoic acid–related orphan receptor (ROR)γt and produce IL-22 and IL-17, although NKp46$^+$ ILC3s can also express Tbet and secrete IFN-γ.[12]

The impact of the microbiota on ILC3 accumulation appears to commence *in utero*, as transient colonization of pregnant germ-free mice with nonpathogenic *Escherichia coli* increases accumulation of NKp46$^+$ ILC3s and F4/80$^+$CD11c$^+$ mononuclear cells in newborn pups.[3] In the normal intestine, ILC3s are the dominant source of IL-22.[12] The microbiota, acting on innate cells, including macrophages and dendritic cells (DCs), induces expression of IL-23, which, in turn, stimulates the production of IL-22 and/or IL-17 by ILC3. IL-22 can induce the secretion of antimicrobial peptides, such as RegIIIγ, by intestinal epithelial cells to limit bacterial invasion, and within GALT to prevent systemic dissemination of resident microbes. Throughout life, ILC3 expressing Major Histocompatibility Complex (MHC) class II molecules (Chapter 5) on their cell surface help regulate the pool size and function of commensal-specific CD4 T cells to prevent spontaneous inflammation.[13]

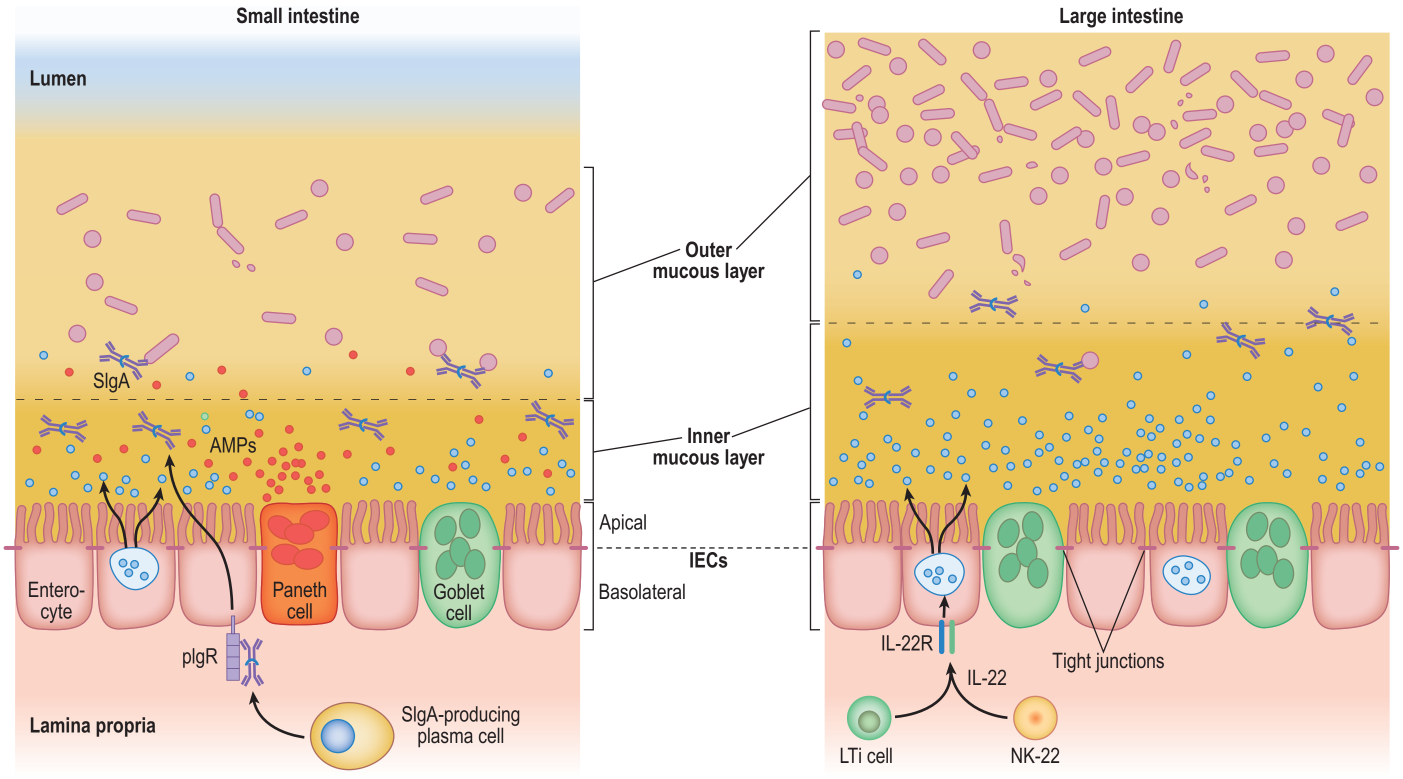


FIG 14.4 Active Barrier Functions of the Intestinal Epithelium. The intestinal epithelium comprises a single layer of polarized columnar epithelial cells (IECs), sealed by tight junctions. Specialized cells known as *goblet cells* secrete mucins that form a bilayered mucus sheath, which maintains separation between the luminal bacteria and the epithelium. The thicker inner mucus layer is sparsely populated with bacteria, whereas the outer layer is more loosely structured and contains ≈10-fold more bacteria. Paneth cells within the epithelium secrete defensins which, along with epithelial-derived antimicrobial peptides, further help with bacterial containment. Epithelial cells also transport antibodies into the lumen and transmit cytokine signals that help maintain barrier integrity.

Invariant Natural Killer T Cells

A unique subset of thymus-derived cells, invariant NK T (iNKT) cells (Chapter 17) express an invariant T-cell receptor (TCR) (Chapter 4) that recognizes lipid antigens presented by the nonclassic MHC class I molecule CD1d (Chapter 5). iNKT cells are essential for the ulcerative colitis–like phenotype that develops in experimental mice treated with the chemical compound oxazolone. Studies on the role of the microbiota in iNKT cell expansion and function provided an elaborate example of the how early postnatal colonization can profoundly impact the maturation and long-term function of the immune system. Germ-free mice harbor increased numbers of iNKT cells in both the colon and the lung, and correspondingly demonstrate increased severity of experimental colitis and asthma. The disease phenotype can be reduced to that of conventionally raised mice if germ-free mice are colonized with commensals in the neonatal period, but not during adulthood.[14]

Regulatory T Cells

Regulatory T cells (Tregs) (Chapter 18) are crucial to the establishment and maintenance of immune homeostasis. Mice or humans with absent or defective expression of the signature Treg transcription factor Foxp3 (FOXP3 in humans) succumb to spontaneous multiorgan autoimmune disease very early in life. Commencing in the first few days of life, FOXP3$^+$ Tregs are exported from the thymus to seed every lymphoid and nonlymphoid tissue. Most Tregs in the intestine express Foxp3, and most of these coexpress IL-10, TGF-β, and IL-35. However, other FOXP3$^-$ subsets also exist including IL-10-producing Tr1 cells, as well as the CD4CD8αα T cells in the human lamina propria.[15]

Commensal-specific Foxp3-expressing Tregs in the intestinal lamina propria are potentially a mixture of thymic-derived Tregs (tTregs) and peripherally induced Tregs (pTregs).[16] In support of the latter, mice deficient in their ability to generate extrathymic Foxp3$^+$ Tregs spontaneously develop intestinal pathology late in life. Colonization of germ-free mice with a benign cocktail of eight commensal microbes, collectively referred to as *altered Schaedler flora,* is sufficient to induce expansion and accumulation of Foxp3$^+$ cells in the large intestine, approaching the levels seen in conventionally raised mice. However, a case has been made for the ability of specific bacteria and/or bacterial components to uniquely promote Treg induction. For example, a cocktail of 46 mouse *Clostridium* strains potently induced expansion of colonic Foxp3$^+$ cells.[17] This cocktail, which has since been reduced to 17 strains derived from a single human donor, and recapitulates the phenotype in germ-free mice.

Several factors produced or induced by gut microbes have been shown to be essential for microbiota-dependent Treg accumulation, particularly in the colonic lamina propria. The capsular polysaccharide A (PSA) moiety of *Bacteroides fragilis,* which mediates the interaction between the bacterium and the colonic mucosa, can also act via the TLR–MyD88 pathway to promote expansion of IL-10–producing colonic Tregs.[18] Clostridial strains stimulate production of TGF-β by intestinal immune cells. This, in turn, enhances Foxp3$^+$ Treg induction. Clostridia are also very adept at anaerobic fermentation of indigestible fiber. Thus their production of butyrate might be another mechanism whereby they can enhance accumulation of Foxp3$^+$ Tregs in the colonic lamina propria.[19]

CD4 T-Helper Cells

In the intestines of both mice and humans, both IFN-γ–producing Th1 and T-helper "IL-17–producing" (Th17) cells are present (Chapter 16). Although detectable even at steady state, these populations, reactive to microbial antigens, are largely held in check by the intestinal immunoregulatory system. In some cases, expression of the Th cell signature (transcription factors and cytokines) of distinct lineages can overlap with each other or with the Treg transcription factor Foxp3, denoting either a common progenitor or the dynamic lineage transitions that can ensue in response to competing immune signals. In the absence of regulatory pathways, for example, in IL-10 deficiency, the numbers and frequencies increase gradually, coincident with the onset of chronic inflammation. The same is true during GI infection (see below), where the ability of invasive bacteria or viruses to enter a cell, or the physical interaction of the bacteria with the intestinal epithelium, culminates in induction and expansion of Th1 and Th17 cells, respectively, in the intestinal lamina propria.

Mucosal B Cells

With increasing colonization comes an increased likelihood of epithelial breach, particularly by bacteria capable of penetrating the inner mucus layer and gain access to the intestinal epithelium. By intercalating dendrites between epithelial cells and into the lumen, DCs in the intestinal lamina propria are able to sample the luminal bacteria. These antigen-loaded DCs migrate to the mesenteric lymph nodes, where they interact with B and T cells to induce B-cell production of anticommensal IgA. These IgA-producing plasma cells migrate to the intestinal lamina propria, where they secrete IgA dimers that migrate across the epithelium and into the lumen. There they bind commensals and thus limit their translocation. In germ-free mice, the absence of the microbiota and resultant absence of IgA production is seemingly replaced by IgE class switching (Chapter 4) in mucosal lymphoid tissues and a corresponding increase in susceptibility to oral antigen-induced systemic anaphylaxis.[20] Therefore the microbiota itself helps limit hyperreactivity to allergens and parasite challenge throughout an individual's life.

IMMUNE CONSEQUENCES OF EARLY MICROBIAL MANIPULATION

Because the microbiota is so essential to the early postnatal maturation of the immune system, the consequences of manipulation or insufficiency of the microbiota can impact host immunity throughout life. Microbial disruption, particularly during the neonatal–infancy period, is being associated with increased risk of autoimmunity and chronic inflammation later in life.

The increasing incidence of autoimmune disease in industrialized countries has been linked to use of antibiotics,[21] improved sanitation, and consumption of processed foods rich in fat and carbohydrates but negligibly low in fiber.[22] These practices limit microbiota diversity and deplete the bacteria that "educate" the developing immune system, leaving it prone to overreaction to subsequent challenges.

There are several observational, clinical, and epidemiological reports, supported by a growing body of experimental findings in support of this concept. For example, vancomycin treatment in young mice increases their susceptibility to asthma as well as food allergy, in large part as a result of depletion of clostridial strains known to promote colonic Treg induction and expansion. Early life treatment with low-dose antibiotics results in increased risk susceptibility to obesity and corresponding alterations in immune gene expression in the ileum. Individuals who had limited sanitary amenities during childhood or were raised around livestock are at reduced risk of developing IBD (Chapter 75) during adulthood. This latter phenomenon is commonly attributed to the acquisition of a diverse microbiota in these microbially enriched environments. However, such living conditions also pose an increased risk of parasitic helminth infection. This concept is supported by experimental data showing that treatment with *Trichuris muris* prevents development of experimental colitis in mice deficient for the IBD-related gene *Nod2*.

IMMUNE SYSTEM—MICROBIOTA CROSS-TALK IN INTESTINAL INFLAMMATION

The immune system and the microbiota are in constant dialogue at steady state. Throughout an individual's life, the microbiota undergoes transient shifts in response to external influences. These include infections; medications, such as antibiotics; and dietary changes. The timing, magnitude, and targets of these perturbations can result in immune responses aimed at resetting this balance or limiting host collateral damage. Furthermore, feces of patients with certain extraintestinal chronic inflammatory diseases also display reduced microbial abundance and diversity relative to their healthy counterparts. This may reflect the role of the intestinal microbiota in the etiology of diseases, the impact of tissue-specific inflammation on the microbiota, or both.

Gastrointestinal Infection

The microbiota helps provide resistance to pathogenic invasion at mucosal sites. This can occur either indirectly by enhancing barrier defenses through various immune and nonimmune mechanisms or directly by competing with harmful microbes. Nevertheless, several organisms still manage to breach these defenses. Such pathogen invasion induces production of proinflammatory cytokines by innate cells that can appropriately expand the magnitude of the immune response by signaling the differentiation of Th cells. The soluble products of these cells recruit other immune cells and collectively aid in pathogen eradication and sometimes in repairing any physical damage to the intestinal barrier. Unimpeded inflammatory responses can be destructive to host tissue, and thus immunosuppressive mechanisms are induced that help to restore immune homeostasis.

In a healthy host, GI infections are largely self-limiting. However the temporary disruption of microbial homeostasis (*dysbiosis*) can have lasting effects on host health. In the cases of *Salmonella enterica* serovar *typhimurium* (*S. typhimurium*) and *E. coli* infection in mice, the associated inflammation can drive production of substances that favor pathogen growth or enable the pathogen to outcompete resident microbes. The inflammatory response to the enteric pathogen *Citrobacter rodentium* promotes a restructuring of the microbiota that can predispose to chronic inflammation. Similarly, *Yersinia pseudotuberculosis* infection results in chronic inflammation and long-term defective lymphatic communication between the gut and mesenteric lymph nodes. These changes are supported by the dysbiotic microbiota.[23]

Infection of mice with the protozoan parasite *Toxoplasma gondii* (*T. gondii*) induces production of IL-12, which, in turn, promotes the differentiation of IFN-γ–secreting Th1 cells. IFN-γ is essential for pathogen control and promotes dysbiosis via targeted destruction of Paneth cells. The production of IL-10 by a subset of the Th1 cells helps overcome the inflammation, whereas IL-10–deficient mice succumb to the infection. Even when the infection is controlled, however, the brief disruption in Treg homeostasis enables systemic dissemination of commensal bacteria and a temporary disruption in tolerance to the microbiota.[24]

Certain microbes also induce specific immune responses simply by their physical interaction with host cells. The ability of certain bacteria to bind directly to the intestinal epithelium is central to their ability to provoke a Th17 cell response that culminates in the accumulation of these cells in the underlying lamina propria. In mice, this is vividly displayed following infection with bacterial species, including the *Clostridium* spp., *Candidatus arthromitus* (segmented filamentous bacteria [SFB]), *Citrobacter rodentium,* and enterohemorrhagic *E. coli* (EHEC) O157:H7, which can breach the mucous barrier and adhere directly to intestinal epithelial cells. In the case of *C. rodentium* and EHEC, this results in a temporary effacing of the epithelial cell layer, local inflammation, diarrhea, and weight loss and in robust induction of Th17 cells. The same phenomenon can be reproduced if germ-free mice are colonized with *E. coli* –adherent bacteria derived from patients with ulcerative colitis.[25] Production of IL-17 as a result of the presence of SFB in the terminal ileum provides colonization resistance to *C. rodentium.* Induction of this immune pathway is stimulated by a microbial breach and helps restore epithelial integrity and limit further invasion. This provides an explanation for the therapeutic blockade of IL-17 being ineffective in the treatment of IBD, and even exacerbating the disease, even though it was successful in reversing symptoms of psoriasis (Chapter 64).[26]

Inflammatory Bowel Disease

IBD is a collective term that refers to a group of chronic relapsing–remitting inflammatory disorders that can occur anywhere along the GI tract. The two main forms of IBD, Crohn disease and ulcerative colitis, have similar clinical presentations but can differ in terms of histopathological features, affected sites, and risk of malignancy. Although disease onset can be impacted by a vast array of genetic, environmental, and immune factors, IBD is characterized by dysregulated immune responses to microbial antigens.

Although no single causative microbe or microbial cluster has been identified, there is strong correlative evidence in support of a dominant role for the microbiota in disease development and function. First, antibiotic therapy continues to be very effective in patients with inflammation in the lower bowel.[27] Second, in treatment-naïve patients with IBD, there is increased abundance of mucosa-associated pathobionts, including Enterobacteriaceae, Pasteurellaceae, Veillonellaceae, and Fusobacteriaceae. Conversely, a decrease in "beneficial" microbes, including Erysipelotrichales, Bacteroidales, and Clostridiales, has been observed. Reduced abundance of *Faecalibacterium prausnitzii,* a member of *Clostridium* cluster IV that induces an antiinflammatory phenotype in immune cells, has also been associated with IBD and risk of relapse in distinct patient cohorts. Whether dysbiosis is a cause or consequence of disease is debatable. Dysbiosis can occur in disease-free relatives of patients with IBD,[28] which is consistent with a genetic influence on microbiota composition. However, several inflammatory mediators very effectively promote dysbiosis in experimental systems, supporting the notion that inflammation precedes microbial disruption.

Genome-wide association studies (GWAS; Chapter 33) of diverse populations of patients with IBD of different races, ethnicities, and geographical locations have all identified over 160 genetic loci that harbor polymorphisms that segregate with clinical disease. Several of the genes identified encode immune-related proteins involved in detecting and/or responding directly to microbial products, inducing or amplifying the immune response to microbial challenge, or restoring and/or maintaining immune homeostasis with intestinal microbes.[29]

The first gene to be causally linked to IBD is *NOD2,* which encodes the nucleotide-binding oligomerization domain–containing protein 2, a microbial sensor that enables the immune system to recognize and respond to intracellular fragments of bacterial peptidoglycan.[30] Loss-of-function (LOF) mutations in *NOD2* (*i.e.,* defective microbial recognition) predisposes to intestinal inflammation, but only in response to infection or injury. This is consistent with the "multiple hit" model of IBD pathogenesis. *NOD2* deficiency is associated with reduced numbers of intestinal goblet cells and therefore reduced mucous production, hyperactive IELs in the small intestine, and increased expansion of the commensal *Bacteroides vulgatus.*[31] Interestingly, all of these abnormalities can be prevented in mice infected with the helminth *Trichuris muris.* Nevertheless, to date, helminth infection as a therapy for human IBD has been largely unsuccessful.

Another immune-related IBD risk allele identified by GWAS is *IL10,* which encodes the immunosuppressive cytokine, IL-10. In addition, rare LOF mutations in *IL10RA,* or *IL10RB,* which encode the IL-10 receptor α and β chains, respectively, are found in a subset of patients with very-early-onset IBD. Mice with global or even CD4- or Foxp3-specific deletion of *IL10* develop spontaneous colitis, which is dependent on the presence of the microbiota. The importance of the microbiota is highlighted by the varying kinetics and severity (ranging from complete protection to severe inflammation) of disease in different animal colonies or even in the same colony at different points in time. Disease is more consistent and severe and can even progress to colorectal cancer in mice colonized with *Helicobacter* spp. Thus IL-10, which can be produced by multiple hematopoietic and nonhematopoietic cells, is critical for maintaining tolerance to the microbiota, particularly in the presence of species that can progressively disrupt the homeostasis of the microbiota.

Experimental studies have also suggested a role for the nonprokaryotic inhabitants of the intestines in maintaining the status quo, thus limiting susceptibility to IBD. Fungi interact with the immune system via the receptor dectin-1. A single polymorphism in *CLEC7A,* which encodes dectin-1, has been linked to a severe form of ulcerative colitis. Accordingly, dectin-1 deficiency in mice precipitates increased susceptibility to chemical-induced colitis, and long-term antifungal treatment results in increased severity of acute and chronic experimental colitis.[32]

Members of the virome can either promote or prevent development of IBD-like symptoms via their interactions with intestinal bacteria. Murine norovirus induces inflammation in mice with disruption in the IBD susceptibility gene *Atg16L1*. Antibiotic administration reverses this viral-induced disease.[33] Conversely, antiviral pretreatment of otherwise healthy mice can exacerbate acute experimental colitis by disrupting the production of antiinflammatory IFN-β.[34]

EXTRAINTESTINAL MANIFESTATIONS OF GUT MICROBIOTA–IMMUNE SYSTEM INTERACTION

Obesity develops when energy intake exceeds expenditure and culminates in deposition of excess adipose tissue. The associated chronic complications, collectively and clinically referred to as *metabolic syndrome,* include hyperglycemia, hypertriglyceridemia, dyslipidemia, and hypertension. Studies in mice and humans have revealed that there are alterations in the gut microbiota in certain conditions, such as obesity and associated metabolic diseases brought on by increased gut permeability, immune responsiveness, and aberrant bacterial translocation.

Perhaps the most striking evidence for the direct role of the microbiota in at least sustaining the obese phenotype was seen in mice transplanted with human fecal microbiota from obese or lean cotwins. Mice transplanted with the obese microbiota displayed increased weight gain and adipose tissue relative to those transplanted with the microbiota of the lean cotwins.[35]

Obesity has been associated an increased abundance of Firmicutes relative to Bacteroidetes, and differential responses of the lean and obese microbiota to the caloric content of the diet.[36] The altered composition can also lead to a reduction in microbial gene richness, which, in turn, affects the "inflammatory tone" of the microbiota and likely contributes the chronic low-grade inflammation characteristic of obesity. Lipopolysaccharides (LPSs), also known as endotoxins, derived from the outer cell membrane of gram-negative bacteria, are found at low concentrations in the circulation of healthy individuals but are dramatically increased in obese individuals and even more so in those who develop type 2 diabetes. LPS infiltrates tissues, including the liver and adipose tissues, and the activation of macrophages via TLR4 initiates an innate immune response characterized by secretion of cytokines, including IL-6 and TNF-α. Obesity-related inflammation also contributes to defective insulin signaling or insulin resistance, a major player in the transition from metabolic syndrome to diseases, including type 2 diabetes, hepatic steatosis, and cardiovascular disease.[37]

Although the microbiota (usually fecal) is distinct in certain non-GI diseases, including ankylosing spondylitis (Chapter 57), multiple sclerosis (Chapter 66), and asthma (Chapter 41), whether these changes precede or follow disease development is not clear. One measure might be the effect of microbiota-induced immune mediators on disease pathology in extraintestinal tissues. For example, although gut microbes that can adhere to the epithelial cells induce robust expression of IL-17, their functions in the gut are largely protective. In contrast, microbiota-induced IL-17 can be proinflammatory in extraintestinal tissues. In germ-free mice colonized with SFB and subjected to experimental models of arthritis or multiple sclerosis, severe IL-17–dependent disease develops.[38] Elevated induction of RORγt and IL-17 in the central nervous system is also a hallmark of virus-induced maternal immune activation (MIA), which can lead to autism spectrum disorder (ASD)–like symptoms in a murine model.

CANCER AND THE MICROBIOTA

The tissue-specific immune inflammatory response that is necessary for pathogen eradication or to restore host–microbiota homeostasis can be deleterious to the host if allowed to reach chronicity. In addition to causing permanent tissue damage (scarring), such a chronic inflammatory response predisposes to tumor development and the production of neo-self antigens that are now recognized as foreign to the host. The antitumor immunity that ensues possesses characteristics similar to those of the antipathogen response and ultimately has the same goal—eradication of a foreign body. Likewise, antiinflammatory or regulatory mechanisms are deployed to limit the magnitude and duration of the inflammatory response in an effort to restore immune homeostasis (Fig. 14.5). However, continued tumor growth means that the "foreign" antigens have not been eradicated, and/or new ones are continually being generated. In response

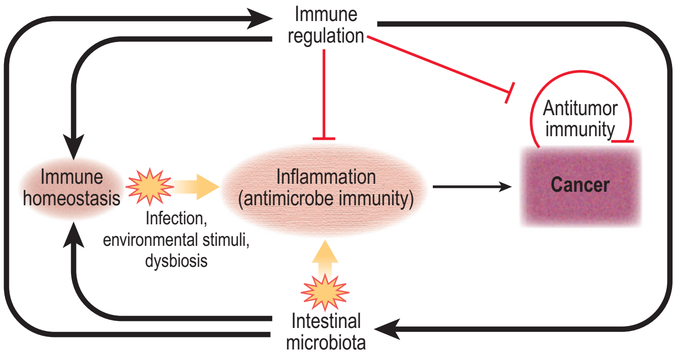


FIG 14.5 **Microbiota–Immunity–Cancer Triad.** The reciprocal interaction of intestinal microbiota and the immune system induces default regulatory pathways that help maintain intestinal immune homeostasis. Factors that promote microbial dysbiosis, such as infection and other environmental insults, can disrupt immune homeostasis. This can result in proinflammatory responses targeting the causative agent, accompanied by immune regulatory responses that aim to reset immune balance. The perpetuation of the inflammatory response enhances the likelihood of progression to cancer. In response to the novel antigens generated in the tumor, an immune response is initiated with the goal of eradicating the tumor. As this response bears similar characteristics to an antipathogen response, it is subject to the same immune suppressive mechanisms.

to cancer, the host immune response needs to persist, rather than being opposed by the host's antiinflammatory arsenal. Thus similar nondiscriminatory immune regulatory mechanisms that are beneficial in curbing inflammation can impede antitumor immunity, thereby enabling tumor growth and eventual dissemination.[39]

There are several immune-related mechanisms whereby microbes directly impact the inflammation–cancer continuum. Some pathogens can promote an inflammatory milieu that encourages tumor development, whereas others can directly transform the eventual tumor-initiating cells. Oncogenic bacteria, including certain strains of *Enterococcus faecalis,* produce carcinogenic reactive oxygen species (ROS) capable of inducing DNA-damaging compounds or can induce production of carcinogenic compounds by activated immune cells. Inflammation-induced cell turnover directly increases the likelihood of introducing mutations in replicating DNA.

Microbiota–Immune System Interactions in Cancer Susceptibility and Development

Not surprisingly, animal models in which deletions of immune genes favor the emergence of a dysbiotic microbiota tend to develop spontaneous nonremitting intestinal inflammation. These include mice deficient in IL-10, Nod1, Nod2, Tbet, or Rag1. In several cases, microbiota reduction using antibiotics, rederivation in an axenic (germ-free) environment, or colonization with the microbiota derived from a wild-type animal is sufficient to significantly inhibit the development of inflammation, as well as the severity of cancer. Thus as a community, the microbiota has the potential to drive both gut inflammation and the eventual progression to cancer.

The gut microbial community has also been implicated in modulating carcinogenesis outside the intestines in experimental systems. Infection with *Helicobacter hepaticus* enhances mammary carcinoma in mice via mechanisms dependent on innate immune activation and TNF production. In addition, TLR5 signaling promotes progression of sarcomas in mice deficient for the tumor suppressor p53, and in which the protooncogene *Kras* has been activated. This disease phenotype can be abrogated by antibiotic-mediated reduction in commensal bacterial load.

In humans, approximately one of every six cancers develop downstream of a pathogenic infection.[40] Notable pathogen–cancer axes include *Helicobacter pylori* and gastric carcinoma, human papillomavirus (HPV) and cervical cancer, and hepatitis B and C viruses and hepatocellular carcinoma.[41] As with intestinal inflammatory diseases, the composition of the fecal and mucosal microbiota in patients with colorectal cancer (CRC) is distinct from that of healthy individuals. The differences between mucosal bacterial populations present "on" and "off" the tumors in the same patient suggest a role for site-specific bacterial community structure in disease development and/or production.[42] More specifically, correlations have been made between CRC and the presence of colonic microbes, including enterotoxigenic *Bacteroides fragilis*, *Fusobacterium nucleatum, Enterococcus faecalis,* and *E. coli.* Animal studies support a role for all these strains in Wnt signaling and myeloid cell activation of nuclear factor-κB (NF-κB)–dependent inflammatory pathways.

The Microbiota in Cancer Immunotherapy

Because of its ability to educate the immune system to be constantly poised to respond to challenge, the microbiota is now also recognized as an important ally in the fight against cancer. Seminal discoveries have been made recently in delineating the pathways whereby the microbiota-primed immune system is essential for the success of common anticancer therapeutic agents.[1] Commensal bacteria support the potency of anti–IL-10R/CpG ODN treatment (a form of immunotherapy) as well as oxaliplatin (a form of platinum salt chemotherapy) in treating colon carcinoma by enhancing the production of myeloid-derived proinflammatory cytokines and ROS, respectively. The efficacy of the alkylating drug cyclophosphamide (CTX) is reduced in germ-free mice or mice treated with vancomycin, which depletes gram-positive bacteria, as a result of diminished antitumor adaptive immune responses. The microbiota is also critical for the antitumor effects of immune checkpoint inhibitors anti–programmed death ligand 1 (PD-L1) and anti–cytotoxic T lymphocyte antigen-4 (CTLA-4).[43] In human, anti–CTLA-4 treatment induces mucosal damage and microbiota modification, partly as a result of partial depletion of gut Tregs. The modified microbiota and the consequent Th1-like immune response are critical for anti-CTLA_4 antitumor functions. In this study, the microbiota of patients treated with anti–CTLA-4 was enriched with bacterial species including *Bacteroides thetaiotaomicron, Bacteroides fragilis,* and *Burkholderia cepacia*. Transplantation of germ-free mice with *B. fragilis* and *B. cepacia* partially rescued the efficacy of anti–CTLA-4 and prevented the mucosal toxicity of the antibody.

THE SKIN MICROBIOTA AND THE IMMUNE SYSTEM

Skin Microbes Maintain Barrier Integrity in the Steady State

With a surface area of approximately 1.8 m^2, skin is the largest organ in the body. Skin functions as a physical barrier against foreign agents (Chapter 19) and also participates in thermoregulation. Unlike the warm, nutrient-rich intestinal tract, skin is cool, dessicates, and is limited in available nutrients for microbial species. Thus skin is populated by microbial communities capable of tolerating its diverse physiology (Fig. 14.6).[44] The total bacterial content of healthy human skin averages about 1 million/cm^2, for upward of 10^{10} total cells covering a single individual (or ≈1% of the number of bacterial cells per milliliter in the distal colon). The skin microbiota is generally acquired in concert with the colonization of other barrier surfaces in infancy. However, coincident with the individual's sexual maturation during adolescence, his or her skin bacterial communities undergo a major shift.[45]

Unlike the gut microbiota, skin commensals are dispensable for the maturation of the immune compartment of the tissue. However, their involvement in resistance to infection is clear. In response to pathogen challenge, skin-resident microbes mount a robust innate immune response characterized by some of the more primitive and evolutionarily conserved immune system messengers—antimicrobial peptides (AMPs) including cathelicidins and β defensins, components of the complement system, and IL-1. Epithelial cells constitutively express some AMPs, which can target a vast array of skin pathogens, including bacteria, fungi, viruses, and parasites. Other AMPs are induced in a microbiota-specific manner and are expressed secondary to activation of the complement system.[45]

In addition to being part of the innate response to microbial encroachment of skin, IL-1 stimulates robust adaptive immune responses that are essential for containment of both pathogens

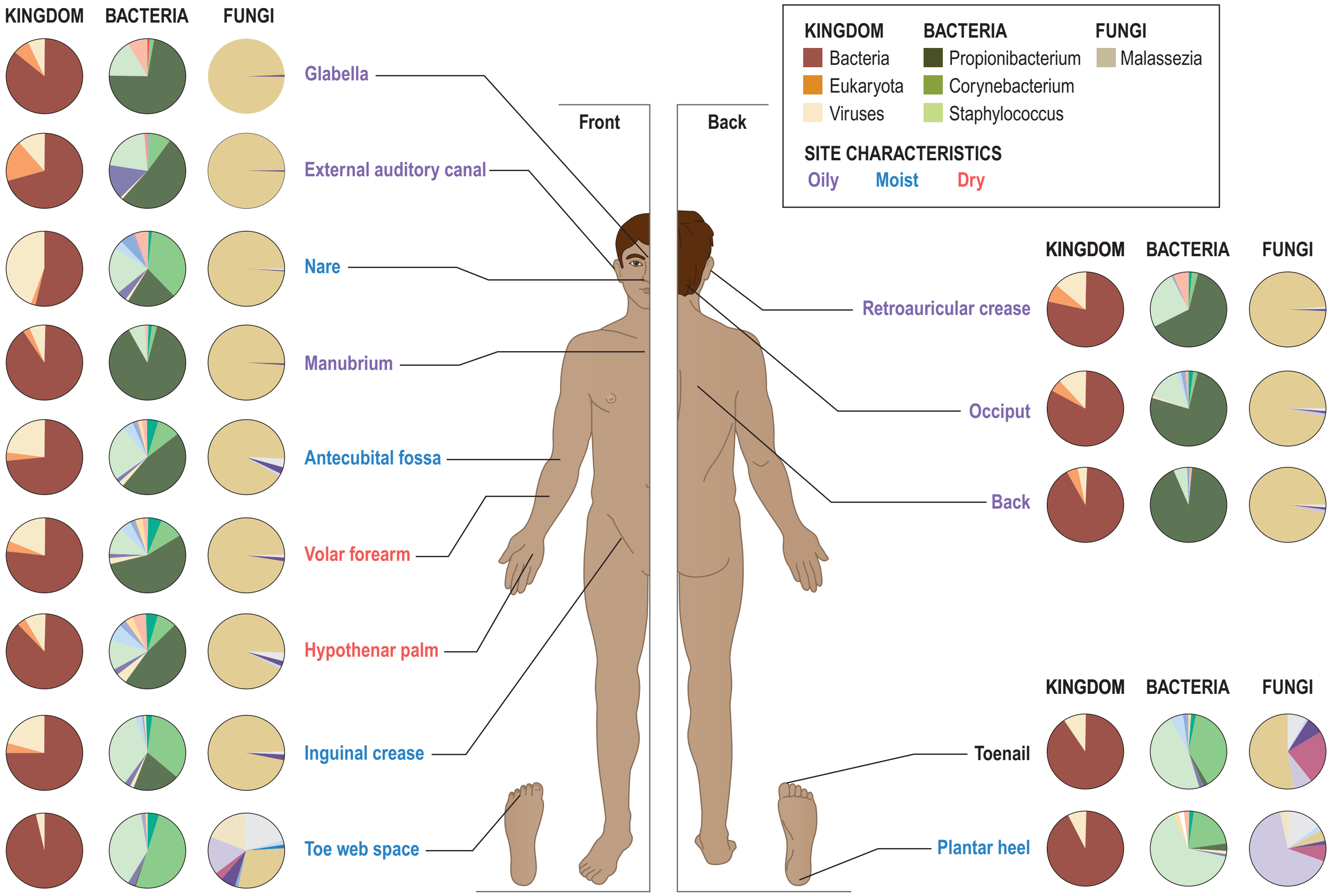


FIG 14.6 **Relative Abundance of Viral, Bacterial, and Fungal Components of the Microbial Community of Skin.** Sites represent three microenvironments: sebaceous (blue), dry (red), and moist (green). The toenail (black) is a site that does not fall under these major microenvironments and is treated separately. Pie charts represent consensus relative abundance of the different categories kingdom, bacteria, and fungi. For bacteria and fungi, major taxa colors are identified in the legend. The minor taxa are colored to represent their relative proportion. (From Belkaid Y, Segre JA. Dialogue between skin microbiota and immunity [Figure 1]. Science 2014;346(6212): 954–959.)

and commensals. In the skin of germ-free mice, there is reduced expression of IL-1, impaired induction of skin Th1, Th17, and IL-17-producing γδ T cells, and an elevated frequency of Foxp3$^+$ Tregs. Consequently these mice also mount a suboptimal response to skin infection. Additionally, in germ-free mice, the reduction in adaptive immunity results in impaired containment of skin commensals and their dissemination to the draining lymph nodes. In contrast, in immune-replete mice, introduction of commensal organisms triggers the induction of CD8 T-cell responses, including the production of IL-17, which is further reinforced by DC-derived IL-1, and helps preserve the integrity of the barrier.

Skin Microbes in Chronic Inflammatory Disease

Skin dysbiosis and resultant dysregulated immune responses have been associated with inflammatory disorders, including acne vulgaris, psoriasis, and atopic dermatitis (AD). Modulation of AMP production is critical, although the specific AMP, and whether it is upregulated or downregulated, can vary from one condition to the next.

In acne vulgaris, sebaceous hyperplasia and the release of lipids into the follicular lumen ultimately clogs the follicle and promotes a self-perpetuating outgrowth of *Propionibacterium acnes*. The follicular wall is breached, triggering an influx of inflammatory neutrophils and pustule formation. The expansion of *P. acnes*, as well as *Staphylococcus epidermidis*, leads to dysregulated immune responses, including elevated expression of AMPs and TLR expression by keratinocytes. These factors sustain the inflammatory response.[46]

Psoriasis (Chapter 64) has been associated with alterations in the relative distributions of two bacterial phyla. Psoriatic lesions in human skin have been found to contain a reduced abundance of and Actinobacteria, including the genus *Propionibacterium*, but an overrepresentation of Firmicutes. Overexpression of antimicrobial peptides, particularly IL-37 produced by stressed cells, is also detectable in diseased tissue.[47] IL-37 primes the innate immune response that subsequently induces development of Th1 and Th17 cells in the draining lymph nodes. The culpability of IL-17 expression in psoriasis has been confirmed by early clinical trials in which antibody-based targeting of IL-17 or the IL-17 receptor resulted in improvements in Psoriasis Area Severity Index (PASI) in at least 80% of patients after 12 weeks.[47]

Atopic dermatitis (AD; Chapter 44) is characterized by dry skin (xerosis), which, along with the associated change in skin pH, favors the colonization and expansion of some microbes more than others. Colonization with *Staphylococcus aureus* has been linked to development of AD, and *S. aureus* is detectable in the skin lesions of over 90% of patients with AD.[48] *S. aureus* infection induces both innate and adaptive immune activation. The expression of IL-37 transcript is markedly reduced in AD lesions, whereas other AMPs, including psoriasin, human β defensin-2, and RNase 7, are overexpressed.

THE RESPIRATORY TRACT MICROBIOTA IN HEALTH AND DISEASE

Inhaled air contains bacteria, viruses, and fungi. Thus the respiratory tract serves as the main entry portal for these airborne microbes. Yet, for over a century, the healthy lung was considered a sterile environment, free of culturable and/or resident reproducing microbes. The detection of microbes in samples collected using instruments that had to traverse the mouth or nasal cavities was often assumed to have resulted from contamination with microbes from these sites. The emergence of culture-independent techniques for detection of microbial communities has precipitated a revision of this concept and an appreciation of the presence and diversity of microbial communities along the respiratory tract even in the absence of overt disease.[49] Thus although the airway and lung microbiota does resemble the bacterial populations of the upper respiratory tract, there are differences that ultimately point to the existence of a specific lung microbiota in healthy individuals.

Although the concept of a lung microbiota during health is relatively new, the role of microbial agents in the pathogenesis of chronic lung diseases has been widely examined. In the healthy lung, the immune system eradicates potential pathogens and overcomes environmental disturbances that threaten to impair lung function. Direct impairment or intrinsic failure of the pulmonary defense mechanisms can lead to infection and/or chronic lung diseases, including asthma, chronic obstructive pulmonary disorder (COPD), cystic fibrosis, and bronchiectasis. Although microbial expansion in the lung is a characteristic of many lung diseases, it is not always clear whether the dysbiotic microbiota is the cause, the consequence, or both.

The Shaping of the Healthy Respiratory Microbiota

The bacterial density of inhaled air is approximately 10^4–10^6 cells/m^3. Thus with each breath, mammals are constantly exposing their lungs to airborne bacteria. There are system-intrinsic mechanisms that serve to regulate entry and colonization of the lungs by microbes. The lung, like skin, and in stark contrast to the GI tract, is a very low nutrient resource. The thin mucus layer, which might seem to represent reduced barrier protection relative to the intestine, also means the absence of a valuable source of nutrients for certain mucophilic microbes. Furthermore, the combination of oxygen tension, temperature, regional pH, the diverse architecture of the respiratory tract, and the proximity of inflammatory cells all help keep the bacterial biomass of the lungs extremely low relative to that of the gut. In a healthy individual, community composition at any given time is largely determined by the relative amount of microbial immigration and elimination. The major routes of immigration are microaspiration, air inhalation, and direct dispersion along the mucosal surface. Elimination is a continuous process mediated by the ciliated epithelial cells, by coughing, and by the actions of the pulmonary immune system. Colonization and growth contribute minimally to microbiota composition during homeostasis but are favored by changes in regional conditions that promote and/or sustain chronic inflammatory diseases of the lung.

Similar to the intestinal microbiota, the microbiota of the upper respiratory tract stabilizes in early childhood and can be influenced by external factors, including breastfeeding and use of antibiotics.[50] In a healthy lung, there is little spatial variation within the same individual. This lack of variation supports the theory that the composition of the lung microbiota is more heavily influenced by immigration and elimination than by local growth. The most abundant bacterial phyla in the lung are Bacteroidetes and Firmicutes. At the genus level, *Prevotella, Veillonella,* and *Streptococcus* predominate. There are also prominent fungal communities, with dramatic differences in composition observed in different healthy cohorts. The oral cavity is home to several genera of fungi, including *Candida, Cladosporium,* and *Aspergillus,* but the fungal colonization of a healthy lung is unclear. Considering the constant exposure of the lung to oral and inhaled fungi, it is likely that the immune apparatus of the lung does, indeed, encounter fungal antigens in the steady state.

The involvement of the bacterial communities of other body sites, particularly the intestines, in lung immune homeostasis is still being actively explored. This phenomenon is commonly referred to as the "gut–lung axis" and involves the action of soluble mediators produced and/or induced by the gut microbiota that can enter the systemic circulation. By this rationale, this phenomenon is more a reflection of the intestinal microbiota on the organism as a whole, and not representative of a unique relationship between these two organs. Nevertheless, there is compelling evidence that modulation of the gut microbiota, particularly during infancy, can have lifelong effects on lung immunity and susceptibility to chronic diseases.

MICROBES AS THERAPY

Certain infections or chronic inflammatory disorders are associated with a severely damaged microbiota. Transplantation of healthy donor microbiota has emerged as a successful therapeutic approach to repair and/or restore microbial communities. Fecal microbiota therapy (FMT), or fecal transplantation, has been safely and effectively utilized as a last resort to treat chronic *Clostridium difficile* infection. FMT presents the risk of adverse effects because the transplanted microbes, which are dormant in the donor, experience propathogenic conditions in the recipient. Because of the successes to date, empirical information regarding the long-term stability and resilience of one individual's microbiota transplanted into another should emerge over time. Other, more focused approaches that employ distinct microbes, groups of microbes, or microbial products known to have immune cell-specific effects to treat inflammatory diseases, such as IBD, are also being considered.

In Western, industrialized societies, there has been a consistent upward trend in the incidence of autoimmune diseases, and this has been linked to improved sanitation and diets low in fiber but rich in unhealthy fats and carbohydrates. These factors have contributed to dramatic shifts in microbial diversity away from that of our ancestral microbiota. Restoration of the richness of the microbiota using *probiotics*—live microorganisms that, when administered in adequate amounts, provide a measurable health benefit to the host—may be helpful in slowing and possibly reversing this trend. Long-term persistence of probiotic organisms within the complex microbiota of the host is essential if these organisms are to induce durable shifts in microbiota dynamics. Complementary *prebiotic* regimens that support the growth and survival of the introduced microbes without detrimental alterations in the distribution of functions of other species will likely be helpful in this regard.

SUMMARY

ON THE HORIZON

- Improved understanding of the structure and function of the microbiota in health and disease:
  - Factors that determine the assembly of the microbiota during early life colonization, and maintenance of community structure throughout life
  - Commensal microbes as a rich source of biological agents that can be used to regulate community development and resistance
  - The mycobiome and viriome—currently understudied mostly because of technical limitations
- Precision medicine and the microbiome:
  - Use of advanced genomic techniques for detailed assessment microbiota functions in health and specific diseases
  - Application of host genome and microbiome data to predict disease susceptibility and responses to therapy
  - Identification and targeted eradication and replacement of deleterious organisms
  - Assembly of limited "designer microbiotas" to initiate therapeutic reconstruction of dysbiotic patient microbiota

The peaceful interaction between humans and the microbes that colonize them is essential for a healthy life. As such, the coexistence of the two entities occurs in an organized fashion, and we now know that this commences *in utero* and is subject to perturbations that can have dire consequences. Perhaps most notable are the findings within the last decade regarding the detrimental and long-lasting effects of broad-spectrum antibiotics used during infancy. The involvement of the microbiota as an instigator and/or sustainer of chronic inflammatory diseases and cancer continues to be explored. In the future, these studies may identify novel avenues for targeting certain conditions and restoring the immune balance. Our growing understanding of the essential roles of the microbiota in diverse physiological processes presents new opportunities to harness these capabilities across the clinical spectrum.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Perez-Chanona E, Trinchieri G. The role of microbiota in cancer therapy. Curr Opin Immunol 2016;39:75–81.
2. Kuss SK, Best GT, Etheredge CA, et al. Intestinal microbiota promote enteric virus replication and systemic pathogenesis. Science 2011;334:249–52.
3. Gomez de Aguero M, Ganal-Vonarburg SC, Fuhrer T, et al. The maternal microbiota drives early postnatal innate immune development. Science 2016;351:1296–302.
4. Spits H, Di Santo JP. The expanding family of innate lymphoid cells: regulators and effectors of immunity and tissue remodeling. Nat Immunol 2011;12:21–7.
5. Ismail AS, Severson KM, Vaishnava S, et al. γδ intraepithelial lymphocytes are essential mediators of host-microbial homeostasis at the intestinal mucosal surface. Proc Natl Acad Sci USA 2011;108:8743–8.
6. Rogier EW, Frantz AL, Bruno ME, et al. Secretory antibodies in breast milk promote long-term intestinal homeostasis by regulating the gut microbiota and host gene expression. Proc Natl Acad Sci USA 2014;111:3074–9.
7. Niewiesk S. Maternal antibodies: clinical significance, mechanism of interference with immune responses, and possible vaccination strategies. Front Immunol 2014;5:446.
8. Koch MA, Reiner GL, Lugo KA, et al. Maternal IgG and IgA Antibodies Dampen Mucosal T Helper Cell Responses in Early Life. Cell 2016;165:827–41.
9. Letterio JJ, Geiser AG, Kulkarni AB, et al. Maternal rescue of transforming growth factor-β 1 null mice. Science 1994;264:1936–8.
10. Garofalo R, Chheda S, Mei F, et al. Interleukin-10 in human milk. Pediatr Res 1995;37:444–9.
11. Bouskra D, Brezillon C, Berard M, et al. Lymphoid tissue genesis induced by commensals through NOD1 regulates intestinal homeostasis. Nature 2008;456:507–10.
12. Klose CS, Artis D. Innate lymphoid cells as regulators of immunity, inflammation and tissue homeostasis. Nat Immunol 2016;17:765–74.
13. Hepworth MR, Fung TC, Masur SH, et al. Immune tolerance. Group 3 innate lymphoid cells mediate intestinal selection of commensal bacteria-specific CD4(+) T cells. Science 2015;348:1031–5.
14. Olszak T, An D, Zeissig S, et al. Microbial exposure during early life has persistent effects on natural killer T cell function. Science 2012;336:489–93.
15. Sarrabayrouse G, Bossard C, Chauvin JM, et al. CD4CD8αα lymphocytes, a novel human regulatory T cell subset induced by colonic bacteria and deficient in patients with inflammatory bowel disease. PLoS Biol 2014;12:e1001833.
16. Cebula A, Seweryn M, Rempala GA, et al. Thymus-derived regulatory T cells contribute to tolerance to commensal microbiota. Nature 2013;497:258–62.
17. Atarashi K, Tanoue T, Shima T, et al. Induction of colonic regulatory T cells by indigenous Clostridium species. Science 2011;331:337–41.
18. Round JL, Lee SM, Li J, et al. The Toll-like receptor 2 pathway establishes colonization by a commensal of the human microbiota. Science 2011;332:974–7.
19. Kim CH, Park J, Kim M. Gut microbiota-derived short-chain Fatty acids, T cells, and inflammation. Immune Netw 2014;14:277–88.
20. Cahenzli J, Koller Y, Wyss M, et al. Intestinal microbial diversity during early-life colonization shapes long-term IgE levels. Cell Host Microbe 2013;14:559–70.
21. Blaser MJ. Antibiotic use and its consequences for the normal microbiome. Science 2016;352:544–5.
22. Deehan EC, Walter J. The fiber gap and the disappearing gut microbiome: Implications for human nutrition. Trends Endocrinol Metab 2016;27:239–42.
23. Fonseca DM, Hand TW, Han SJ, et al. Microbiota-dependent sequelae of acute infection compromise tissue-specific immunity. Cell 2015;163:354–66.
24. Hand TW, Dos Santos LM, Bouladoux N, et al. Acute gastrointestinal infection induces long-lived microbiota-specific T cell responses. Science 2012;337:1553–6.
25. Atarashi K, Tanoue T, Ando M, et al. Th17 Cell induction by adhesion of microbes to intestinal epithelial cells. Cell 2015;163:367–80.
26. Hueber W, Sands B, Vandemeulebroecke M, et al. Inhibition of IL-17A by secukinumab is ineffective for Crohn's disease. J Crohns Colitis 2011;5:S7.

27. Feller M, Huwiler K, Schoepfer A, et al. Long-term antibiotic treatment for Crohn's disease: systematic review and meta-analysis of placebo-controlled trials. Clin Infect Dis 2010;50:473–80.
28. Joossens M, Huys G, Cnockaert M, et al. Dysbiosis of the faecal microbiota in patients with Crohn's disease and their unaffected relatives. Gut 2011;60:631–7.
29. Jostins L, Ripke S, Weersma RK, et al. Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. Nature 2012;491:119–24.
30. Philpott DJ, Sorbara MT, Robertson SJ, et al. NOD proteins: regulators of inflammation in health and disease. Nat Rev Immunol 2014;14:9–23.
31. Ramanan D, Tang MS, Bowcutt R, et al. Bacterial sensor Nod2 prevents inflammation of the small intestine by restricting the expansion of the commensal *Bacteroides vulgatus*. Immunity 2014;41:311–24.
32. Wheeler ML, Limon JJ, Bar AS, et al. Immunological consequences of intestinal fungal dysbiosis. Cell Host Microbe 2016;19:865–73.
33. Cadwell K, Patel KK, Maloney NS, et al. Virus-plus-susceptibility gene interaction determines Crohn's disease gene Atg16L1 phenotypes in intestine. Cell 2010;141:1135–45.
34. Yang JY, Kim MS, Kim E, et al. Enteric viruses ameliorate gut inflammation via Toll-Like Receptor 3 and Toll-Like Receptor 7-mediated interferon-β production. Immunity 2016;44:889–900.
35. Ridaura VK, Faith JJ, Rey FE, et al. Gut microbiota from twins discordant for obesity modulate metabolism in mice. Science 2013;341:1241214.
36. Turnbaugh PJ, Ley RE, Mahowald MA, et al. An obesity-associated gut microbiome with increased capacity for energy harvest. Nature 2006;444:1027–31.
37. Boulange CL, Neves AL, Chilloux J, et al. Impact of the gut microbiota on inflammation, obesity, and metabolic disease. Genome Med 2016;8:42.
38. Lee YK, Menezes JS, Umesaki Y, et al. Proinflammatory T-cell responses to gut microbiota promote experimental autoimmune encephalomyelitis. Proc Natl Acad Sci USA 2011;108(Suppl. 1):4615–22.
39. Goldszmid RS, Dzutsev A, Trinchieri G. Host immune response to infection and cancer: unexpected commonalities. Cell Host Microbe 2014;15:295–305.
40. de Martel C, Ferlay J, Franceschi S, et al. Global burden of cancers attributable to infections in 2008: a review and synthetic analysis. Lancet Oncol 2012;13:607–15.
41. Parkin DM. The global health burden of infection-associated cancers in the year 2002. Int J Cancer 2006;118:3030–44.
42. Dejea C, Wick E, Sears CL. Bacterial oncogenesis in the colon. Future Microbiol 2013;8:445–60.
43. Viaud S, Saccheri F, Mignot G, et al. The intestinal microbiota modulates the anticancer immune effects of cyclophosphamide. Science 2013;342:971–6.
44. Findley K, Oh J, Yang J, et al. Topographic diversity of fungal and bacterial communities in human skin. Nature 2013;498:367–70.
45. Oh J, Conlan S, Polley EC, et al. Shifts in human skin and nares microbiota of healthy children and adults. Genome Med 2012;4:77.
46. Nakamizo S, Egawa G, Honda T, et al. Commensal bacteria and cutaneous immunity. Semin Immunopathol 2015;37:73–80.
47. Morizane S, Gallo RL. Antimicrobial peptides in the pathogenesis of psoriasis. J Dermatol 2012;39:225–30.
48. Cho SH, Strickland I, Boguniewicz M, et al. Fibronectin and fibrinogen contribute to the enhanced binding of *Staphylococcus aureus* to atopic skin. J Allergy Clin Immunol 2001;108:269–74.
49. Bassis CM, Erb-Downward JR, Dickson RP, et al. Analysis of the upper respiratory tract microbiotas as the source of the lung and gastric microbiotas in healthy individuals. MBio 2015;6:e00037.
50. O'Dwyer DN, Dickson RP, Moore BB. The lung microbiome, immunity, and the pathogenesis of chronic lung disease. J Immunol 2016;196:4839–47.

MULTIPLE-CHOICE QUESTIONS

1. The ability of microbes to stably colonize a certain barrier tissue is directly influenced by all of the following *except:*
   A. Environmental factors, such as pH, oxygen tension, and nutrient supply
   B. The musculature of the underlying tissue
   C. The presence or absence of other microbes in that locale
   D. The immune pathways that get activated following colonization

2. Microbe X can be detected in the fecal material of 40% of healthy, immune competent laboratory mice. When pure isolates of microbe X are transplanted into mice with a deletion in gene Y, or into healthy or germ-free mice on a specific diet, the mice develop inflammatory lesions in the intestines and the liver. The most appropriate characterization of microbe X is:
   A. Parasite
   B. Commensal
   C. Pathobiont
   D. Enteric pathogen

3. Puberty has a profound impact on the microbiota of:
   A. Skin
   B. The large intestine
   C. The small intestine
   D. Lungs

4. All gnotobiotic organisms are also germ-free.
   A. True
   B. False

15

Immunoglobulin Function

Neil S. Greenspan, Lisa A. Cavacini

Antibody-mediated immunity generally requires noncovalent contact between an antibody and the antigen (Chapters 4 and 6). The ability of an antigen to bind noncovalently to an antibody, termed *antigenicity*, is a physical–chemical property that is evaluated with respect to a given antibody population. In contrast, *immunogenicity*, the ability to induce the biosynthesis and secretion of soluble antibody molecules, is a biological property. Thus its measurement requires *in vivo* studies. Although antigenicity is necessary for immunogenicity (as defined by the production of antibodies), it is not sufficient. Moreover, the immunogenicity of a given molecule or molecular complex is influenced by host genetic variation. When an antibody binds to a macromolecular antigen, it directly contacts only a portion of the molecular surface of that antigen. Similarly, only a portion of the antibody molecule makes direct contact with the antigen. By convention, the portion of an antibody or T-cell receptor (TCR) that makes physical contact with an antigen is referred to as the *paratope* or *combining site*. Conversely, the region of the antigen in physical contact with the paratope, the antigenic determinant, is termed the *epitope*. Most of the amino acids in an antibody variable domain that contact a given antigen are located in the hypervariable regions (also termed *complementarity determining regions* [CDRs]). However, X-ray crystallographic analyses of antibody–protein antigen complexes have shown that contact residues can reside in the framework regions[1] as well.

Although an epitope (paratope, etc.) is usually defined in terms of intermolecular contact, the region of a molecule involved in physical contact with another molecule may not correspond exactly to the structural correlates for energetics and specificity.[2]

ANTIGEN BINDING AND MOLECULAR IDENTITY

Physical Aspects of Binding

Antibody–antigen interactions are, with rare exceptions, noncovalent. This fact is significant in that these interactions are, in principle, spontaneously reversible under the conditions of temperature, pressure, pH, and ionic strength that generally prevail in living organisms.

Several types of noncovalent bonds have been shown to contribute to antibody–antigen binding. These include van der Waals forces, hydrogen bonds, ionic bonds, and hydrophobic interactions. Individually the strength of these bonds is in the range of one to a few kilocalories per mole (kcal/mole), versus 50–100 kcal/mole for covalent bonds. Since the potential to engage in these types of bonds is shared by many of the components of biological macromolecules, individual weak bonds do not usually confer a high degree of specificity. For example, any two atoms can interact through van der Waals forces. It is only through the simultaneous action of many such bonds that molecular specificity arises. Hence, the importance of a close fit, often referred to as *complementarity,* between the epitope and the paratope.

Complementarity can be maximized by matching the physical–chemical properties of the epitope and the paratope. For example, binding can be facilitated when one molecule is concave and the other is convex, when one molecule is positively charged and the other is negatively charged, or when one molecule is a hydrogen bond donor whereas the other offers a hydrogen bond acceptor. It is expected that the greater the complementarity between receptor and ligand, the stronger is the interaction (greater affinity) between the two molecules. Specificity (see below) is also expected to be influenced by complementarity.[2] In rationalizing the strength of interactions between antibodies and antigens, it is important to remember that the antibody competes with solvent for binding to antigen. The thermodynamics of the interaction between these two structures reflects the influence of the interaction on the solvent and other solutes. Moreover, bound water molecules may make important, even crucial, contributions to an interaction between two biomolecules.

Antibody recognition of antigen serves as a paradigm for understanding molecular recognition in the immune system and in biology in general. As will be discussed below, this fact, coupled with the inducibility of antibodies, permits them to be used as surrogate ligands for almost any receptor (and vice versa).

Affinity is the concept used to convey how strongly two molecules bind to each other. Antibody–antigen interactions can be categorized with respect to the numbers of different kinds of paratope–epitope bonds and the absolute number of such bonds of each kind (Table 15.1). Reflecting the different types of antibody–antigen interaction, two categories of affinity merit consideration: intrinsic affinity and functional affinity. It should be noted that some immunologists use the term “avidity” in place of “functional affinity.”

Intrinsic affinity is a measure of the strength of the monovalent interaction between a particular paratope and a particular epitope under defined conditions of temperature, pressure, ionic strength, and pH (Fig. 15.1). By convention, the intrinsic affinity is taken to be the equilibrium association constant characterizing the paratope–epitope pair. It is the reciprocal of the concentration of monovalent antigen at which half of the paratopes will be occupied. It is not an intrinsic property of either the paratope or the epitope, but rather intrinsic affinity characterizes the relationship between two molecules under defined conditions.

The intrinsic affinity of an antibody for a small molecule, such as a drug (*e.g.*, digoxin) or a hormone (*e.g.*, insulin), can be clinically important both *in vivo* and *in vitro*.[3] For example, the *in vivo* effectiveness of antibody F(ab) fragments in removing toxic levels of the drug, digoxin, from patients being treated for congestive heart failure likely depends on the intrinsic affinity of the F(ab) fragments for the drug. Alternatively, antibody intrinsic affinity can limit the analytical sensitivity of an *in vitro* immunoassay designed to determine the concentration of an analyte, such as a hormone (*e.g.*, insulin, parathyroid hormone) or a drug (*e.g.*, digoxin).

In contrast, functional affinity is defined as the equilibrium association constant characterizing the interaction between an intact antibody and an intact antigen. For a monovalent immunoglobulin G (IgG) antibody–antigen interaction, the intrinsic affinity and the functional affinity will be the same. However, if two paratopes interact simultaneously with two epitopes on the same antigen, referred to as *monogamous bivalency* (Fig. 15.2), the functional affinity of the antibody for the multivalent antigen may be substantially greater (as much as 10 000-fold for IgG) than the intrinsic affinity of that antibody for the relevant epitope on that same antigen.[2]

TABLE 15.1 Antigens and Valence

Number of Types of Epitope	Epitope Copy Number	Examples
Monodeterminant	Monovalent	Hapten: DNP, digoxin
Monodeterminant	Multivalent	Polysaccharide: dextran[a]
Multideterminant[b]	Monovalent	Monomeric protein: myoglobin
Multideterminant	Multivalent	Virion: influenza virus

[a] Even a polysaccharide composed of one type of hexasaccharide can have two or more different kinds of epitope: terminal versus internal residues, for instance. However, a given antiserum may preferentially contain antibodies specific for only one such epitope.

[b] Typically, multideterminant recognition occurs with respect to a polyclonal antibody.

Adapted from Benjamini E, Leskowitz S. Immunology: a short course, 2nd ed. New York: Wiley-Liss; 1991, with permission from Wiley-Liss, Inc., a subsidiary of John Wiley & Sons, Inc.

Functional affinity is also influenced by the degree to which the geometrical relationships among the epitopes are optimal for the paratopes, which will depend on the quaternary structure and segmental flexibility of the antibody molecule. In the presence of nonoptimal geometry, the average number of engaged sites may be less than maximal, and energy may be expended in achieving some epitope–paratope contacts. Therefore the functional affinity for a multivalent interaction does not necessarily increase in direct relationship to the maximal number of binding sites that can be engaged simultaneously by an antibody molecule. For example, the effective valency of pentameric IgM with 10 paratopes is typically half that.

Intrinsic affinity provides information, qualified as above, on the degree of complementarity between the epitope and the paratope. Functional affinity accounts for properties influenced by structural features outside of the epitope and the paratope, as normally conceived. Both concepts of affinity are valuable. Maximization of intrinsic affinity may be of prime importance for antibody-mediated inactivation of toxins or enzymes, which frequently involve monovalent interactions. However, in cases where antibodies bind to repeated epitopes on the surfaces of bacteria, viruses, fungi, parasites, or mammalian cells, the functional affinity may play a much larger role influencing the biological consequences of the interaction.

Bivalent (IgG, IgE) or multivalent (IgA, IgM) antibodies carry with them the potential to bind simultaneously to two or more epitopes on different antigenic particles, cross-linking them rather than engaging in monogamous bivalency or monogamous multivalency (see Fig. 15.2). This phenomenon has played an important historical role in immunology. It is the basis for the clinical method for typing erythrocyte antigens (*e.g.*, ABO and Rh antigens), which still rely on agglutination of red cells by antibodies (or lectins). Such antibodies may also participate in the inactivation of complex antigens such as are found on the surface of human immunodeficiency virus (HIV). Many neutralizing epitopes of HIV are normally not well exposed on native

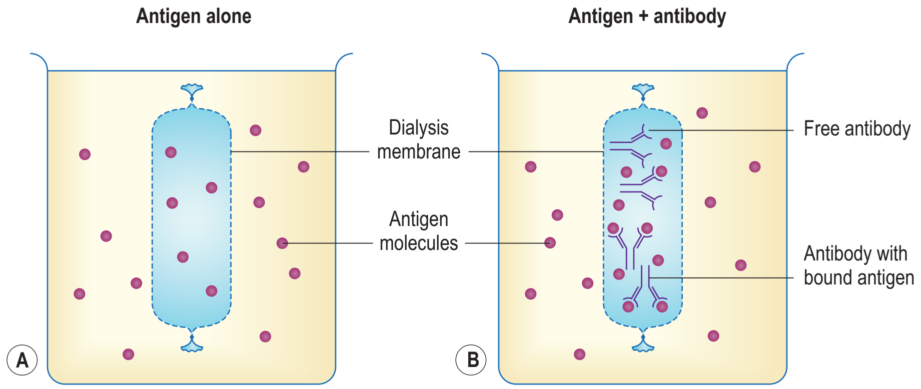


FIG 15.1 Measurement of the Intrinsic Affinity Characterizing an Interaction Between Antibody and Antigen (Hapten) by Equilibrium Dialysis. At equilibrium (A), the amount of diffusible free hapten inside the dialysis bag will be equal to the amount of free hapten outside of the dialysis bag. However, in the presence of hapten-specific antibody (B), the total hapten concentration will be greater inside of the dialysis bag (free + antibody-bound) than outside of the bag (free). The extent of this difference can be used to determine the intrinsic affinity of the antibody for the hapten. With permission from Abbas AK, Lichtman AH, Pober JS. Cellular and molecular immunology. W. B. Saunders Company; 1991.

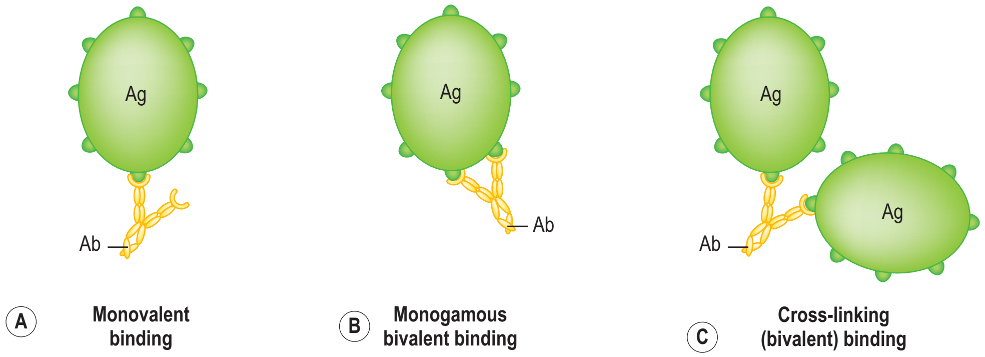


FIG 15.2 **Interaction of a Bivalent Antibody, Such as Immunoglobulin G (IgG), With Multivalent Antigens Can Result in Monogamous Bivalent Binding (B) or Cross-Linking (C).** The complexes in (B) are referred to as cyclic antibody–antigen complexes. With permission from Eisen HN. General immunology. J. B. Lippincott Company; 1990.

virus; however, conformational changes following antibody binding may expose these epitopes for neutralization by other antibodies.

KEY CONCEPTS

Binding of Antibody to Antigen

- Paratopes interact with epitopes through multiple noncovalent (weak) bonds, each of which is reversible at room temperature.
- For the clinician, immunological specificity needs to be viewed in its biological rather than physical context.
- Intrinsic affinity and functional affinity both refer to the strength of antigen–antibody interactions but often play separate roles in the biology of that interaction.
- Intrinsic affinity measures the strength of the interaction between a monovalent epitope and the paratope on the antibody. This can be determined by the equilibrium association constant. Intrinsic affinity is influenced both by the degree of complementarity between the epitope and the paratope and by ambient conditions, including the temperature, pressure, ionic strength, and pH.
- Functional affinity is a measure of the average strength of the interaction between a multivalent antigen and an intact antibody. It is influenced by the spatial relationships characterizing the epitopes that are being recognized as well as the physical properties of both the underlying substrate and the antibody.

IMMUNOLOGICAL SPECIFICITY

The concept of specificity is fundamental to an understanding of the nature and consequences of interactions between immunological receptors and antigens. However, in the immunological context, the term *specificity* encompasses multiple different aspects.[2]

One aspect of specificity focuses on the goodness of fit between the paratope and the epitope. Intrinsic affinity is regarded as a reasonable measure of this goodness of fit. However, substantial conformational adjustments of either the paratope or the epitope may be necessary for formation of the complex.[4] Such conformational changes will generally incur energetic costs. Consequently, intrinsic affinity and final complementarity may not be perfectly correlated.

A second aspect of specificity focuses on the ability of a paratope to distinguish among different epitopes. Such specificity is most readily studied when the epitope is in monovalent form and evaluated relative to a specified set of ligands. Thus one should be cautious about extrapolating claims that one antibody is more or less specific than another antibody without any reference to the relevant universe of ligands. However, there are practical cases where it is justifiable to speak globally of more-or less-specific antibodies. Polyspecific antibodies have been described in the neonatal primary repertoire.[5] These antibodies appear to be globally less discriminating than antibodies typical of the immune repertoire (secondary or later response) when tested on large panels of antigens.

Nevertheless, it is important to note that even antibodies derived from secondary (or later) responses are not,[6] and cannot be, absolutely specific.[2] The impossibility of perfect recognition or discrimination can be understood in both thermodynamic and structural terms. First, perfect fit and absolute discrimination would imply infinite intrinsic affinity (negative free energy change of complex formation), which is not physically plausible.[2] Second, the convexity of atoms prevents perfect shape complementarity between antibody (receptor) and antigen (ligand).[2] Recent results also indicate that at least some antibodies can adopt two or more different unbound conformations, each of which exhibits a different ligand-binding profile. Such paratopes may undergo further structural adjustment in the process of binding to an epitope.[7] This property can be advantageous to the function of an antibody. Antibodies that react with multiple conformations of a viral surface antigen may be much more likely to interfere with viral infection because they can bind more rapidly to the virus than the virus can bind to its receptor, as demonstrated for HIV.[8]

Whereas the first two aspects of specificity focused on the epitope, a third relates to the ability of an antibody to discriminate among antigens that display multiple copies of one or more distinct epitopes. An antigen expressing many copies of one epitope is termed *multivalent,* and an antigen that expresses two or more structurally distinguishable epitopes is referred to as *multideterminant* (see Table 15.1). Because two different cells, bacteria, viruses, and so on may both express multiple copies of the same or nearly the same epitope, an antibody that is highly specific (in the first aspect above) for such a shared epitope may be a poor discriminator between such multivalent particles.[2] Yet, an antibody with a relatively poor degree of complementarity and intrinsic affinity for an epitope found on only one of two or more multivalent targets may be superior at discriminating among these antigens. Furthermore, antibodies (or other molecules)

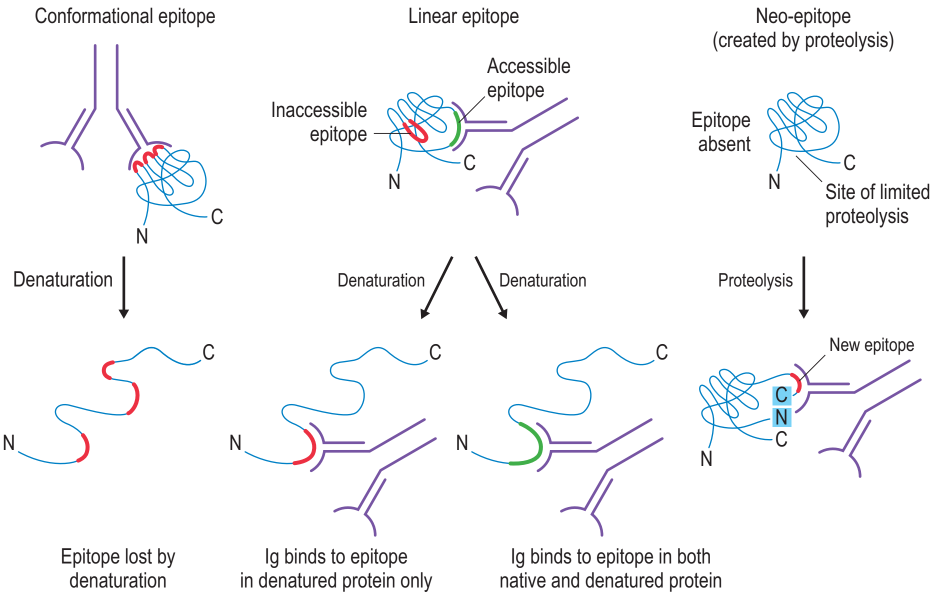

Conformational epitope
Linear epitope
Neo-epitope (created by proteolysis)
Accessible epitope
Inaccessible epitope
Epitope absent
Site of limited proteolysis
N
C
Denaturation
Denaturation
Denaturation
Proteolysis
New epitope
Epitope lost by denaturation
Ig binds to epitope in denatured protein only
Ig binds to epitope in both native and denatured protein

FIG 15.3 **Types of Protein Epitopes.** Some antibodies recognize structural features of proteins that arise from the folding of the polypeptide backbone (conformational epitope). Other antibodies recognize groups of amino acid residues that are contiguous, or nearly so, in the primary (covalent) structure of the protein (linear epitope). If such a linear determinant is inaccessible in the native structure of the protein, the corresponding antibodies may only be elicited by the denatured form of the protein. Neo-epitopes are created by covalent posttranslational modifications, such as proteolytic cleavage. With permission from Abbas AK, Lichtman AH, and Pober JS. Cellular and molecular immunology, 3rd ed. W. B. Saunders Company; 1997.

expressing two or more binding sites of identical structure may not discriminate identically among antigens displaying the same epitope in different two- or three-dimensional distributions.[2]

We offer some final points regarding specificity. First, the interactions between molecules, such as CD4 and major histocompatibility complex (MHC) class II (Chapter 5), which are not clonally distributed, are often described as nonspecific, that is, not specific for an antigen under consideration. The contrast being drawn is with the antigen-specific receptors Ig or TCR. However, from the point of view of the first or second category of specificity described above (goodness of fit, discrimination between ligands), the interaction between molecules, such as CD4 and MHC class II, may be as biochemically specific as that between an antibody and a class II molecule. Ideally, this sense of "nonspecific" should be modified by the preceding term "antigen" (Chapter 6). If not, the meaning must be inferred from the context.

Second, for many purposes immunological specificity has an ultimately biological, not a physical, definition. If the endpoint of analysis is the triggering of a complex response, such as cell activation or initiation of the complement cascade (Chapters 3, 21), then the presence, absence, or the extent of that response, and not the extent of receptor–antigen interaction, will be the ultimate criterion for evaluating specificity. Such biological specificity is not always directly correlated with specificity as determined by the analysis of binding.[2]

Third, the enormous utility of antibodies is crucially dependent on the discriminatory abilities of these molecules with respect to other molecules or molecular aggregates. However, given that the discrimination mediated by antibodies is not absolute, the usefulness of a particular antibody may depend on the context. For example, which antigens or potential antigens in addition to the preferred target are available for binding to the antibody.

Fourth, apparent antibody specificity may vary with the methods used for analysis, as these methods may differ in sensitivity and the conditions (pH, ionic strength, temperature) of application such that the relevant intrinsic affinities may vary among the different assays.

PROTEIN EPITOPES

On the basis of the proximity of the relevant amino acids in the primary structure of the protein, several categories of epitopes have been defined for protein antigens (Fig. 15.3). The simplest is the *linear* epitope, where all of the relevant amino acids are derived from a contiguous stretch of the polypeptide chain. Linear epitopes are particularly characteristic of the short peptide epitopes presented to T cells by MHC molecules on specialized antigen-presenting cells (Chapters 1 and 6). However, antibodies generally recognize complex epitopes on globular proteins that involve amino acids from two or more stretches of polypeptide that are distant from one another in the primary structure. Such an epitope is referred to as a *conformational*, or *discontinuous*, epitope. In some cases, such as with the capsids of nonenveloped viruses, a conformational epitope may contain amino acids that are derived from separate polypeptide chains but that are next to each other in the final capsid structure.

Another category of protein epitope, the *neo-epitope*, is reserved for those antigenic sites that become recognizable only after a posttranslational event, such as proteolytic cleavage. For example, several neo-epitopes have been defined on cleavage products of human C1q, C3, and C9, components of the complement pathway.[9] And, in rheumatoid arthritis (RA) autoantibodies to citrullinated epitopes on such antigens as filaggrin can be of diagnostic value.

Antibodies recognizing neo-epitopes can be used to monitor the extent of activation of the complement pathway.[9] A chimeric antibody, ensituximab, targeting neo-epitopes associated with colorectal and pancreatic carcinomas, represents an evolving area of discovery of antibodies to target cancer neo-antigens.

Studies in the 1970s on the sizes of epitopes associated with synthetic peptide antigens yielded results suggesting that protein epitopes would maximally involve six or seven amino acids. However, the first structure of an antibody-variable module in complex with a globular protein antigen, determined by X-ray crystallography,[1] indicated that protein epitopes, defined on the basis of intermolecular contact, could be as large as 15–20 amino acids. A similar number of amino acids in the antibody V domains constituted the paratope. And, even peptide antigen–antibody interaction can involve at least twelve peptide amino acids in contact with the antibody. Still, it is possible that there are smaller epitopes on globular proteins, particularly for regions of proteins that protrude or have a high radius of curvature.

Antibodies specific for both linear and conformational epitopes have important practical applications. For example, a synthetic peptide corresponding in amino acid sequence to a segment of the polypeptide chain predicted from the nucleotide sequence can be used to elicit antibodies. Antibodies with the potential to recognize a linear epitope available in a denatured form of the gene product can be used to identify the protein following expression, electrophoresis, and blotting under denaturing conditions. Some antibodies raised by challenge with synthetic peptides that bind to linear epitopes can recognize a protein in denatured form but will not bind to or alter the function of the native protein.

Antibodies with the ability to neutralize protein function generally recognize conformations accessible to the native protein, usually at discontinuous epitopes. Thus antibodies specific for peptides (that correspond in amino acid sequence to a portion of a native protein) or denatured protein that can cross-react with the protein in a native (folded, functional) state can be extremely valuable. Such cross-reactivity is more likely to occur when the region being recognized is relatively disordered in the native structure.[10]

CARBOHYDRATE EPITOPES

The classic studies of Kabat on the binding of antibodies to dextran led to the concept that epitopes on carbohydrate antigens could be as large as six or seven monosaccharides. However, minimal carbohydrate epitopes can probably be as small as one or two monosaccharides. Even in the case of larger epitopes, it is typical for the terminal groups to play a dominant role in determining antibody specificity for carbohydrate antigens. Recent studies have suggested that polysaccharide epitopes can sometimes also result from conformational properties of polysaccharides.

Interactions between antibodies and polysaccharides have typically been characterized by relatively low intrinsic affinities in comparison to antibody–protein interactions.[2] Relatively weak antibody–carbohydrate binding can result from biological constraints related to protection against self-recognition and consequent tissue damage or from physical–chemical constraints related to the conformational freedom and high degree of solvation of unbound carbohydrates. However, antibodies produced in response to pathogens, such as HIV, may be much more effective at interacting with carbohydrate antigens.[11]

Another important feature of polysaccharide antigens is that they are generally multivalent. Bacterial and perhaps viral polysaccharide epitope densities can approach values in the millions per square micrometer, which is probably one to several orders of magnitude greater than the epitope densities for protein determinants on mammalian cells. Therefore multipoint attachment and functional affinity are likely to be critical factors in the mediation of immunity by antipolysaccharide antibodies or other carbohydrate-specific proteins.

IMMUNE COMPLEXES *IN VIVO*

KEY CONCEPTS

Immune Complexes

- Immune complexes are aggregates of antibody and antigen.
- Immune complexes can form in tissues, or they can form in the circulation and subsequently deposit in tissues.
- Immune complexes can activate complement or Fc receptor–bearing cells, leading to tissue damage.
- The composition, size, charge, and antibody isotypes characterizing a given population of immune complexes will influence the pathogenic potential of the complexes.

Interactions between antibodies and antigens *in vivo* can result in the formation of molecular aggregates, referred to as *immune complexes.* Deposition of immune complexes in tissues, such as blood vessels, renal glomeruli, renal tubules, the thyroid gland, and the choroid plexus, can result in pathological conditions.[12] Immune complexes can form in the circulation before deposition in a given tissue, or they can form directly in the affected tissue.

A clinical situation associated with immune complex formation is therapy with immunoglobulin (Chapter 84). During administration, infused antibodies may bind endogenous antibodies and the numbers of such complexes are greater at higher infusion rates. These "induced" immune complexes can promote adverse reactions, ranging from mild to severe, including headache, nausea, myalgia, fever, chills, chest discomfort, skin and anaphylactic reactions, and aseptic meningitis.

Some variables, such as concentration, composition, size, charge, and antibody isotype, will influence the magnitude and sites of tissue deposition of immune complexes. The magnitude of complement activation and the extent of interaction with Fc and complement receptors, in conjunction with the sites and extent of tissue deposition, determine the biological properties of the complexes. Antigen–antibody lattice size is determined by antigen valence, epitope geometry, antibody valence, the intrinsic affinity of paratope for epitope, antibody and antigen flexibility, the ratio of antibody to antigen, and the absolute concentrations of antibody and antigen. The potential diversity of immune complex morphologies is illustrated in Fig. 15.4. These complexes, between a monoclonal antibody (mAb) specific for a bacterial polysaccharide and various antiidiotypic or antiisotypic mAbs, are visualized with electron microscopy.

Immune complexes have also been found to have immunoregulatory effects,[13] particularly with respect to antibody responses. Immune complexes can bind simultaneously to B-cell surfaces through antigen (to B-cell surface Ig), antibody (to Fc receptors), and associated complement components (to complement receptors). The interaction with FcγRIIB, on the B-lymphocyte membrane, has the effect of diminishing the B-cell response

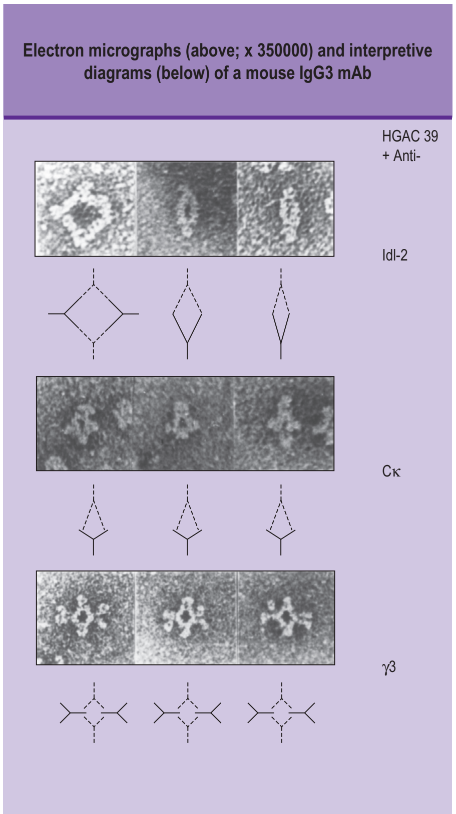

Electron micrographs (above; x 350000) and interpretive diagrams (below) of a mouse IgG3 mAb
HGAC 39
+ Anti-
Idl-2
Cκ
γ3

FIG 15.4 Antigen–Antibody Interactions. Electron micrographs (above; × 350 000) and interpretive diagrams (below) of a mouse immunoglobulin G3 (IgG3) monoclonal antibody (mAb) (HGAC 39, which is specific for a bacterial polysaccharide) in complex with mAbs specific for an idiotope (Idl-2, top), a light chain isotypic determinant (Cκ, middle), and a heavy chain isotypic determinant (γ3; bottom), respectively. The different antibodies are not intrinsically distinguishable in the electron micrographs, but the interpretations take into account information in addition to that provided directly by the electron microscopic images. In the top series of micrographs, the choice of which molecules to represent as solid or dotted figures is arbitrary. With permission from Greenspan NS. Analyzing immunoglobulin functional anatomy with monoclonal antiimmunoglobulin antibodies. BioTechniques 1989;7:1086.

(Chapter 4). The molecular events underlying these immunoregulatory effects are being intensively studied, and they have been exploited clinically for many years. For example, antibody to the erythrocyte Rh antigens is used to prevent immunization of an Rh-negative mother by an Rh-positive fetus, thereby avoiding hemolytic disease of the newborn in a subsequent Rh-positive fetus.

CORRELATIONS BETWEEN C$_H$ REGION STRUCTURE AND ANTIBODY FUNCTION

Antibodies are heterodimeric proteins that can be functionally divided into variable domains, which bind antigen, and constant domains, which define the effector function of the Ig (Chapter 4). This dual function allows an antibody to physically link a specific antigen to a separate antigen-nonspecific effector molecule, such as a component of complement or a cell-bound Fc receptor.

KEY CONCEPTS

Structure–Function Relationships of C$_H$ Domains

- Immunoglobulin (Ig) function is influenced by differences in quaternary structure and segmental flexibility.
- The hinge region helps control segmental flexibility in IgG molecules.
- The extended hinge region of IgA1 molecules allow their Fab arms to range from the "Y" configuration of an IgG or IgA2 to a more open "T" configuration.
- When interacting with multivalent antigens, IgM molecules can adopt a dislocated configuration (staple configuration) in which the Fab arms are bent out of the plane of the Fc regions.

Many features of C$_H$ domain structure exhibited by the Ig isotypes can be understood in the context of this requirement for linkage between antigens and antigen-nonspecific effector molecules.

One property of prime significance for antibody function is intramolecular mobility, often referred to as *segmental flexibility*. Hydrodynamic methods, electron microscopy, X-ray crystallography, and fluorescence polarization have all been used to evaluate the degree of flexibility exhibited by IgG, IgM, IgA, and IgE molecules. In the case of the best-studied isotype, IgG, it is clear that the structural feature most associated with relative motion of one subunit relative to another is the hinge, which connects the C$_H$1 domain to the C$_H$2 domain and is encoded by a separate exon. The human IgG3 subclass has an extended hinge region that can impart increased flexibility. In the case of IgA, the IgA1 hinge is flexible such that the F(ab) arms can range from the typical "Y" configuration to a "T" configuration whereas IgA2 molecules are relatively constrained.

Ig flexibility has important functional consequences for the antibody. First, inter-F(ab) movements can play an important role in permitting antibodies to bind in monogamous bivalent (or multivalent) fashion to antigenic surfaces that display repetitive epitopes. Second, efficiency in precipitation of multivalent antigen molecules or agglutination of multivalent antigen particles can be correlated with inter-F(ab) flexibility. And, third, optimal interactions of effector molecules with IgG antibody Fc regions has been postulated to depend on the ability of the Fc region to bend out of the plane of the F(ab) arms (*dislocation*) (but see discussion on complement activation below).

FUNCTIONS MEDIATED BY ANTIBODY ALONE

Although it is clear that in many *in vivo* situations antibodies mediate their effects with the aid of other molecules and, in selected cases, cells (see next section), there are circumstances where the antibody can influence antigenic targets directly, at least *in vitro*. The very name "antibody" implies the negation of some activity, and antibodies were first defined as factors that could inactivate or neutralize toxins. Subsequent studies have shown inactivation of viruses, parasites, and enzymes, as well.

Virus Neutralization

A phenomenon of fundamental medical and biological importance is the *neutralization* of viruses by antibodies.[14] Although neutralization is defined as the elimination or reduction of the virus's ability to replicate, it does not imply a particular mechanism of interference with the process of replication. Moreover, the measurement of neutralization can depend on the choice of host cell. Thus the neutralizing activity of a given

KEY CONCEPTS

Virus Neutralization

- Antibodies can neutralize (decrease the replication of) viruses by blocking attachment to the host cell, preventing penetration of the host cell membrane, or interfering with uncoating of the virus within the cell.
- Neutralizing antibodies typically recognize proteins or glycoproteins on the virion surface.
- Some antibodies that bind to virion surface proteins or glycoproteins are not neutralizing. In some cases, such antibodies may contribute to immunity, whereas in others, they may enhance infection.
- The magnitude of neutralization mediated by a given antibody may vary with the host cell used for the measurement.
- Neutralization *in vitro* is usually related to protection *in vivo,* but these two properties are not always perfectly correlated.

antibody for a given virus is not an intrinsic property of the antibody but is a property of the relationship between the antibody and the virus, under defined conditions. Consequently, neutralization titers in serum do not always correlate perfectly with protection from infection or disease *in vivo.*

There are several mechanisms by which antibodies can inactivate viruses (Chapter 25). The process by which a virus infects a cell involves multiple steps. These include attachment to one or more membrane components, penetration of or fusion with the membrane, uncoating, and genome expression. Although the most obvious mechanism of neutralization is prevention of viral attachment to the host cell surface, some antibodies can block other steps. For example, neutralizing antibodies for enveloped viruses, such as influenza virus, have been shown to prevent fusion between the virion and cell membranes, and neutralizing antibodies for poliovirus have been shown to interfere with viral uncoating in the host cell.

Different isotypes of antibodies may employ different neutralization mechanisms to varying degrees, although this statement should not be interpreted to mean that there is a one-to-one correspondence between isotypes and neutralization mechanisms. For example, IgG or IgM antibodies in blood can mediate protection against a virus either directly, in some cases, or with the assistance of complement components in others. However, IgA, the dominant isotype in mucosal secretions, operates under conditions where complement is less plentiful than in blood. Thus virus-specific IgA is more likely to utilize virus-inactivating mechanisms that do not require complement, such as prevention of attachment.

Traditional thinking maintains that antibody mediates any protective effects extracellularly. However, it has been reported that IgA antibodies, being transported by the polymeric Ig receptor, can mediate protection against intracellular influenza virus.[15] Similar phenomena have been reported for rotavirus and HIV.

There are several other notable features of antibody–virus interactions. Not all antibodies that bind to molecules on the virion surface will neutralize the virus in all conditions. For a given virus-encoded gene product, such as the influenza virus hemagglutinin, binding of antibodies to some sites, but not others, will effect neutralization. Some gene products on the virion surface may fail to routinely support viral neutralization (*e.g.,* influenza neuraminidase). However, antibody to influenza neuraminidase, while nonneutralizing, is thought to slow the spread of infection by interfering with the escape of progeny virions from infected cells. Nonneutralizing antibodies, or neutralizing antibodies at suboptimal concentrations, have been found in some instances to enhance the infection of host cells by virus (*e.g.,* HIV-1 or dengue virus). It should be noted, however, that the clinical relevance of this enhancement, at least in the case of HIV, remains to be determined. Finally, some nonneutralizing antibodies, or those antibodies that fail to directly neutralize virus in an *in vitro* assay, can mediate protective effects *in vivo,* presumably by engaging antigen-nonspecific effector mechanisms (*i.e.,* complement or Fc receptor-bearing cells) or perhaps through cellular signal transduction.[16]

Neutralization of Toxins and Enzymes

KEY CONCEPTS

Bacterial Immunity

- By neutralizing exotoxins, antibodies can prevent disease mediated by bacterial pathogens such as *Corynebacterium diphtheria* and *Clostridium tetani.*
- Antibodies alone or with complement-derived split products can opsonize pyogenic pathogens such as *Streptococcus pneumoniae.*
- Through activation of the classical pathway of complement leading to assembly of the membrane attack complex, antibodies can mediate destruction of some bacteria, most notably *Neisseria meningitidis* and *N. gonorrhoeae.*
- Antibodies can bind to bacterial adhesins thereby interfering with pathogen attachment to mucosal epithelial cells.
- Antibodies can bind to and inhibit bacterial proteins other than toxins performing critical metabolic or virulence-related functions.

In many bacterial infections, the clinical consequences of infection result from toxic molecules liberated by the bacterial cells rather than from the presence of the microorganisms themselves. Antibodies to such toxins can provide life-saving protection from disease while not directly eliminating the bacteria producing the toxins. A classic example is infection with *Corynebacterium diphtheriae,* which secretes a potentially lethal exotoxin. A more recent example is the emergence of *Clostridium difficile,* which secretes both an enterotoxin (toxin A) and cytotoxin (toxin B). Not only is there a correlation between antibody titers to toxin A and B and prevention of relapse, passive immunotherapy with antibody also prevented relapse.[17]

Bacteria can also produce additional virulence factors, such as enzymes that facilitate spreading of the pathogen through tissues. Host antibodies that inactivate such enzymes can have a beneficial influence on the clinical course. Inactivation of toxins or enzymes is presumed to result from direct competition between antibody and the target molecule or substrate of the toxin or enzyme or from the stabilization or induction of conformations incompatible, to some degree, with the normal function(s) of the toxin or enzyme. However, recent evidence in mice suggests that the protection afforded by exotoxin-neutralizing antibodies can depend on the presence of Fcγ receptors.[18]

FUNCTIONS MEDIATED BY ANTIBODY AND ADDITIONAL MOLECULES OR CELLS

Complement Activation

Regardless of whether binding of antibody to antigen directly mediates protective effects, antibody bound *in vivo* will activate antigen-nonspecific effector mechanisms. The exact mechanisms will depend on the isotype of the antibody as well as on other factors. One critical set of these effector mechanisms is encompassed by the classical pathway of complement activation (Chapters 3, 21). The human antibody isotypes vary considerably in their intrinsic ability to activate this pathway. The consensus view is that IgM, IgG1, and IgG3 isotypes are effective activators. Although some sources state that IgG2, IgG4, and IgA are weak or nonactivators of the classical complement pathway, evidence suggests that when epitope density is high, IgG2 can also activate the classical pathway effectively.[19] Instead of attempting to decide which subclass is absolutely superior, it is more useful to recognize that complement-fixing ability may not be determined solely by the subclass of an IgG antibody.

One obvious source for the isotype-related variation in complement activating ability is variation in affinity for C1q (IgG3 > IgG1 > IgG2 > IgG4), the portion of the first component in the classical pathway that physically contacts the C_H2 domains of antibodies. The intrinsic affinity of the C1q globular heads for Fc regions of any isotype is relatively low, which may account, in part, for the observation that two or more IgG molecules in proximity are required for activation of the classical pathway beginning with C1. Thus in the activation of the classical pathway, the functional affinity of C1q for antibody Fc regions is a crucial parameter.

IgG subclass–associated differences in some measures of complement activation have been found, under some experimental conditions, to depend on quantitative differences in steps of the cascade subsequent to the binding of C1q to antibody. Although there have been speculations regarding the role of segmental flexibility in complement activation, there is no simple correlation between this physical property and activity in fixing the classical complement pathway.[20]

It is generally agreed that IgA does not activate the classical pathway; however, its ability to activate the alternative complement pathway has been controversial. Studies with recombinant IgA molecules have suggested that neither IgA1 nor IgA2 activates either complement pathway. However, aberrantly glycosylated IgA and polymeric IgA may activate the lectin pathway and/or the alternative pathway, and this activation has been postulated to be associated with IgA nephropathy.

Antibody-mediated activation of the classical complement pathway has a variety of potential consequences, including creation of additional sites for attachment to a foreign particle, thereby facilitating ingestion (opsonization), elaboration of substances that mediate leukocyte chemotaxis, additional metabolic changes involved in the destruction of pathogens by leukocytes, and changes in vascular permeability (Chapter 21). In this process, it is the antibody that provides the specificity, whereas the other molecules function without specificity for the epitopes involved.

Receptors for Fc Regions

KEY CONCEPTS

Antibody Effector Systems

- Multivalence of Fc regions is often important in activating antibody effector functions.
- The occurrence and magnitude of effector function activation varies with antibody isotype.
- Effector mechanisms are inherently nonspecific with respect to antigen.
- Antibody Fc regions provide the mechanistic link between antigen-specific V domains and antigen nonspecific effector mechanisms.

The other major system by which antibodies mediate effector functions is cellular. The specific molecules with which cells recognize antibodies are called *Fc receptors* (FcR).[21] In humans, there are several Fc receptors for IgG (FcγRI, FcγRIIa, FcγRIIb, FcγRIIIa, FcγRIIIb), as well as other Fc receptors for IgA, IgE (FcεRI, FcεRII), and IgM (Fig. 15.5). We describe selected features of Fc receptors that illuminate the principles by which they function.

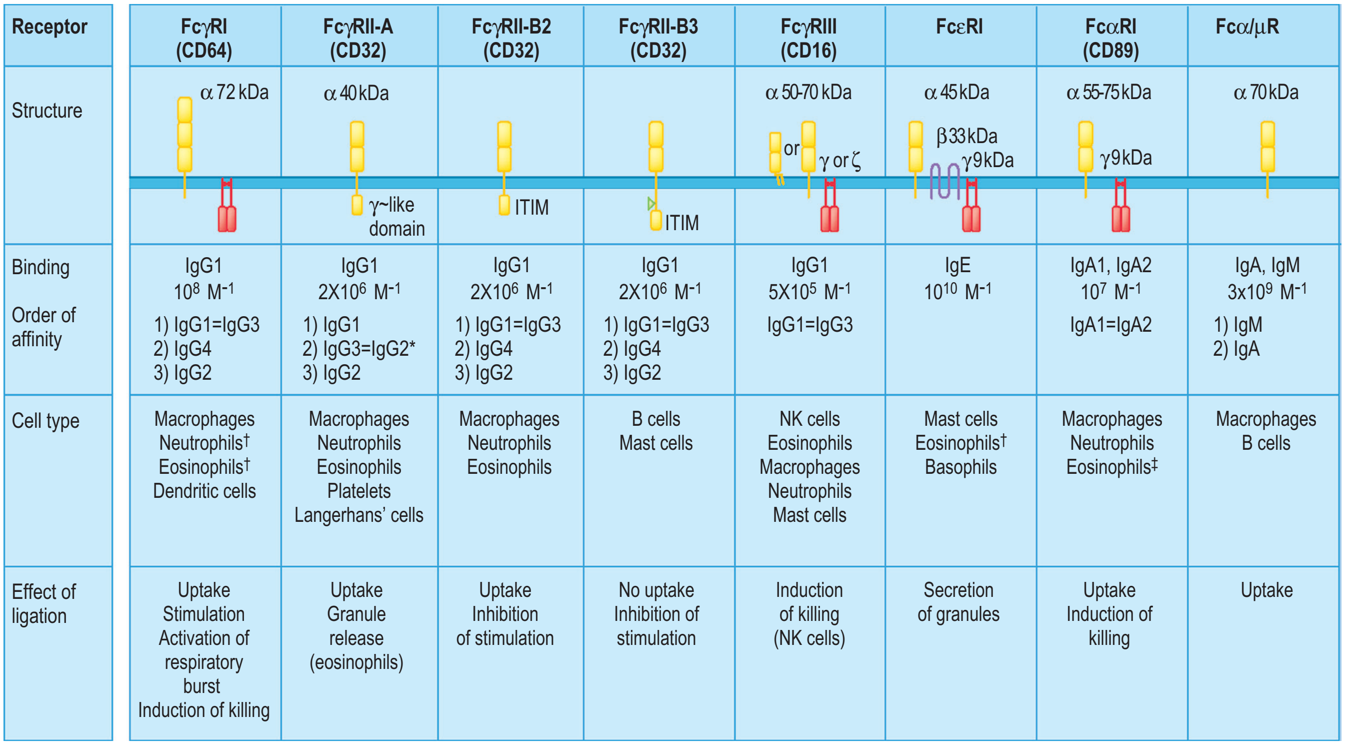
Some receptors (FcγRI, FcεRI) have relatively high intrinsic affinities for antibody molecules and can, therefore, bind significant fractions of monomeric Ig at physiological concentrations. For example, the high-affinity receptor for IgE (FcεRI) binds IgE with an intrinsic affinity of approximately 1×10^{10} M^{-1}. Therefore single IgE molecules can bind to mast cells or basophils through cell surface FcεRI before interacting with allergen (antigen). In contrast, FcγRII and FcγRIII have relatively weak intrinsic affinities for IgG Fc regions. Consequently, multivalent forms of IgG, such as are found in complexes of antibody and multivalent antigens (immune complexes), are much more readily bound to these FcR. Thus for both the complement-dependent and the FcR-dependent effector function pathways, multivalency of Fc regions (functional affinity) plays a critical role.

Several types of functional consequences can follow ligation of FcR by antibody–antigen complexes. These include activation and metabolic alteration of the FcR^+ cells, phagocytosis of antibody-coated particulate antigens, antibody-dependent cellular cytotoxicity (ADCC), and release of mediators that promote inflammation. The end result of Fc binding depends not only on the receptor but also on the cell on which it is expressed and on costimulation, if any, of additional receptors on that cell. As an example, the most studied FcR are those that bind IgG, and these receptors are expressed on many hematopoietic and even nonhematopoietic cells. Within the three classes of receptors (I, II, and III), the latter two FcR exist in two isoforms (A and B). Of interest to the regulation of the immune response, the B isoform for FcγRII transmits an inhibitory signal, and the A isoform transmits an activating signal.

CD89 has been identified in humans as a receptor for IgA, and it is expressed on myeloid cells including polymorphonuclear neutrophils (PMNs), monocytes, and a population of dendritic cells (DCs).[22] Signaling through CD89 involves an ancillary chain that transmits an activating signal. However, not all CD89 molecules associate with this chain, in which case, bound IgA is endocytosed and recycled back to the surface of the cell.[22] Interestingly, Fc binding to CD89 may be more potent at mediating antibody-dependent cellular cytotoxicity than Fc binding to one of the FcγR. Recent data have suggested another possible function that depends on the interaction between antibody (IgA) and a cell-surface receptor able to bind to polymeric Igs (pIgR). Transport of IgA–antigen complexes across epithelial surfaces by pIgR may represent a form of antibody-facilitated antigen excretion.[23]

ANTIBODIES AS SURROGATE LIGANDS

The notion that one molecule can mimic a second molecule, in one respect or another, is of extraordinarily broad applicability and profound biological significance. At least three types of mimicry can be distinguished, and each type can be regarded as a continuous (as opposed to discrete) variable.[24] First, one can conceive of limited structural mimicry of one molecule by another. By chance, two otherwise different molecules could have regions that happen to contain the same or similar (in key



Receptor	FcγRI (CD64)	FcγRII-A (CD32)	FcγRII-B2 (CD32)	FcγRII-B3 (CD32)	FcγRIII (CD16)	FcεRI	FcαRI (CD89)	Fcα/μR
Structure	α 72 kDa	α 40 kDa γ~like domain	ITIM	ITIM	α 50-70 kDa or γ or ζ	α 45 kDa β 33 kDa γ 9 kDa	α 55-75 kDa γ 9 kDa	α 70 kDa
Binding	IgG1 10^8 M^{-1}	IgG1 $2X10^6$ M^{-1}	IgG1 $2X10^6$ M^{-1}	IgG1 $2X10^6$ M^{-1}	IgG1 $5X10^5$ M^{-1}	IgE 10^{10} M^{-1}	IgA1, IgA2 10^7 M^{-1}	IgA, IgM $3x10^9$ M^{-1}
Order of affinity	1) IgG1=IgG3 2) IgG4 3) IgG2	1) IgG1 2) IgG3=IgG2* 3) IgG2	1) IgG1=IgG3 2) IgG4 3) IgG2	1) IgG1=IgG3 2) IgG4 3) IgG2	IgG1=IgG3		IgA1=IgA2	1) IgM 2) IgA
Cell type	Macrophages Neutrophils† Eosinophils† Dendritic cells	Macrophages Neutrophils Eosinophils Platelets Langerhans' cells	Macrophages Neutrophils Eosinophils	B cells Mast cells	NK cells Eosinophils Macrophages Neutrophils Mast cells	Mast cells Eosinophils† Basophils	Macrophages Neutrophils Eosinophils‡	Macrophages B cells
Effect of ligation	Uptake Stimulation Activation of respiratory burst Induction of killing	Uptake Granule release (eosinophils)	Uptake Inhibition of stimulation	No uptake Inhibition of stimulation	Induction of killing (NK cells)	Secretion of granules	Uptake Induction of killing	Uptake

FIG 15.5 **Domain Structures, Binding Properties, Cellular Expression Patterns, and Functional Effects of Human Fc Receptors.** A given FcR may exhibit differences in composition depending on the cell type expressing it. For example, FcγRIII is expressed on neutrophil plasma membranes bearing a glycosylphosphatidylinositol anchor, without FcR γ chains, whereas it is expressed on natural killer (NK)-cell plasma membranes as a conventional transmembrane protein in association with FcR γ chains. Similarly, FcγRIIB1 contains an additional stretch of polypeptide encoded by an exon whose product is not represented in the intracellular domain of FcγRIIB2. This additional portion of the polypeptide is believed to prevent the internalization of FcγRIIB1 subsequent to cross-linking. [a]A subset of FcγRIIA allotypes bind to human IgG2. [b]For these cells, FcR expression is inducible, not constitutive. [c]The molecular weight of CD89α chain is 70–100 kDa in eosinophils. With permission from Janeway CA Jr, Travers P, Walport M, Shlomchik M. Immunobiology: the immune system in health and disease, 6th ed. New York: Garland Science; 2004.

respects), atoms in the same, or almost the same, three-dimensional arrangement. Second, there is mimicry at the level of noncovalent interaction. In this case, the question of interest is whether the model (object of mimicry) and the mimic bind the same receptor sites and with the same affinities. Third, there is mimicry of more complex biological functions, such as cellular or enzymatic inactivation.

It is important to make these distinctions because the extent of mimicry of one type is not a perfect predictor of the extent of mimicry of another type. We have already noted that slight changes in structure sometimes have slight effects on binding affinity or specificity, whereas in other cases, they have dramatic effects on binding affinity or specificity. Thus structural similarity (mimicry), as we perceive it, is not perfectly correlated with mimicry at the level of binding or the level of elicitation of higher biological function.

There are two aspects of receptor–ligand interaction that antibodies can potentially mimic. First, as noted earlier, the inducibility of a vast repertoire of antibody specificities suggests the potential for identifying, through screening or selection, antibodies that can bind any given target molecule at (near) a given site. Thus there should be a reasonable probability of obtaining antibodies that bind to a particular receptor at a site bound by some other, perhaps physiological, ligand or coreceptor. Evidence that antibodies can mimic the functional effects of other molecules is provided by many investigations of antiidiotypic antibodies and conventional antireceptor antibodies.[24]

Second, as noted earlier, the triggering event for many cellular and effector processes in the immune system is the aggregation of receptor molecules by clustered ligands. Therefore the ability of antibodies, which naturally have a maximal valence of ≥2, to cross-link cell-surface molecules and initiate signal transduction contributes to the abilities of antibodies to serve as surrogate coreceptors for cell-surface molecules. This property of antibodies has greatly facilitated the identification and functional characterization of many of these molecules and is also being exploited for therapeutic uses.[25]

FUNCTIONAL PROPERTIES OF ENGINEERED ANTIBODY MOLECULES

Monoclonal Antibodies

Many modern applications of antibodies in research, medicine, veterinary medicine, and other fields rely heavily, although not exclusively, on mAbs. By definition, a mAb population is derived from a clonal population of B-lineage cells (Chapter 7). Therefore all of the molecules express identical variable domains and identical antigen specificities. It is the homogeneity of V-domain structure associated with mAbs that most crucially distinguishes them from polyclonal antibodies, such as those conventionally derived from the serum. Homogeneous antibodies give more reproducible and more easily interpreted results for many kinds of assays.

KEY CONCEPTS

Monoclonal Antibodies From Hybridomas

- Hybridomas are created by fusing normal lymphocytes from immunized animals or immune humans with transformed cells of B-lymphocyte lineage (myeloma cells).
- Monoclonal antibodies (mAbs) produced by a hybridoma are generally homogeneous, expressing a single amino acid sequence.
- Hybridomas can be grown indefinitely in tissue culture.
- Hybridomas can be selected on the basis of the antigenic specificity or functional properties of the mAb secreted by the hybridoma.

mAbs of selected specificity were first produced by cells referred to as *hybridomas*.[26] In contrast, a myeloma protein would be an example of an mAb of unselected specificity. Hybridomas are hybrid transformed cells that are created by the fusion of two types of cells, with each parental cell endowing the hybridoma with desirable properties. One parent of a hybridoma (the fusion partner) is a transformed cell, usually a myeloma cell line, which contributes a metabolism that supports unlimited growth in tissue culture and high rates of Ig synthesis and secretion. Myeloma cell lines currently used as hybridoma fusion partners have been genetically manipulated in such a way that they no longer synthesize an Ig molecule, and they can be selected against in special culture media. The second parental cell is a B lymphocyte that provides the genetic information for the production of a particular antibody. Note that the choice of specificity on the part of the investigator is influenced at two stages in the process, including the choice of immunogen and the nature of the screening assay. By screening a population of hybridomas, most of which will secrete different antibodies, with an appropriately designed assay, it is possible to identify the minority of cells that secrete a mAb of desired specificity.

mAbs are useful for the identification and quantitation of diverse molecules of biological or synthetic origin, including human Igs (*e.g.,* paraproteins), antigens from infectious agents (*e.g.,* HIV p24), hormones, drugs, and toxins. They have also been exploited for therapeutic purposes, such as reversing allograft rejection, killing tumor cells, or preventing cytokine activity contributing to autoimmune disease (Chapter 89).

Recombinant Antibodies

The ability to manipulate the genes that encode antibodies, and thereby manipulate the structures of antibodies, has opened a new era in the study and application of antibodies (Fig. 15.6). Progress includes expression of recombinant intact IgG molecules,[27] expression of Ig fragments [F(ab), Fv] in eukaryotic and prokaryotic host cells, proteomic mining of combinatorial libraries of antibody fragments displayed on the surfaces of filamentous phage,[28] or yeast and bispecific or multispecific antibodies. Recombinant antibodies are also being designed to improve distribution and half-life of administered antibodies (Chapter 89). Not only does antibody engineering contribute to the design of potential therapeutics, but important basic science information has also been gained by using recombinant antibodies. For example, advances in understanding the role of specific effector functions in tumor cell destruction have been possible with the use of mutant and engineered recombinant antibodies, with rituximab being an example.

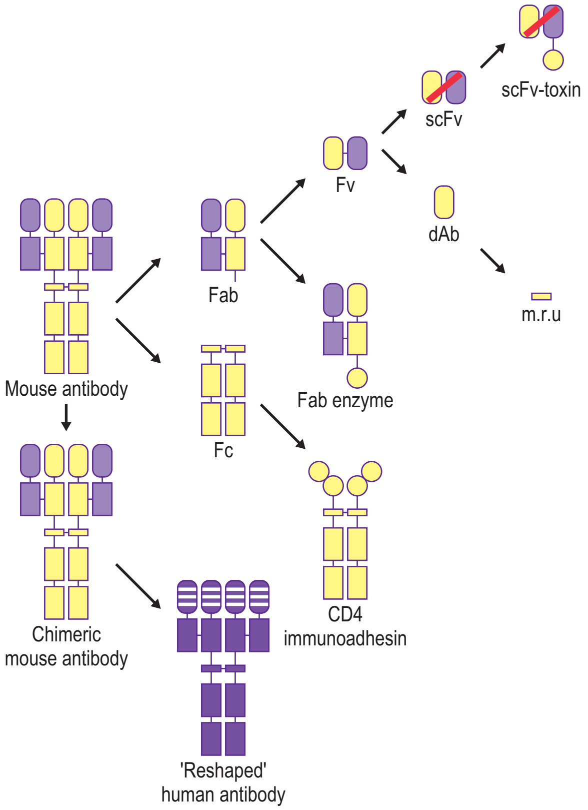


FIG 15.6 **Examples of Engineered Antibodies and Antibody-Derived Fragments That Can Be Created Through the Manipulation of Antibody Genes.** Each closed rectangular (constant) or rounded (variable) box represents a domain. The molecule at the bottom of the figure represents a humanized antibody, where the constant domains and variable domain framework regions correspond to human amino acid sequences. Only the hypervariable regions, and in some cases a small number of framework residues, correspond to mouse or rat antibody amino acid sequences. Other structures depicted include an Fab fragment; an Fv fragment; a single-chain Fv fragment (scFv) in which the C-terminus of the V_H domain is linked covalently by a linker peptide to the N-terminus of the V_L domain; an Fab-enzyme fusion protein; an scFv-toxin fusion protein; an immunoadhesin in which extracellular domains from CD4 have been covalently attached to human heavy chain constant domains; a single V_H domain (dAb); and a peptide derived from a hypervariable region (minimal recognition unit, or mru). With permission from Winter G, Milstein C. Man-made antibodies. Nature 1991;349:293.

The ability to clone Ig genes from single B cells has revolutionized the production of human mAbs. This technology has made possible the identification of broadly neutralizing antibodies for rapidly evolving pathogens, such as HIV-1 and influenza A viruses. These antibodies may prove to be clinically useful.[29] A second recent advance that could be important clinically is the ability

ON THE HORIZON

Recombinant Antibodies—Antibody Engineering

- Enhance effector function potency and half-life through Fc modification.
- Reduce or eliminate bystander effects.
- Isolate and express antibody genes using single B-cell cloning or combinatorial libraries.
- Allow expansion of the antibody repertoire and targeting possibilities.
- Facilitate agents for molecular recognition of minimal size (*i.e.,* nanobodies).
- Develop drugs or toxins that can be covalently attached to novel antibodies for therapy.
- Create multispecific or multimeric constructs.
- Guide vaccine design.
- Evaluate the roles of isotypes in protection or other functions.
- Define how differences in glycosylation can alter effector function.

ON THE HORIZON

Immunity Through Vector-Mediated Expression of Antibodies (Vectored Immunoprophylaxis)

- Host muscle cells transduced with a viral vector encoding antigen- or pathogen-specific antibodies.
- For highly variable pathogens, such as human immunodeficiency virus (HIV)-1, the antibodies are preferably broadly neutralizing.
- Based on model systems in mice, vector-mediated antibody expression can be long-term without immunization.
- Antibodies carrying different variable domains and expressing overlapping reactivities can potentially be simultaneously expressed.

to encode protective antibodies in viral expression vectors that can be introduced into experimental animals and probably humans and make possible the long-term expression of these mediators of immunity at relevant serum concentrations.[30]

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Amit AG, Mariuzza RA, Phillips SEV, et al. Three-dimensional structure of an antigen-antibody complex at 2.8 Å resolution. Science 1986;233:747–53.
2. Greenspan NS. Dimensions of antigen recognition and levels of immunological specificity. Adv Cancer Res 2001;80:147–87.
3. Steward MW. The biological significance of antibody affinity. Immunol Today 1981;2:134–9.
4. Rini JM, Schulze-Gahmen U, Wilson IA. Structural evidence for induced fit as a mechanism for antibody–antigen recognition. Science 1992;255:959–65.
5. Holmberg D, Forsgren S, Forni L, et al. Reactions among IgM antibodies derived from neonatal mice. Eur J Immunol 1984;14:435–41.
6. Kramer A, Keitel T, Winkler K, et al. Molecular basis for the binding promiscuity of an anti-p24 (HIV-1) monoclonal antibody. Cell 1997;91:799–809.
7. James LC, Roversi P, Tawfik DS. Antibody multispecificity mediated by conformational diversity. Science 2003;299:1362–7.
8. Kwong P, Doyle M, Casper D, et al. HIV-1 evades antibody-mediated neutralization through conformational masking of receptor-binding sites. Nature 2002;420:678–82.
9. Mollnes TE, Harboe M. Neoepitope expression during complement activation—a model for detecting antigenic changes in proteins and activation of cascades. The Immunologist 1993;1:43–9.
10. Berzofsky JA. Intrinsic and extrinsic factors in protein antigenic structure. Science 1985;229:932–40.
11. Horiya S, MacPherson IS, Kruas IJ. Recent strategies targeting HIV glycans in vaccine design. Nat Chem Biol 2014;10:990–9.
12. Mannik M. Physicochemical and functional relationships of immune complexes. J Immunol 1980;74:333–8.
13. Heyman B. The immune complex: possible ways of regulating the antibody response. Immunol Today 1990;11:310–13.
14. Dimmock NJ. Neutralization of animal viruses. Curr Top Microbiol Immunol 1993;183:1–149.
15. Mazanec MB, Kaetzel CS, Lamm ME, et al. Intracellular neutralization of virus by immunoglobulin A antibodies. Proc Natl Acad Sci USA 1992;89:6901–5.
16. Binder GK, Griffin DE. Immune-mediated clearance of virus from the central nervous system. Microbes Infect 2003;5(5):439–48.
17. Lowy I, Molrine DC, Leav BA, et al. Treatment with monoclonal antibodies against *Clostridium difficile* toxins. N Engl J Med 2010;362:197–205.
18. Abboud N, Chow SK, Saylor C, et al. A requirement for FcgR in antibody-mediated bacterial toxin neutralization. J Exp Med 2010;207:2395–405.
19. Garred P, Michaelsen TE, Aase A. The IgG subclass pattern of complement activation depends on epitope density and antibody and complement concentration. Scand J Immunol 1989;30:379–82.
20. Tan LK, Shopes RJ, Oi VT, et al. Influence of the hinge region on complement activation, C1q binding, and segmental flexibility in chimeric human immunoglobulins. Proc Natl Acad Sci USA 1990;87:162–6.
21. Nimmerjahn F, Ravetch JV. Fcγ receptors as regulators of immune responses. Nat Rev Immunol 2008;8:34–47.
22. Monteiro RC. Role of IgA and IgA fc receptors in inflammation. J Clin Immunol 2010;30:1–9.
23. Kaetzel CS, Robinson JK, Chintalacharuvu KR, et al. The polymeric immunoglobulin receptor (secretory component) mediates transport of immune complexes across epithelial cells: a local defense function for IgA. Proc Natl Acad Sci USA 1991;88:8796–800.
24. Greenspan NS. Reflections on internal images. Nat Biotechnol 1997;15:123–4.
25. Cragg MS, French RR, Glennie MJ. Signaling antibodies in cancer therapy. Curr Opin Immunol 1999;11:541–7.
26. Kohler G, Milstein C. Continuous cultures of fused cells secreting antibody of predefined specificity. Nature 1975;256:495–7.
27. Morrison SL. In vitro antibodies: strategies for production and application. Annu Rev Immunol 1992;10:239–65.
28. Hudson PJ. Recombinant antibody constructs in cancer therapy. Curr Opin Immunol 1999;11:548–57.
29. Corti D, Lanzavecchia A. Broadly neutralizing antiviral antibodies. Annu Rev Immunol 2013;31:705–42.
30. Schnepp BC, Johnson PR. Vector-mediated in vivo antibody expression. Microbiol Spectr 2014;2.

MULTIPLE-CHOICE QUESTIONS

1. For which of the following phenomena is intrinsic affinity necessarily more important than functional affinity?
   - **A.** Antibody-mediated opsonization of encapsulated bacterial pathogens
   - **B.** Assay using antibodies to detect small drugs in solution
   - **C.** Binding of C1q to immunoglobulin G (IgG) or IgM
   - **D.** Binding of neutralizing IgG antibodies to nonenveloped viruses

2. Which of the following human immunoglobulin G (IgG) subclasses is the least effective in activating the classical pathway of complement?
   - **A.** IgG1
   - **B.** IgG2
   - **C.** IgG3
   - **D.** IgG4

3. Which of the following Fcγ receptors transduces signals that inhibit activation of B lymphocytes?
   - **A.** FcεRI
   - **B.** FcεRII
   - **C.** FcγR1
   - **D.** FcγRIIA
   - **E.** FcγRIIB1

4. Which of the following immunoglobulin (Ig) isotypes is dominant in colostrum and intestinal secretions?
   - **A.** IgA1
   - **B.** IgG2
   - **C.** IgG3
   - **D.** IgD
   - **E.** IgE

16

Helper T-Cell Subsets and Control of the Inflammatory Response

Todd N. Eagar, Stephen D. Miller

T cells are key regulators of the inflammatory processes required to limit infection. CD4 T cells were initially described as "helper" cells because of their ability to promote B-cell antibody production. It is now known that CD4 T cells have a much broader influence through regulation of the functions of phagocytes, granulocytes, and other lymphocyte subsets. This is accomplished through direct cell-to-cell interactions and via the production of cytokines. The contributions of CD4 T cells to host protection are highlighted by the recurrent bacterial, fungal, and viral infections arising in patients with impaired T-cell responses as a result of genetic diseases (Chapter 35), such as severe combined immunodeficiency (SCID), or in patients on immunosuppressive therapies.[1] Conversely, dysregulated T-cell responses are associated with rampant inflammation and autoimmunity, as seen in patients with such conditions as autoimmune lymphoproliferative syndrome (ALPS), autoimmune polyendocrinopathy syndrome type 1 (APS-1) and immune dysfunction, polyendocrinopathy, enteropathy, and X-linked (IPEX) syndrome.[2] Thus a balance must be struck between allowing T-cell inflammation to protect the hosts from pathogens and ensuring control mechanisms to prevent injury from T cell–mediated inflammation. The requisite control is established as T cells transit through stages or checkpoints of the immune response. These include activation, clonal expansion, migration, differentiation, and response termination.

- **Activation:** T-cell function requires antigen-specific T-cell receptor (TCR) signals and accessory signals. This provides fine specificity against discrete protein antigens and limits self-reactivity.
- **Clonal expansion:** Activated T cells rapidly expand in number to scale the immune response.
- **Migration:** T cells traffic through blood, lymphatics, and organs. This allows for surveillance and localized effector responses in diverse tissue types.
- **Effector responses:** T cells adapt their phenotypes to tailor the response to the organism and infected tissue. Long-term memory cells provide protection from reexposure to the same pathogens.
- **Termination:** T-cell responses are limited through inhibitory receptors, cell death pathways, and suppression.

T cells are generated in the thymus from lymphocyte progenitor cells (Chapter 8). Mature T cells enter the peripheral immune system as inexperienced or naïve cells that are incapable of carrying out effector immune responses. To participate in inflammation, T cells must go through an instructional process through defined checkpoints. This chapter will discuss the functional heterogeneity within the CD4 T-cell compartment in the context of the checkpoints involved in the generation and regulation of T-cell biology.

KEY CONCEPTS

T Cell–Mediated Inflammation Requires Activated or Memory T Cells

Naïve T Cells

- Low frequency
- Traffic through circulatory and lymphatic systems
- Require professional antigen-presenting cells (APCs) for activation
- Require strong costimulation
- Delayed expansion
- Delayed cytokine production
- Dependent on interleukin-7 (IL-7)

Activated T Cells

- Traffic through most tissues
- Respond to antigen presented by nonprofessional APCs
- Require less costimulation
- Rapidly expand following antigen encounter
- Rapidly produce effector cytokines
- Dependent on IL-2

Memory T Cells

- Traffic through most tissues
- Dependent on IL-7 and IL-15
- Respond to antigen presented by nonprofessional APCs
- Require less costimulation
- Rapidly expand following antigen encounter

Activation

Activation is required for the naïve T cell to function in effector or memory roles (Fig. 16.1). Full activation requires the cell to receive and integrate signals from a minimum of three diverse types of receptors; TCR (signal 1), costimulatory receptors (signal 2), and cytokines (signal 3).[3] TCR signals are essential for T-cell activation. Each TCR is generated through the process of somatic recombination to recognize major histocompatibility complex (MHC) and peptide complexes (Chapter 4). Engagement of cognate MHC–peptide complex by the TCR activates biochemical signals that lead to changes in transcription and ultimately to T-cell function (Chapter 12). Costimulation and cytokine signals can augment TCR signals by amplifying TCR-induced signal pathways or by eliciting additional signaling cascades. Activation

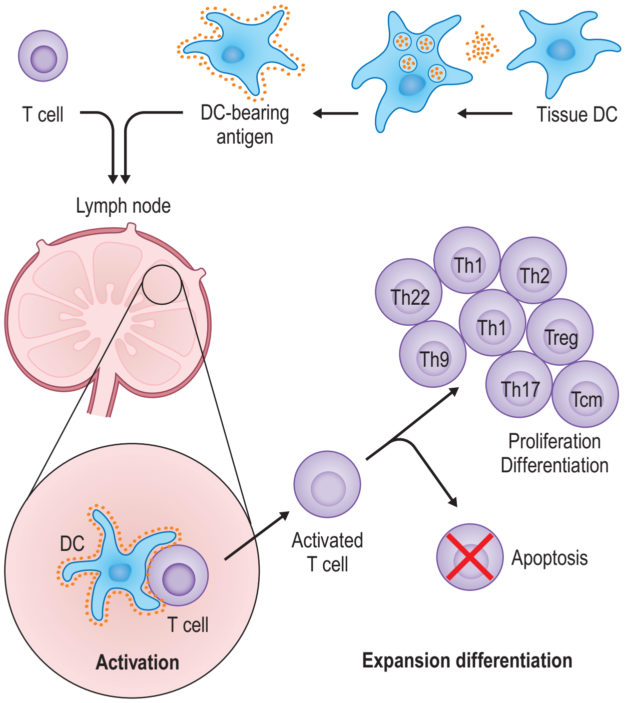


FIG 16.1 **T-Cell Expansion and Differentiation.** Following emigration from the thymus, CD4 T cells travel through the lymph nodes. Inflammation results in the recruitment of tissue dendritic cells (DCs) into the draining lymph nodes. In the lymph nodes, DCs present antigen to naïve T cells. Effective antigen presentation stimulates the activation, proliferation, and differentiation of T cells into Th1, Th2, Th9, Th17, Th22, or uncommitted T central memory (T_{cm}) cells.

requires extensive interactions between T cells and antigen-presenting cells (APCs; Chapter 6). These interactions lead to the formation of a multimolecular signaling complex known as an *immunological synapse*. This synapse brings critical receptors and intracellular signaling components that are required for activation into close proximity to elicit the activation process.[4]

TCR activation is dependent on the interactions between T cells and APCs. The primary APC type involved in activating naïve T cells is the dendritic cell (DC).[5] DCs are specialized phagocytes that are widely distributed in tissues throughout the body (Chapter 2). Mature DCs are efficient APCs and potent stimulators of naïve T cells. In response to infection or damage, DCs migrate from the tissue into the regional lymph nodes, where they enter the T cell–rich areas.[6] In the lymph node, T cells interact with numerous DCs as they scan for cognate antigen. In the absence of TCR signaling, these interactions are of short duration; however, TCR engagement elicits T cell–DC interactions that may last up to 20 hours.[7] Long-term interactions are required for full activation. In the presence of incomplete stimulation, that is, if costimulation is lacking, T cells are prevented from participating in inflammation and may become unresponsive to further stimulation (anergic) or undergo apoptosis. This is a regulatory mechanism used to prevent inappropriate T-cell activation and helps maintain self-tolerance in T cells.

T-cell activation depends on the maturation state of the DC. Immature DCs are incapable of stimulating functional T-cell responses and may instead induce unresponsiveness, whereas mature DCs readily elicit T-cell activation. DC maturation is regulated by external stimuli, such as Toll-like receptor (TLR) stimulation. T cells also stimulate DC maturation by expressing CD40 ligand (CD40L or CD154). Mature DCs instruct T cells through expression of MHC, costimulatory ligands, and cytokines. The expression of these molecules provides context to the T-cell activation process and helps tailor the immune response. For instance, TLR stimulation may elicit the expression of specific types or levels of costimulation and cytokine allowing mature DCs to promote antibacterial T-effector responses. Tissue-specific factors, such as vitamin D, transforming growth factor-β (TGF-β) and interleukin-10 (IL-10), alter costimulatory ligand and cytokine expression by the DC and reduce T-cell activation. Thus DCs transmit information from tissues to either promote specific immunity or suppress potentially damaging T-cell responses.[8]

T-cell activation defects contribute to several immunodeficiency diseases. T cells from patients with Wiskott-Aldrich syndrome, for instance, have defects in the formation of stable TCR signaling complexes. T cells in leukocyte adhesion deficiency 1 have a reduced capacity to become activated in response to stimulation due to defects in integrin β_2. Similarly, patients with mutations in the TCR signaling components lck and Zap70 show defective T-cell activation. In contrast, mutations in inhibitory receptors, including CTLA-4 and PD-1, are thought to contribute to susceptibility in autoimmune diseases.

The T-cell activation process is an important target for immunosuppression. Therapies targeting key biochemical pathways leading to activation are used to prevent rejection in solid organ transplantation, to treat graft-versus-host disease (GvHD) in bone marrow transplantation, and to reduce the severity of autoimmune diseases or other immune-mediated disorders. Cyclosporine and tacrolimus block the enzyme calcineurin and thereby inhibit TCR-induced nuclear factor of activated T cells (NFATs) signals. Rapamycin inhibits the kinase mammalian target of rapamycin (mTOR), which is induced by TCR and CD28 or by IL-2 receptor signals.[9] Monoclonal antibodies (mAbs) and fusion proteins that target T-cell activation are also applied for the treatment of autoimmune diseases and the prevention of graft rejection. A fusion of the extracellular domain of cytotoxic T lymphocyte antigen-4 (CTLA-4) with the Fc portion of immunoglobulin G (IgG; CTLA-4-Ig) is used to prevent T-cell activation by blocking the interactions between CD28 and its B7 ligands. Antibodies against the IL-2R block IL-2–driven T-cell activation. On the opposite side of the spectrum, antibodies that block the inhibitory receptors PD-1 and CTLA-4 are employed to promote or prolong T-cell activation and thereby enhance antitumor T-cell responses.[10]

Clonal Expansion

T-cell activation is detectable within hours of antigen exposure; however, days are required until T-cell responses can be detected from a naïve host. This delay partially reflects the need for T-cell expansion to produce a detectable and effective response. The initial frequency of T cells for a single epitope is estimated at approximately one per million total CD4 T cells or approximately 1×10^5 cells per human.[11] Following antigen exposure, cellular proliferation is initially slow, taking 2 days for the first divisions to occur in the lymph nodes and the spleen. Cell numbers increase rapidly between days 2 and 7 after antigen exposure, when many T cells have undergone as many as eight rounds of division. This process is known as *clonal expansion* (see Fig. 16.1). For instance, a single precursor T cell following eight rounds of division can yield 256 daughter cells, each possessing the same TCR. Through the combined proliferative responses of multiple T-cell clones to the same or different epitopes produced by the same pathogen, a rapid net increase in pathogen-reactive T cells occurs.

In the clinical setting, T-cell proliferation is a target for immunosuppression. Some drugs, such as azathioprine, methotrexate, and mycophenolate mofetil, impair T-cell proliferation by inhibiting nucleotide synthesis. Cyclophosphamide suppresses DNA replication by alkylating guanine residues. These agents are used in the treatment of various autoimmune diseases and for the prevention of graft loss in transplantation.

TRAFFICKING

Following activation, T cells change their migration patterns (Chapter 11). T-cell migration relies on the actions of selectins, chemokines (Chapter 10), integrins, and matrix proteases. Together, these permit vessel transmigration into lymphatic tissues or sites of inflammation.[12] The differential trafficking of naïve and effector T cells results from varied expressions of selectins, chemokine receptors, and integrins. The circulation of naïve T cells is restricted to blood, lymph nodes, and lymphatic ducts. This pattern of movement relies on the expression of chemokine receptor 7 (CCR7), L-selectin or (CD62L), and LFA-1($\alpha_L\beta_{2,}$ integrin, CD11a, CD18). CCR7-expressing T cells follow gradients of CCL21 into T cell–rich areas in the lymph nodes, Peyer patches, and spleen. L-selectin allows rolling on the high endothelial venule (HEV). LFA-1 and intercellular adhesion molecule-1 (ICAM-1) permit firm adhesion and extravasation. Migration to the Peyer patches also requires $\alpha_4\beta_7$ integrin in addition to leukocyte function–associated antigen-1 (LFA-1). T-cell retention in and egress from the lymph node are regulated by sphingosine 1 phosphate (S1P), a secreted phospholipid. S1P sensitivity is mediated by sphingosine 1 phosphate receptor 1 (S1P1), a G protein–coupled receptor. Recirculating T cells downregulate S1P1. As a result of low S1P levels in the lymph node, S1P1 expression allows exit from the lymph nodes in 12–18 hours after entry.[13]

During activation, T cells alter the patterns of homing receptors to reduce the cell's ability to migrate within the secondary lymphatic system while gaining receptors needed for trafficking in the peripheral tissues. Very early in the activation process, T cells downregulate S1P1 and express CD69, which suppresses chemotactic migration in response to S1P, shed L-selectin, and express the ligands for E- and P-selectin. Additionally, T cells downregulate CCR7 expression and express a new set of chemokine receptors. Following activation, LFA-1 expression is markedly increased, and the expression of additional integrin molecules is also induced. One such integrin, very late antigen-4 (VLA-4), is induced by activation and permits T-cell migration into the central nervous system (CNS), lungs, and intestines.[14]

Differential expression of homing receptors between naïve and activated T cells facilitates their diverse functions. Naïve T cells in the lymphatic tissues are in close proximity to a concentrated collection of DCs from many tissues; this facilitates antigen encounter. The function of activated T cells in immune surveillance, in contrast, is facilitated by migration through diverse tissues.

The multistep process of migration and the tissue selectivity of migration receptors have been targeted for the treatment of autoimmune diseases. Antibodies targeting CD62L and LFA-1 have been developed to prevent the entry of T cells into lymph nodes. In addition, the S1P1 inhibitor drug, fingolimod, has been recently approved for the treatment of multiple sclerosis. Fingolimod is thought to prevent the egress of recently activated T cells from the lymph nodes. Natalizumab, an mAb against the α_4 integrin, is utilized in the treatment of multiple sclerosis and Crohn disease to prevent the migration of T cells to the CNS and intestines, respectively.

DIFFERENTIATION OF CD4 TH SUBSETS

Naïve T Cells

Naïve T cells are precursors for effector and memory T cell subsets. Phenotypically, naïve T cells are small cells with little cytoplasm; they express surface markers, such as CD45RA, CCR7, CD62L, CD127, and CD132. They lack expression of markers of previous activation, such as CD25, CD44, CD69, CD45RO, or HLA-DR. Naïve T cells have low metabolism and are unable to produce proinflammatory cytokines.[15] Naïve T cells migrate within the secondary lymphoid organs, where they interact with DCs. Functionally, naïve T cells are not capable of mediating effector immune responses. Naïve T cells are predominant after birth but decline in percentage after puberty. Because of their role as precursors for all T-effector subsets, their loss is thought to be an important contributor to the immunosenescence seen in older adults.

Early studies demonstrated that CD4 T-cell clones and cell lines would express reproducible patterns of cytokine expression designated T-helper cell-1 (Th1) and T-helper cell-2 (Th2).[16] It is now appreciated that CD4 effector T cells can differentiate into many functional phenotypes (see Fig. 16.1). These can be grouped into four general categories: *(i)* those possessing proinflammatory effector characteristics, *(ii)* those possessing regulatory or antiinflammatory activities, *(iii)* those that function to promote B-cell follicle development, and *(iv)* those that provide long-term memory. Our understanding of the phenotype diversification process is rapidly expanding. From the T-effector differentiation models, we learn of three steps. First, TCR signals lead to activation and elicit the expression of cytokine receptors.[17] Second, signals through specific cytokine receptors differentially promote the expression of lineage-specific "master" transcription factors that promote expression of genes associated with a particular T-cell phenotype and suppress the expression of genes associated with other T-cell phenotypes. Third, the phenotype specific transcription factors induce epigenetic changes that control gene accessibility and maintain the T-cell phenotype in a cell-intrinsic manner.

EFFECTOR CELL PHENOTYPES

T-effector cells promote inflammation through the release of cytokines (Fig. 16.2). Cytokine release elicits the activity of accessory cells, which ultimately mediate the inflammation to clear the antigen. The pattern of cytokine production ultimately dictates the type of inflammation (Table 16.1). Effector T cells are divided into five basic groups: Th1, Th2, Th9, Th17, and Th22. Table 16.2 summarizes responses the T-effector responses associated with selected pathogens.

Th1

Th1 cells are defined by production of the cytokines interferon (IFN)-γ, granulocyte macrophage–colony-stimulating factor (GM-CSF), IL-2, and lymphotoxin (LT, tumor necrosis factor

KEY CONCEPTS

CD4 Effector Phenotypes

Th1

- Produce interferon (IFN)-γ, interleukin (IL)-2, tumor necrosis factor (TNF)-α and TNF-β (LT-α)
- Stimulate Immunoglobulin (Ig)G2 and IgG3 class switching
- Responses mediated by macrophage activity
- Promotes phagocytic activity through:
  - FcγRIII cross-linking
  - Complement deposition
  - Opsonization
  - IFN-γ-mediated macrophage activation

Th2

- Produce IL-4, IL-5, IL-6, and IL-10
- Stimulate IgG4 and IgE class switching
- Responses mediated by mast cells and eosinophils
- Increases degranulation through:
  - FcεRI cross-linking
  - IL-5–mediated eosinophil activation

Th17

- Produce IL-6, IL-17, and TNF-α
- Activate local endothelium
- Induce cytokine and chemokine production
- Increase infiltration by neutrophils
- Activates cell-mediated inflammation

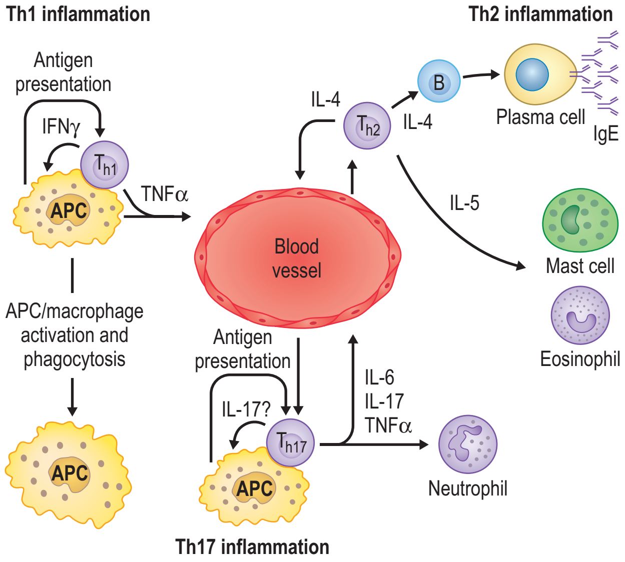


FIG 16.2 Generalized Model for T Helper (Th)–Mediated Inflammation. *Top left.* Introduction of infectious agent stimulates the release of chemokines and tumor necrosis factor (TNF)-α from tissue macrophages, stimulating the recruitment *(upward arrows)* of T cells and monocytes through the local vasculature. Antigen recognition by T cells stimulates the local production of Th1 cytokines. Interferon (IFN)-γ activates macrophages enhancing the clearance of infectious agents. *Top right.* Trafficking to sites of Th2 responses is stimulated by local chemokine expression leading to T-cell recruitment. Antigen recognition leads to interleukin (IL)-4 production by Th2 cells, which stimulates B-cell immunoglobulin (Ig)E class switching. Production of IL-5 activates eosinophils. Cross-linking of FcεR1 molecules bound to IgE leads to the degranulation of mast cells (Ma) and eosinophils (E). *Bottom.* Recruitment of Th17 cells and restimulation by antigen results in the release of IL-6, IL-17, and TNF-α, which promotes the recruitment, activation, and function of many cells, particularly neutrophils.

[TNF]-β). Th1 differentiation is elicited in response to infection by intracellular bacteria, fungi, and viruses. Infection by these pathogens leads to the production of cytokines by DCs and natural killer (NK) cells. Th1 differentiation is promoted by IL-12, IL-18, IFN-γ, and type 1 interferons and is inhibited by IL-4, IL-10, and TGF-β (Fig. 16.3). During activation, IFN-γ receptor signals activate signal transducer and activator of transcription 1 (STAT1) and promote the expression of the Th1-restricted transcription factor T-bet and IL-12R expression. IL-12R signals through STAT4 then drives the expression of high levels of T-bet. T-bet serves to reinforce the Th1 phenotype by promoting IFN-γ and IL-12Rβ_2 expression.[17] IL-18 plays a dual role in Th1 function by promoting Th1 commitment and eliciting IFN-γ production by fully differentiated Th1 cells.

Th1 cells promote cell-mediated inflammatory responses through inducing the activation of macrophages, NK cells, B cells, and CD8 T cells. Th1 cells regulate macrophage function at several levels. GM-CSF promotes the production of monocyte lineage cells from the bone marrow. IFN-γ is a potent macrophage activator, enhancing microbicidal activity by initiating nitric oxide (NO) production, upregulating production of oxygen radicals, and increasing phagocytic function. IFN-γ also promotes antigen presentation by upregulating MHC class I molecules, MHC class II molecules, and costimulatory molecule expression by macrophages (see Fig. 16.2). IFN-γ can activate NK cells and also promote humoral responses in B cells to mediate antibody class switching to an IgG1 (IgG2a in mice) isotype.[18] IgG1 activates the classical complement pathway and can bind Fcγ receptors expressed on phagocytic cells thereby promoting opsonization. Finally, IFN-γ acts in conjunction with another Th1 cytokine, IL-2, to promote the differentiation of CD8 cells into cytotoxic effector cells (Chapter 17). Macrophage-dependent Th1-mediated inflammatory responses are known as *delayed-type hypersensitivity* (DTH) responses. *In vivo*, DTH responses are critical for protection from intracellular pathogens including bacteria, fungi, and viruses. Individuals with defects in components of the IFN-γ or IL-12 signaling pathways are susceptible to atypical mycobacterial infections and have altered responses to other microorganisms. Additionally, Th1 cells are thought to contribute to the pathogenesis of autoimmune diseases, including multiple sclerosis, type 1 diabetes mellitus, rheumatoid arthritis, and Crohn disease. The Th1 pathway has been explored therapeutically. IL-12 has been studied as a way to boost vaccine immune responses against microorganisms and antitumor immunity. IL-12 neutralizing mAb-based therapies have been investigated in inflammatory and autoimmune diseases, including psoriasis, arthritis, Crohn disease, and multiple sclerosis.

Th2

Th2 cells are defined by their production of IL-4, IL-5, IL-9, IL-10, and IL-13. *In vitro*, the critical step in Th2-cell differentiation is the presence of exogenous IL-4 and the absence of IFN-γ during T-cell activation. *In vivo*, Th2 differentiation is thought to require IL-4 produced by basophils, eosinophils, mast cells, NKT cells, or even previously differentiated Th2 cells. Naïve T

TABLE 16.1 **T-Helper (Th) Cell Effector Function Through Cytokine Secretion**

Th1	
Interleukin (IL)-2	T-cell growth and potentiation of Fas-mediated apoptosis
	Natural killer (NK) cell growth and cytolytic activity
	B-cell growth and antibody production
Interferon (IFN)-γ	Increases class I and class II molecule expression on numerous cell types
	Promotes Th1 differentiation
	Activates macrophages; stimulates phagocytic killing and oxidative bursts
	Induces immunoglobulin (Ig)G2α and IgG3 class switching
	Inhibits IgG1 and IgE class switching
	Inhibits Th2 proliferation
	Activates neutrophils
	Activates endothelium to promote CD4 T-cell adhesion
	Stimulates NK cell cytolytic activity
	Required for the activation of CD8 cytotoxic T lymphocytes (CTLs)
Tumor necrosis factor-α (TNF-α)	Activates vascular endothelial cells; enhances leukocyte recruitment
	Activates neutrophils, eosinophils, and macrophages
	Stimulates IL-1, IL-6, TNF, and chemokine expression by macrophages
	Protects against viral infections, similar to interferons (IFNs)
	Increases class I molecule expression
	Induces fever
	Activates coagulation
Lymphotoxin (LT) (TNF-β)	Activates neutrophils and osteoclasts
	Activates vascular endothelial cells; enhances leukocyte recruitment
	Cytotoxic activity against tumor cells
	Stimulates adhesion molecule expression
	Increases class I molecule expression
Th2	
IL-4	Growth and differentiation of Th2 cells
	B-cell growth and class II molecule upregulation
	Induces IgG1 class switching (IgG4 in humans) and IgE
	Inhibits IgG2a and IgG3 class switching
	Mast cell growth
	Inhibits macrophage activation
	Induces vascular cell adhesion molecule (VCAM) expression on endothelium; recruits eosinophils/monocytes
IL-5	B-cell growth and activation
	Eosinophil differentiation
	Eosinophil activation and survival
IL-10	Inhibits cytokine and chemokine production in monocytes, especially TNF, IL-1, and IL-12
	Inhibits macrophage activation and function
	Inhibits T cell–mediated inflammation
	Induces IgG4 class switching (IgG1 in mice)
IL-13	Upregulates class II molecules on monocytes and B cells
	Inhibits proinflammatory cytokine production by monocytes
	B-cell costimulation
	Induces IgG4 class switching (IgG1 in mice) and IgE
	Increases class I molecule expression
Th9	
IL-9	Mast cell expansion and recruitment
	T-cell growth factor
IL-21	Promotes Th17 differentiation
	B-cell survival; antibody class switching
	Enhances NK cell proliferation; promotes killer function
Th17	
IL-17	Increases T-cell proliferation
	Promotes neutrophil recruitment and activity
	Promotes cytokine production including IL-6 and TNF-α
	Induces chemokine production
Granulocyte macrophage–colony-stimulating factor (GM-CSF)	Promotes myeloid cell development and maturation
	Promotes dendritic cell differentiation and survival
	Enhances macrophage activation
IL-6	Activates acute-phase response, inducing fever and antibacterial responses
	Promotes Th2 and Th17 differentiation
	Activates and elicits s NK responses
	Promotes plasma cell differentiation and Ig production
IL-21	See above.
TNF-α	See above.
Th22	
IL-22	Activates acute phase response, inducing fever and antibacterial responses
IL-13	See above.
TNF-α	See above.

TABLE 16.2 T-Cell Effector Responses to Selected Pathogens

Organism	Nature of the Immune Response
Bacteria	
Borrelia burgdorferi	T-helper (Th1) responses associated with protection and joint pathology
Chlamydia trachomatis	Th1 responses protective and source of pathology
Helicobacter pylori	Th1 responses involved in ulcer formation
Legionella pneumophila	Th1 responses associated with immunity
Listeria monocytogenes	Th1 responses are protective: interferon (IFN)-γ from γδ T cells and CD8 T cells is important
Mycobacterium leprae	Severity and phenotype of disease depends on Th1 and Th2 predominance
Mycobacterium tuberculosis	Th1 responses control infection
Treponema pallidum	Th1 resolves infection; Th2 chronic
Yersinia pestis	Th1 and Th17 responses associated with immunity
Fungi	
Aspergillus fumigatus	Th2 production predominates; does Th1 offer protection?
Blastomyces dermatitidis	Th1 protects; Th2 switch in progressive disease
Candida albicans	Th17 responses are protective
Cryptococcus neoformans	Susceptibility associated with Th2 response; Th17 response associated with protection
Paracoccidioides brasiliensis	Infection stimulates Th2 response, but Th1 response protects
Parasites	
Leishmania spp.	Th1 responses are protective; Th2 responses allow chronic infection
Filaria	Initiates Th2 production; Th1 response appears protective
Schistosoma mansoni	Th1 and humoral responses protect; typically Th2 responses directed against eggs
Trypanosoma cruzi	Th1 responses inhibit parasite replication, but protection is not completely CD4 dependent
Giardia lamblia	Th1 and Th2 responses protect
Viruses	
Measles	Th1 responses are protective
Hepatitis B	Th1 responses seen in spontaneous recovering patients
Human immunodeficiency virus	Shift from Th1 to a Th0 (Th2?) response late in disease correlates with susceptibility to pathogens
Respiratory syncytial virus	Th1 response protects; Th2 response kills

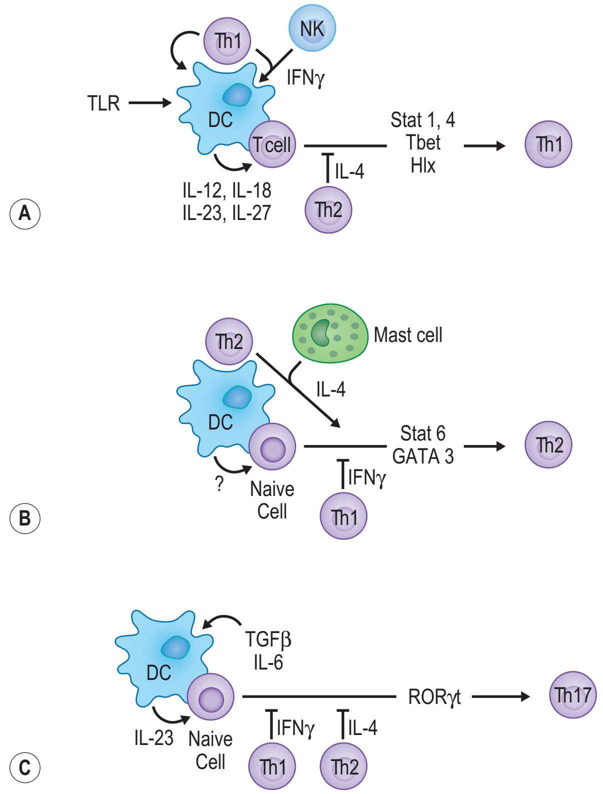


FIG 16.3 Factors Influencing T-Effector Differentiation. Cytokine exposure during the activation stage of naïve T cells strongly influences T-effector differentiation. Depicted here are the factors promoting and inhibiting Th1, Th2, Th17, and T-follicular helper (Tfh) cells following functional activation of undifferentiated T cells.

cells express the IL-4R, and the combination of TCR, costimulatory (CD28 and ICOS), and IL-4R/STAT6 signaling promotes IL-4 transcription and the production of the transcription factors c-Maf and GATA3. c-Maf helps to establish Th2 polarity by promoting IL-4 and suppressing IFN-γ production. GATA3 also appears to inhibit IFN-γ production but additionally plays a critical role in establishing Th2 cells by promoting IL-5 and IL-13 production (see Fig. 16.3).[19]

Th2 cells release IL-4 and IL-5, which attract and activate the function of eosinophils and mast cells. Th2-type cytokines enhance B-cell class switching toward IgG2, IgE, and sIgA. High levels of IgE in Th2 reactions combined with antigen exposure and FcεR1 receptor expression by eosinophils or mast cells results in triggering and release of inflammatory factors, such as histamine, platelet-activating factor, prostaglandins, and leukotrienes (see Fig. 16.2). These factors act on the local environment, producing vascular dilation and leakage, bronchial constriction, and intestinal hypermotility. On a more systemic level, anaphylaxis may be produced. Eosinophil- and mast cell–dependent reactions are known as *immediate-type hypersensitivity* (ITH) responses. ITH responses are important for ridding the body of intestinal helminths; in fact, components of helminth eggs strongly promote Th2 differentiation. Th2 responses are also associated with atopy and hyperresponsive airway conditions, such as asthma and allergies.

Th17

Th17 T cells are characterized by the production of IL-17a/f, IL-21, IL-22, IL-26, GM-CSF, and TNF-α. Th17 differentiation is promoted by IL-1β, IL-6, IL-23 and TGF-β, and the absence of type 1 interferons, IFN-γ, and IL-4 (see Fig. 16.3). Much of the potency attributed to Th17 cells derives from the production of IL-17 and GM-CSF. IL-17, together with TNF-α, strongly promotes inflammation by inducing the expression of adhesion molecules, proinflammatory cytokines (including IL-6, GM-CSF, granulocyte–colony-stimulating factor [G-CSF]), chemokines, prostaglandin E2, and matrix metalloproteinases (see Fig. 16.2). IL-21 regulates B-cell, T-cell, and NK-cell functions.[20] In B cells, IL-21 has been found to regulate plasma cell differentiation and promote IgM and IgG1 antibody production. IL-21 is a T-cell growth factor, and in the presence of TCR signals, IL-21 promotes T-cell activation, proliferation, and survival. It therefore appears that Th17 cells are important for the recruitment of effector cells, including neutrophils and monocytes.

In many ways, Th17 cells are potent cell-mediated effectors. Th17 responses are elicited in response to infection with extracellular bacteria and fungi. In addition, Th17 cells have been identified in patients with atopic dermatitis, Crohn disease, psoriasis, and multiple sclerosis. Experimental models demonstrate that Th17 cells, via the production of GM-CSF and IL-17, are important mediators of autoimmune diseases, including experimental autoimmune encephalomyelitis (EAE).[21] Patients with impaired IL-17 expression or mutations causing Th17 differentiation experience recurrent infections with *Staphylococcus aureus* and *Candida albicans.* Alleles of the IL-23R have been found to confer protection or susceptibility to Crohn disease and psoriasis. Clinical trials are underway to examine the effects of disrupting Th17 immune responses in patients with autoimmune diseases. mAbs to IL12/23 p40, IL23 p19, and IL-17a have been approved for treatment of psoriasis and are being investigated for application in other conditions.

Th9

Although Th2 cells have long been known to produce IL-9, recently, it has been proposed that an IL-9 producing T-cell subset is a separate lineage. Human Th9 cells have been shown to produce IL-9 and IL-21. IL-9 is an important growth factor for T cells, mast cells, and hematopoietic stem cells (HSCs). IL-9 functions by preventing apoptosis. The differentiation of naïve T cells into a Th9 phenotype is thought to be dependent on the presence of IL-4 and TGF-β.[22] Th9 cells express the transcription factor PU.1, which is thought to repress Th2-associated GATA3 and Th1-associated T-bet expression. Similar to Th2 cells, Th9 cells are thought to play a role in protection from helminths. In addition, Th9 cells play a role in asthma, where they promote airway restriction and mucus secretion. Experimental models also suggest that Th9 cells may contribute to the pathogenesis of autoimmune disease.

Th22

Although IL-22 was originally described as a Th17 cytokine, research has shown that a distinct CD4 T population expressed IL-22 independently of IL-17. Further work identified IL-22, IL-13, and TNF-α-secreting cells in the skin of patients with inflammatory skin disorders.[3] These IL-22–expressing cell populations also express the skin homing receptors CCR10. These cells, labeled Th22 cells, have been subsequently found to differentiate in the presence of IL-6 and TNF-α.[23] Phenotypically, Th22 cells express a different gene profile than other Th cell types, suggesting that they are a distinct effector population. They also produce fibroblast growth factors that promote epidermal repair. The hallmark cytokine IL-22 is a member of the IL-10 family with important functions in regulating skin homeostasis and protection from infection. Experimental models suggest that IL-22 and Th22 cells may play roles in regulating inflammation in other epithelial tissues.

KEY CONCEPTS

Regulation of T-Cell Responses

- Cell death pathways
  - Tumor necrosis factor (TNF) family receptors: Fas/FasL, TNF-related apoptosis inducing ligand (TRAIL), and TNFR1
- Inhibitory receptor activity
  - Cytotoxic T lymphocyte antigen-4 (CTLA-4) and programmed death-1 (PD-1)
  - Inhibition of T-cell receptor (TCR) signaling
- Cytokine-mediated regulation
  - Antiinflammatory cytokine production (interleukin [IL]-10, transforming growth factor-β [TGF-β])
  - Cytokines that promote and inhibit inflammation at different stages of a response (interferon-γ [IFN-γ], IL-2, IL-27)
  - Regulation of apoptosis (LT, TNF-α)
  - Growth factor withdrawal (IL-2, IL-7, IL-15)

REGULATORY T CELLS

Regulatory T cells (Tregs) possess the ability to suppress or otherwise downregulate the function of other proinflammatory T cells. Although multiple T-cell subsets may possess regulatory T cell activity, our understanding of regulatory function was greatly enhanced by the discovery of CD25 and FoxP3 as markers for a regulatory subpopulation of CD4 T cells (Chapter 18). This discussion focuses on the involvement of two populations of peripherally derived Treg phenotypes: adaptive Treg and Tr1 cells.

Adaptive Tregs

In mice, the majority of $CD4^+$ $CD25^+$ $FoxP3^+$ Tregs (natural Tregs) are thought to develop regulatory ability in the thymus. Peripherally derived adaptive Tregs are similar in phenotype to natural Tregs in that they express CD4, CD25, CD38, CD62L, CD103, and FoxP3. Adaptive Tregs are thought to be generated in response to prolonged exposure to antigen and are influenced by antigen presentation by immature DCs, plasmacytoid DCs, or nonprofessional APCs. Experiments suggest that the presence of inhibitory cytokines, such as IL-10 and TGF-β, also favors adaptive Treg differentiation. Similar to natural Tregs, adaptive Tregs are capable of suppressing CD4 and CD8 T-cell responses. Unlike natural Tregs, adaptive Tregs can function through both cell contact–dependent interactions and through the secretion antiinflammatory cytokines. These cytokines may directly regulate effector T cells or may inhibit DC activity. The generation of adaptive Tregs is of great interest as a cell-based therapy for inflammatory diseases, including autoimmunity.[24]

Tr1 Cells

Tr1 cells have been defined as a population of CD4 T cells that produce IL-10 in response to stimulation. Unlike adaptive Tregs, Tr1 cells do not express FoxP3 and do not necessarily express CD25. Tr1 cells are produced in the respiratory tract when antigen exposure is coupled with IL-10 and TGF-β. Experimentally, Tr1 cells can be induced by repeated intranasal antigen delivery.[25] Respiratory tract infections with *Bordetella pertussis* have also been found to elicit Tr1 cells *in vivo*.[26] Tr1 cells function by producing IL-10 to block proliferation of naïve T cells, prevent

Th1 differentiation by blocking IL-12 production, and enhance differentiation of T cells toward a Tr1 phenotype.

FOLLICULAR HELPER T CELLS (TFH)

One of the earliest described roles of CD4 T cells is their ability to promote or “help” generate effective antibody production by B cells. Th1 and Th2 cells have been described for their ability to promote B-cell class switching and effector responses. Evidence supports the role of a distinct subpopulation of T cells, known as *T-follicular helper* (Tfh) cells, as playing a key role in regulating T cell–dependent B-cell responses. Tfh cells are identified by their surface expression of CXCR5, inducible T-cell costimulator (ICOS), programmed death-1 (PD-1), CD200, B- and T-lymphocyte attenuator (BTLA), OX40, and serum amyloid P (SAP), while lacking the expression of CD127 and CCR7.[27] Tfh cells also express the transcription factor BCL6, which is essential for their differentiation and function. Interestingly, BCL6 can be coexpressed with T-bet, GATA3, or RORγt, but BCL6 modifies many of their effector lineage functions. The fact that there is an incomplete separation between Tfh and other T-effector lineages might suggest that the Tfh phenotype is generated in parallel with T-effector cell differentiation and in response to inflammatory cues present in lymph nodes.

Tfh differentiation is promoted by signals through ICOS, OX-40, and the presence of IL-6, IL-12, and IL-21. Tfh cells upregulate CXCR5, which enables homing to the interface between the T-cell and B-cell areas of lymph nodes, where they interact with recently activated B cells via antigen presentation. The B cells provide signals to the Tfh cell through ICOS and IL-6, while the Tfh cells provide CD40L and IL-21 to the B cells. Through successive interactions with B cells, Tfh cells promote the survival, proliferation, Ig class switching, affinity maturation and differentiation of B cells into memory B cells, and antibody-producing plasma cells through the germinal center reaction.[28]

Mutations affecting Tfh-associated pathways have been identified in B-cell immunodeficiency disorders. Common variable immune deficiency (CVID) is often linked to mutations in the ICOS pathway. Recurrent infections in patients with CVID are associated with hypogammaglobulinemia and reduced numbers of Tfh cells. X-linked lymphoproliferative disease type 1 (XLP1) has been associated with mutations in the *SH2D1A* gene. Interestingly, patients with XLP1 have defects in B-cell germinal center formation and decreased memory B-cell numbers but have normal Tfh numbers. XLP1 Tfh cells reduced the ability to support B-cell function.[29] Tfh cells have been implicated in autoimmune diseases, such as myasthenia gravis, Graves disease, Hashimoto thyroiditis, lupus, Sjögren syndrome, and rheumatoid arthritis, where dysregulated antibody responses are thought to contribute to disease progression. In addition, therapies to target or passively transfer Tfh cells are being studied for treatment of autoimmune disease and some forms of lymphoma.[29]

MEMORY T CELLS

Following an immune response, small numbers of T cells persist over the long term and are called *memory T cells*. Memory T cells protect the host from reinfection by the same microorganisms. Protection correlates with the number of specific memory cells present in the host. Memory cells accelerate responses against repeat antigen for several reasons. First, memory cells are maintained at higher frequencies than are naïve cells. Second, a large portion of memory T cells are maintained in peripheral tissues, allowing for rapid local responses to infection. Third, memory T cells proliferate and produce cytokines in response to stimulation with lower doses of antigen, less costimulation, and much faster kinetics compared with naïve T cells. Fourth, memory cells enhance can promote APC function and produce cytokines that promote the effector differentiation of naïve T cells.

Diverse populations of memory T cells can be defined by surface marker expression, localization, and patterns of migration. Two populations of memory T cells are identified on the basis of their reexpression of the lymph node homing receptors CCR7 and L-selectin following activation. Central memory (T_{cm}) cells are identified as CD45RA$^-$ CD45RO$^+$, CCR7$^+$, CD62L$^+$, and stem cell memory cells (T_{scm}) are identified by the markers CD45RA$^+$ CD45RO$^-$, CCR7$^+$, CD62L$^+$, CD95$^+$.[30] By expressing CCR7 and CD62L, these memory T cells permit recirculation through secondary lymphatic organs, allowing interactions with DCs from diverse tissue sites effectively expanding the area of immune surveillance. T_{cm} and T_{scm} produce IL-2 and rapidly enter the cell cycle upon activation. T_{cm} and T_{scm} are able to acquire effector phenotypes, suggesting that they function as a precursor pool for subsequent immune responses.

Two additional memory cell subsets have been described on the basis of their ability to migrate through peripheral tissues. Effector memory T cells (T_{em}) are identified by the expression pattern: CD45RO$^+$, CCR7$^-$, CD62L$^-$. These cells express chemokine receptors, selectins, and integrins to facilitate their recirculation through peripheral tissues. T_{em} cells have previously undergone effector differentiation, and upon stimulation they rapidly express restricted cytokine patterns reflecting their prior phenotype. Additional populations of nonmigratory memory T cells have been identified. These are known as tissue resident memory T cells (T_{rm}). T_{rm} cells are identified by the markers CD45RO$^+$, CCR7$^-$, CD69$^+$. They have been identified in close proximity to epithelial sites in multiple tissues. T_{rm} cells provide rapid and highly localized cytokine production in response to infection. Thus T_{em} and T_{rm} cells bolster immunity by providing immune surveillance at distinct sites.[31]

The processes underlying memory T-cell differentiation are incompletely understood. A number of signaling pathways and APC populations have been proposed to be involved in memory differentiation. It is clear that T_{cm} and T_{em} cells are generated early in the T-cell response. The pathways that lead to the differentiation of Th1, Th2, Th17, and Tfh effector cells are also important in the generation of T_{em} populations.

GENERAL CONSIDERATIONS IN EFFECTOR T-CELL DIFFERENTIATION

Several key concepts of Th maturation should be discussed with regard to the cellular differentiation of CD4 T cells. First and foremost, all T-effector, memory, and adaptive Tregs are thought to arise from naïve thymic emigrant precursors. In addition, T-cell phenotype specification is regulated independent of TCR specificity.

Second, differentiation to a specific effector phenotype is self-promoting and inhibitory to other phenotypes through production of effector cytokines (see Fig. 16.3). Early in Th1 or Th2 differentiation, for example, expression of the transcription factors T-bet or GATA3, promote proeffector differentiation while suppressing alternative cell fates. Once established, T-cell phenotypes are maintained through active signaling and chromosomal modification.

Third, polarization is incomplete at the population level. Following exposure to antigen, responding T cells may collectively possess multiple phenotypes. Heterogeneity of Th phenotypes is a common feature of some diseases, such as asthma, where Th2, Th9, and Th17 responses are thought to contribute to disease; in psoriasis, where Th1, Th17, and Th22 cell types are detectable; and in multiple sclerosis, where Th1, Th17, and Th22 cells are thought to mediate inflammation. A portion of the T cells may persist in a naïve phenotype or as uncommitted T_{cm} or T_{scm} that may give rise to T cells of a different phenotype upon antigen reexposure.

Fourth, there is limited plasticity in T-cell effector phenotypes. Effector phenotypes are reinforced by "master" transcription factors and epigenetic modifications (Fig. 16.4). For example, established Th1 cells do not express Th2 cytokines despite extended culture in Th2-skewing conditions. There is, however, growing evidence of interconversion between certain T-helper subsets.[32] Experiments tracing T-cell lineage commitment have shown that T cells that once expressed FoxP3 could become effector cells. It is not clear if plasticity is occurs between all phenotypes or if phenotype conversion can occur at any stages of phenotype commitment (see Fig. 16.4).

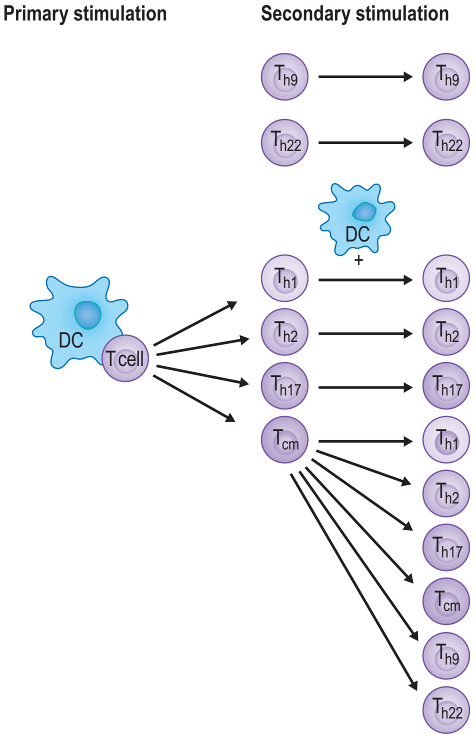


FIG 16.4 Mechanism of T-Helper (Th) Phenotype Shift. Under appropriate conditions, T-cell activation may result in daughter cells adopting one of several T-effector phenotypes. Within a polarized population of Th cells, a subset of undifferentiated T lymphocytes may exist (here referred to in the general category as T central memory [T_{cm}]). Fully committed Th1, Th2, or Th17 cells upon restimulation will produce Th1, Th2, or Th17, respectively. T_{cm} or other uncommitted cell types retain the potential to differentiate into T effectors of any phenotype, depending on the context of the secondary stimulation.

TERMINATION OF T-CELL RESPONSES

CLINICAL RELEVANCE

Diseases Associated With CD4 T-Effector Responses

Th1

- Multiple sclerosis
- Psoriasis
- Type 1 diabetes mellitus
- Tuberculoid leprosy
- Transplant rejection

Th2

- Allergy and asthma
- Helminthic infections
- Lepromatous leprosy

Th9

- Asthma
- Helminthic infections

Th17

- Asthma
- Multiple sclerosis
- Psoriasis
- Rheumatoid arthritis
- Transplant rejection

Th22

- Psoriasis
- Wound healing

As a result of the potential for unregulated inflammation, retention of activated effector T cells present three risks to the organism. First, with each succeeding infection, the body would be increasingly burdened by the energy requirements of maintaining larger T-cell numbers. Second, there would be loss of TCR diversity resulting from the overrepresentation of clonally expanded T cells. Third, there is an increased risk of damage to self by large populations of activated effector T cells. Thus a return to homeostasis is required at the end of an immune response through reductions in T-cell numbers and activity. Homeostasis is restored through the loss of the initiating stimulus and by active control processes, including cell death, inhibitory signaling pathways, and antiinflammatory cytokines.

CELL DEATH PATHWAYS IN T-CELL HOMEOSTASIS

Several cell death pathways regulate the processes of activation-induced cell death (AICD; Chapter 13). The primary pathway involved in AICD involve Fas (CD95) and Fas ligand (FasL, CD95L). Fas and FasL are expressed following T-cell activation. Fas-mediated apoptosis can be stimulated by FasL on neighboring cells ("death signal") or on the same cell ("suicide signal"). Mutations in Fas and FasL result in the accumulation of activated lymphocytes and autoimmunity. Another TNF family member, TNF-related apoptosis inducing ligand (TRAIL), has also been associated with AICD. Mice lacking TRAIL have enhanced susceptibility to autoimmune disease. Some evidence suggests that TRAIL regulates AICD primarily in Th2 cells.[33]

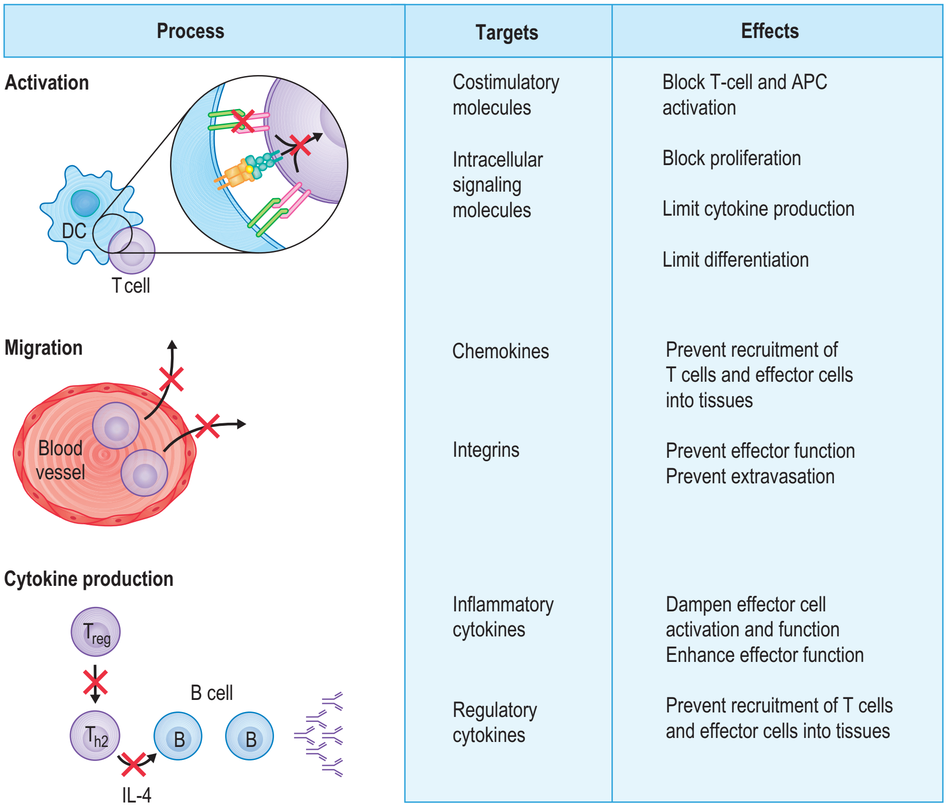
ACTION OF INHIBITORY RECEPTORS

T cells express several Ig family transmembrane proteins containing immune tyrosine inhibitory motifs (ITIMs). Two of these, CTLA4 and PD-1, play important roles in terminating T-cell responses *in vivo.* The first, CTLA4, is expressed by T cells following activation and regulates T-cell activity in two ways: first, CTLA4 binds to B7 with high affinity and competes with CD28 costimulation. Second, CTLA4 inhibits proximal TCR signaling by recruiting the phosphatases SHP-2 and PP2A. Mice deficient in CTLA4 develop a severe CD28-dependent lymphoproliferative disease, tissue infiltration, and early death.[34] PD-1 interacts with two ligands, PD-L1 and PD-L2, which are differentially expressed in immune and peripheral tissues. PD-1 engagement limits proximal TCR signals by recruiting SHP-2, thereby suppressing activation, cytokine production, and proliferation. Loss of PD-1 has been associated with autoimmune cardiomyopathy, arthritis, and a lupus-like glomerulonephritis disease.[35]

CYTOKINE-MEDIATED INHIBITION

Cytokines function to terminate immune responses in three basic methods: *(i)* by loss or withdrawal of growth factors, *(ii)* by inducing cell death, and *(iii)* by direct antiinflammatory properties (Chapter 9). IL-2 functions as a growth factor in early immune responses. Loss of IL-2 signals through decreased production or loss of receptor expression deprives the cell of survival signals, resulting in AICD. Similarly, loss of IL-7 or IL-7 receptor results in increased apoptosis among naive and memory T-cell populations.[36] As discussed in Chapter 13, cell death by growth factor withdrawal results from activation of the mitochondrial pathway of apoptosis.

Cytokines can suppress T-cell responses. This function of cytokines is exemplified by IL-10, IL-27, TGF-β, and TNF-α. IL-10 is produced by adaptive Tregs and Tr1 cells, as well as B cells, monocytes, and macrophages. IL-10 suppresses inflammation by inhibiting macrophage activation, downregulating chemokine production, and suppressing costimulatory molecule expression by APCs. IL-10–deficient mice develop inflammatory bowel disease caused by dysregulated Th1 responses. Mice genetically deficient for the IL-27 receptor develop exaggerated CD4 T-cell responses and inflammatory diseases. These effects of IL-27 can be attributed to its role in suppressing Th1 responses and CD4 T cell proliferation.[37] TGF-β_1 is another antiinflammatory cytokine that is produced by subsets of Tregs and a variety of other cell types throughout the body. Unlike IL-10, which targets APCs, TGF-β_1 directly inhibits T-cell proliferation and Th1 differentiation.[38] Mice deficient in TGF-β_1 have multitissue infiltration by activated lymphocytes and macrophages. TNF-α is well characterized as possessing proapoptotic function. TNFR1 possesses a death domain, and binding of TNF-α results in the activation of the caspase pathway and apoptosis.



Process	Targets	Effects
Activation	Costimulatory molecules	Block T-cell and APC activation
	Intracellular signaling molecules	Block proliferation Limit cytokine production Limit differentiation
Migration	Chemokines	Prevent recruitment of T cells and effector cells into tissues
	Integrins	Prevent effector function Prevent extravasation
Cytokine production	Inflammatory cytokines	Dampen effector cell activation and function Enhance effector function
	Regulatory cytokines	Prevent recruitment of T cells and effector cells into tissues

FIG 16.5 Therapeutic Regulation of Inflammation. Several techniques have been employed to reduce T-cell responses in cases of inflammation. Among these are inhibiting T-cell activation and differentiation by blocking costimulatory interactions with antigen-presenting cells (APCs) *(top)*; inhibiting T-cell and effector-cell trafficking by blocking molecules required for chemotaxis *(middle)*; and limiting the effect of T-helper responses by decreasing the availability of inflammatory cytokines *(bottom)*.

SUMMARY OF THERAPEUTIC REGULATION OF T-CELL RESPONSES FOR TREATMENT OF IMMUNE-MEDIATED DISEASES

ON THE HORIZON

- Development of effective techniques to identify antigens targeted in organ-specific autoimmune and allergic diseases and in transplant rejection
- Development and implementation of an effective tolerance-inducing strategy capable of regulating T-cell responses in an antigen-specific manner
- Identification of tolerance-specific biomarkers or assays to clearly delineate successful therapy of inflammatory diseases

T cell–mediated inflammation is central to immune-mediated diseases. Many of the most effective techniques to control inflammation are aimed at the processes of T-cell activation, differentiation, trafficking, or effector functions (Fig. 16.5). Among many potential approaches currently being employed or in development, include the following: *(i)* blockade of appropriate costimulatory signals or intracellular signaling molecules to prevent T-cell activation (*i.e.*, FK504) or to promote ongoing effector inflammatory responses (*i.e.*, anti-PD-L1); *(ii)* preventing recruitment of effector T cells into inflammatory sites by targeting pathways required for cellular recruitment (*i.e.*, VLA-4); *(iii)* neutralizing the cytokines that promote T-effector differentiation (*e.g.*, IL-6, IL-12, or IL-23); *(iv)* blocking proinflammatory cytokines (*e.g.*, IFN-γ, IL-17, GM-CSF, or TNF-α) and other effector molecules (*e.g.*, iNOS, reactive oxygen intermediates, etc.) involved in the effector limb of destructive inflammatory processes; *(v)* the use of regulatory cytokines, such as TGF-β, IL-4, or IL-10 to regulate proinflammatory responses; and *(vi) in vivo* induction of Treg activity using tolerance strategies or *ex vivo* expansion and administration of protective or regulatory T-cell subsets.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Fischer A. Recent advances in understanding the pathophysiology of primary T cell immunodeficiencies. Trends Mol Med 2015;21:408–16.
2. Bussone G, Mouthon L. Autoimmune manifestations in primary immune deficiencies. Autoimmun Rev 2009;8:332–6.
3. Smith-Garvin JE, Koretzky GA, Jordan MS. T cell activation. Annu Rev Immunol 2009;27:591–619.
4. Friedl P, den Boer AT, Gunzer M. Tuning immune responses: diversity and adaptation of the immunological synapse. Nat Rev Immunol 2005;5:532–45.
5. Watts C, West MA, Zaru R. TLR signalling regulated antigen presentation in dendritic cells. Curr Opin Immunol 2010;22:124–30.
6. Randolph GJ, Angeli V, Swartz MA. Dendritic-cell trafficking to lymph nodes through lymphatic vessels. Nat Rev Immunol 2005;5:617–28.
7. Mempel TR, Henrickson SE, Von Andrian UH. T-cell priming by dendritic cells in lymph nodes occurs in three distinct phases. Nature 2004;427:154–9.
8. Manicassamy S, Pulendran B. Dendritic cell control of tolerogenic responses. Immunol Rev 2011;241:206–27.
9. Thomson AW, Turnquist HR, Raimondi G. Immunoregulatory functions of mTOR inhibition. Nat Rev Immunol 2009;9:324–37.
10. Quezada SA, Peggs KS, Simpson TR, et al. Shifting the equilibrium in cancer immunoediting: from tumor tolerance to eradication. Immunol Rev 2011;241:104–18.
11. Blattman JN, Antia R, Sourdive DJ, et al. Estimating the precursor frequency of naive antigen-specific CD8 T cells. J Exp Med 2002;195:657–64.
12. Ley K, Kansas GS. Selectins in T-cell recruitment to non-lymphoid tissues and sites of inflammation. Nat Rev Immunol 2004;4:325–35.
13. Cyster JG. Chemokines, sphingosine-1-phosphate, and cell migration in secondary lymphoid organs. Annu Rev Immunol 2005;23:127–59.
14. Rebenko-Moll NM, Liu L, Cardona A, et al. Chemokines, mononuclear cells and the nervous system: heaven (or hell) is in the details. Curr Opin Immunol 2006;18:683–9.
15. Takada K, Jameson SC. Naive T cell homeostasis: from awareness of space to a sense of place. Nat Rev Immunol 2009;9:823–32.
16. Mosmann TR, Cherwinski H, Bond MW, et al. Two types of murine helper T cell clone. I. Definition according to profiles of lymphokine activities and secreted proteins. J Immunol 1986;136:2348–57.
17. Zhu J, Paul WE. Peripheral CD4+ T-cell differentiation regulated by networks of cytokines and transcription factors. Immunol Rev 2010;238:247–62.
18. Coffman RL, Lebman DA, Rothman P. Mechanism and regulation of immunoglobulin isotype switching. Adv Immunol 1993;54:229–70.
19. Paul WE, Zhu J. How are T(H)2-type immune responses initiated and amplified? Nat Rev Immunol 2010;10:225–35.
20. Leonard WJ, Spolski R. Interleukin-21: a modulator of lymphoid proliferation, apoptosis and differentiation. Nat Rev Immunol 2005;5:688–98.
21. Dong C. Diversification of T-helper-cell lineages: finding the family root of IL-17-producing cells. Nat Rev Immunol 2006;6:329–33.
22. Soroosh P, Doherty TA. Th9 and allergic disease. Immunology 2009;127:450–8.
23. Eyerich S, Eyerich K, Pennino D, et al. Th22 cells represent a distinct human T cell subset involved in epidermal immunity and remodeling. J Clin Invest 2009;119:3573–85.
24. Masteller EL, Tang Q, Bluestone JA. Antigen-specific regulatory T cells — ex vivo expansion and therapeutic potential. Semin Immunol 2006;18:103–10.
25. Roncarolo MG, Gregori S, Battaglia M, et al. Interleukin-10-secreting type 1 regulatory T cells in rodents and humans. Immunol Rev 2006;212:28–50.
26. Byrne P, McGuirk P, Todryk S, et al. Depletion of NK cells results in disseminating lethal infection with Bordetella pertussis associated with a reduction of antigen-specific Th1 and enhancement of Th2, but not Tr1 cells. Eur J Immunol 2004;34:2579–88.
27. Nutt SL, Tarlinton DM. Germinal center B and follicular helper T cells: siblings, cousins or just good friends? Nat Immunol 2011;12:472–7.
28. Crotty S. T follicular helper cell differentiation, function, and roles in disease. Immunity 2014;41:529–42.
29. Ma CS, Uzel G, Tangye SG. Human T follicular helper cells in primary immunodeficiencies. Curr Opin Pediatr 2014;26:720–6.
30. Lanzavecchia A, Sallusto F. Understanding the generation and function of memory T cell subsets. Curr Opin Immunol 2005;17:326–32.
31. Carbone FR. Tissue-resident memory T cells and fixed immune surveillance in nonlymphoid organs. J Immunol 2015;195:17–22.
32. O'Shea JJ, Paul WE. Mechanisms underlying lineage commitment and plasticity of helper CD4+ T cells. Science 2010;327:1098–102.
33. Roberts AI, Devadas S, Zhang X, et al. The role of activation-induced cell death in the differentiation of T-helper-cell subsets. Immunol Res 2003;28:285–93.
34. Chikuma S, Bluestone JA. CTLA-4 and tolerance: the biochemical point of view. Immunol Res 2003;28:241–53.
35. Okazaki T, Honjo T. The PD-1-PD-L pathway in immunological tolerance. Trends Immunol 2006;27:195–201.
36. Rochman Y, Spolski R, Leonard WJ. New insights into the regulation of T cells by gamma(c) family cytokines. Nat Rev Immunol 2009;9:480–90.
37. Hunter CA. New IL-12-family members: IL-23 and IL-27, cytokines with divergent functions. Nat Rev Immunol 2005;5:521–31.
38. Li MO, Wan YY, Sanjabi S, et al. Transforming growth factor-beta regulation of immune responses. Annu Rev Immunol 2006;24: 99–146.

MULTIPLE-CHOICE QUESTIONS

1. Which of the following is the earliest manifestation of a T-cell response?
   A. Cytokine release
   B. Activation and proliferation
   C. T-cell help
   D. T-cell migration
   E. Stimulation of inflammation

2. Which of the following is NOT true about regulatory T cells (Tregs)?
   A. Tregs express the transcription factor FoxP3.
   B. Tregs control immune responses by lysing effector CD4 T cells.
   C. Tregs are defective in patients with autoimmune polyendocrinopathy–candidiasis–ectodermal dystrophy (APECED).
   D. Tregs can be generated outside of the thymus.
   E. Tregs help promote homeostasis.

3. A patient arrives in the clinic with a history of recurrent infection by *Staphylococcus* bacteria. Laboratory tests indicate that the patient has low serum immunoglobulin G (IgG) levels and reduced numbers of memory B cells in blood. What T-cell population is likely impaired?
   A. CD8 T cells
   B. T-helper (Th)1 effector T cells
   C. Th2 effector T cells
   D. Regulatory T cells (Tregs)
   E. T-follicular helper (Tfh) cells

17

Cytotoxic T Lymphocytes and Natural Killer Cells

Stephen L. Nutt, Nicholas D. Huntington

Cytotoxic T lymphocytes (CTLs) and natural killer (NK) cells represent two distinct but functionally related lineages that contribute to pathogen and tumor immunity. Although the approaches by which CTLs and NK cells kill their target cells and produce immunomodulatory cytokines are quite similar, the mechanisms by which they recognize their targets are distinctly different. CTLs are CD8 T cells (Chapter 8) that recognize targets via the interaction of a diverse repertoire of polyclonally rearranged T-cell receptors (TCRs) (Chapter 4) with a peptide– major histocompatibility complex (MHC) class I complex (Chapter 6) and are components of the adaptive immune response. MHC class I molecules are expressed on virtually all cells in the body and allow CTLs to scan the tissues for cells expressing foreign or cancer-associated peptides. In contrast, NK cells are members of the innate immune system (Chapter 3) and use an array of invariant activating and inhibitory receptors to control their activity and specificity.[1] These fundamentally distinct approaches to the recognition of antigen allow for complementary functions, with CTLs being specialized in detecting cancerous cells or those infected with intracellular pathogens, such as viruses, whereas a prominent function of NK cells is to eliminate those cells where the pathogen or oncogene has impaired MHC class I expression and/or has induced the expression of stress ligands. As one of the principal immune-evasion mechanisms of viruses and tumors is suppression of MHC class I expression, NK cells provide a key line of defense against this strategy.

CLINICAL RELEVANCE

Functions of Cytotoxic T Lymphocytes (CTLs) and/or Natural Killer (NK) Cells

- Protective functions include:
- Host defense against:
  - Viruses, including human immunodeficiency virus (HIV), Epstein-Barr virus (EBV), pox virus, and cytomegalovirus (CMV)
  - Bacteria, including *Listeria monocytogenes*
  - Parasites, including *Plasmodium falciparum* and *Toxoplasma gondii*
  - Primary and metastatic tumors
- Positive regulation of:
  - Graft-versus-leukemia (GvL) effect
  - Placental vascularization by uterine NK cells
- Uncontrolled cytotoxicity contributes to
- Some autoimmune disease, including diabetes and rheumatoid arthritis
  - Hypersensitivity reactions
  - Graft-versus-host disease
  - Transplant rejection

The importance of the lytic function of CTLs and NK cells has been demonstrated in animal models as well as in patients with defective cytotoxicity. A number of recessive genetic syndromes that affect cytotoxic function have been reported, including familial hemophagocytic lymphohistiocytosis (FHL), which results from mutations in the perforin gene.[2] Patients with FHL present with severe immunodeficiency that is often associated with uncontrolled viral infections, including cytomegalovirus (CMV), herpes simplex virus (HSV), and Epstein–Barr virus (EBV) infections. Similarly, mice genetically deficient or depleted for CTLs and NK cells are overtly susceptible to viral pathogens and display impaired immunosurveillance.[2]

With its potent ability to control pathogen-infected and malignant cells, it is not surprising that modulation of cytolytic activity is an aim of many immunotherapies (see Part 10). These strategies involve either dampening of CTL function in such situations as transplantation or autoimmunity or enhancement of CTL and NK-cell function via vaccination, blocking antibodies to inhibitory receptors, or cytokine therapy. Tight controls need to be maintained over these effector cells, as deregulated CTL activity can promote autoimmune diseases, hypersensitivity reactions, graft-versus-host disease (GvHD) and transplant rejection. To maintain the discrimination between killing unwanted or infected cells and not killing healthy neighboring cells, numerous layers of regulation operate to control cytotoxic functions.

EFFECTOR FUNCTIONS/MECHANISMS

KEY CONCEPTS

Cytotoxic T Lymphocyte (CTLs) and/or Natural Killer (NK) Cell Effector Mechanisms

- Cytotoxicity
  - Killing by the perforin/granzyme pathway
  - Death receptor–mediated apoptosis, including Fas and tumor necrosis factor (TNF)-related apoptosis–inducing ligand (TRAIL)
- Immune modulation
  - Inflammatory cytokine production, including interferon (IFN)-γ and TNF
  - Chemokine secretion
  - Immunomodulatory cytokines, including interleukin-10 (IL-10) and granulocyte macrophage–colony-stimulating factor (GM-CSF)

Cytotoxicity

Cytotoxic cells kill their targets via two major pathways: perforin/granzyme–mediated lysis and death receptor–induced apoptosis.

https://archive.org/details/namdhari

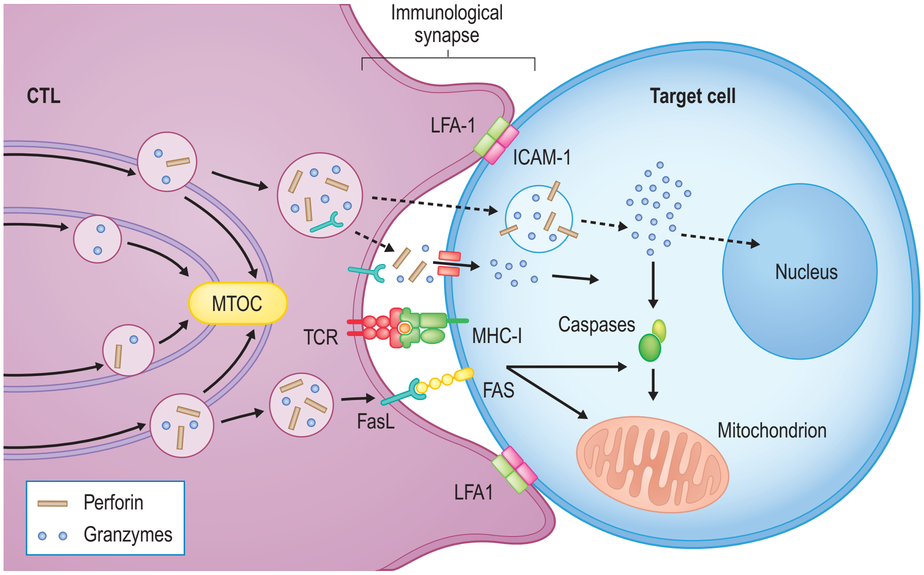


FIG 17.1 Mechanisms of Cytotoxic T Lymphocyte (CTL)–Induced Cell Death. The CTL recognizes its target via the interaction of the T-cell receptor (TCR) and the peptide–major histocompatibility complex (MHC) class I complex on the target cell. TCR signaling induces the formation of an immunological synapse that is stabilized by the binding of leukocyte function–associated antigen-1 (LFA-1) to the intercellular adhesion molecule (ICAM) on the target cell. Lytic granules containing perforin, granzymes, and FasL are polarized along microtubules and move toward the microtubule-organizing center (MTOC). Lytic granules are then secreted into the synapse, allowing perforin to form pores in the target cell membrane, which facilitates granzyme access to the target cells. Granzymes induce apoptosis by caspase-dependent and independent pathways that result in DNA cleavage and mitochondrial damage. Membrane-bound FasL can bind to its receptor on target cells and induce apoptosis through an independent pathway.

Both require intimate contact between the lytic cell and its target (Fig. 17.1).[2] Although the processes are similar for CTLs and NK cells, CTL lytic activity is acquired only after activation and differentiation, whereas NK cells can spontaneously kill target cells without prestimulation. Despite this, NK-cell killing is significantly increased by prior activation by cytokines or inflammatory signals. Both cell types also produce cytokines that further enhance the immune response, most notably interferon-γ (IFN-γ).

Perforin–Granzyme Pathway

Perforin is a membrane-disrupting protein, which, together with a family of serine proteases (granzymes), forms the bulk of lytic granules. The process of lysis has been most extensively studied in CTLs, where, upon interaction between the TCR and an appropriate MHC class I peptide, a synaptic complex forms between the CTL and its target.[2] Lytic granules can then be observed moving along a microtubule network toward the microtubule-organizing center that localizes at the synapse (see Fig. 17.1). This process allows the polarized secretion of lytic granules precisely at the CTL–target cell interface. Perforin forms a pore that disrupts the target cell membrane, including either the plasma membrane or the lysosomal membrane. Once inside the target cell, it is granzymes that are the initiators of cell death. Granzymes function directly by cleaving substrates, such as nuclear proteins and DNA, or indirectly via initiating a protease cascade. One important substrate of granzymes is the proapoptotic protein BID (BH3-interacting domain death agonist), which induces cell death via mitochondrial mediators, such as cytochrome c.

Death Receptor–Induced Apoptosis

Cytotoxic cells also have a receptor-based system to induce apoptosis of target cells (Chapter 13). This pathway uses members of the tumor necrosis factor receptor (TNFR) superfamily that are expressed on the target cells. These receptors have an intracellular signaling motif, called the *death domain,* which recruits molecules that transduce the death signal, such as FADD (Fas-associated death domain). The two most prominent apoptosis-inducing TNFR family members are Fas (CD95)[3] and TRAIL (TNF-related apoptosis–inducing ligand).[4] Fas is expressed on a wide variety of tissues, whereas Fas ligand (FasL) expression is restricted to activated CTLs and NK cells, where it is stored in lytic granules and, upon activation, released to the effector cell membrane. The Fas/FasL pathway is important in controlling T-cell numbers through activation-induced cell death (AICD), as well as in the rejection of some tumors. Cytotoxic cells also express TRAIL, which upon binding to TRAIL receptors, induces apoptosis in a wider selection of targets.[4] Of particular therapeutic interest is the fact that tumor cells are often exquisitely sensitive to TRAIL.

Cytokines

Antigen-stimulated CTLs and activated NK cells modulate the immune response by their ability to produce a variety of cytokines, most notably IFN-γ and TNF. These potent inflammatory cytokines activate macrophages and lymphocytes at the site of infection. IFN-γ helps establish a T-helper cell-1 (Th1) response (Chapter 16) and further stimulates differentiated CTLs. NK cells are also a potent source of a diverse range of cytokines, including granulocyte macrophage–colony-stimulating factor (GM-CSF), interleukin-5 (IL-5), IL-10, and IL-13 (Chapter 9). The capacity to secrete a broad spectrum of cytokines provides NK cells with a wide range of regulatory capabilities. CTLs and NK cells are capable of secreting a number of chemokines (Chapter 10), including chemokine ligand 3 (CCL3), CCL4, and CCL5. These chemokines help recruit additional lymphocytes into the immune reaction.

CYTOTOXIC T CELLS

The Development and Tissue Distribution of CTLs

CD8 T lymphocytes develop in the thymus, where they are selected for their ability to recognize nonself peptides in the context of MHC class I molecules (Chapter 8). Upon thymic export, these cells acquire a quiescent state and are referred to as "naïve" cells. Naïve CD8 T cells circulate between peripheral lymphoid organs, such as the spleen and lymph nodes, via the arterial and lymphatic systems. The tissue distribution of lymphocytes is determined by targeting proteins, which can be divided into three categories: selectins, chemokine receptors, and integrins.[5] Naïve and activated CD8 T cells display distinct sets of these targeting proteins, allowing for the differential homing abilities of these cells (Chapter 11). Naïve CD8 T cells express high levels of the lymph node homing receptor L-selectin (CD62L) and CCR7, a chemokine receptor that recognizes CCL19 and CCL21, which are produced in the T-cell areas of secondary lymphoid organs (Table 17.1; Chapter 10).[5] Here, naïve T cells interact with antigen-presenting cells (APCs), in particular dendritic cells (DCs). If a naïve CD8 T cell does not encounter its specific antigen, it leaves the lymph node.[5] If, however, a CD8 T cell encounters the correctly presented peptide–MHC I complex, a dramatic change in its localization and homing properties ensues. These cells shut down their egress program and undergo multiple rounds of proliferation to become activated CTLs. After the proliferative phase, the CTLs then reacquire egress capacity and travel via the circulation to nonlymphoid sites, where they tether to endothelial cells and extravasate into tissue. This transmigration occurs in both inflamed and noninflamed sites, such as skin, the gut, or the lung. Many of the effector memory CTLs are retained in nonlymphoid tissues, where they are poised to respond rapidly should the antigen be encountered again. The distinct types of memory T cells are discussed below.

THE CTL RESPONSE

The CTL response to an acute infection consists of three phases: first, the initial activation and proliferation of the CTL; second, the contraction of effector populations; and third, the long-term maintenance of memory cells.

Initial Activation

Naïve T cells constantly circulate through secondary lymphoid organs, where antigen encounter occurs. For a CTL response, antigen is brought to the lymph node via the lymphatic system by APC (Fig. 17.2). These APCs are typically DCs that mature after antigen acquisition in nonlymphoid tissues and migrate to the lymph node. These antigens are recognized only when complexed to MHC molecules. APCs efficiently degrade self or foreign (pathogen-derived) proteins into shorter fragments (generally 8–10 amino acids in length) by the action of proteases in the cytosol. They are then transported into the lumen of the endoplasmic reticulum, where they are loaded onto newly synthesized MHC class I molecules for presentation at the cell surface (Chapter 6). This then enables the APCs to communicate with the antigen-specific CD8 T cells via interactions between the TCR and MHC molecules.

In the lymph node, CTLs scan the APCs for the presence of antigenic peptides complexed with the MHC class I molecules, a process termed *immune surveillance.* In the absence of a specific recognition by the TCR, the encounter is only transient, and the T cell continues on to another APC to repeat the process. If the MHC class I–peptide complex is bound by the TCR and initiates signaling, a more lasting interaction occurs.

TCR activation promotes polarization of the T cell and formation of the "immunological synapse."[6] The immunological synapse is a highly structured body that functions to concentrate TCR signaling in a defined area. It is associated with the selective recruitment of signaling molecules and exclusion of negative regulators. The synapse is stabilized by a ring of adhesion molecules, including, for example, LFA-1, which binds to ICAM1 on the APC (see Fig. 17.1). For a T cell to become fully active, costimulation through a second signaling pathway is required.[6] Many costimulators that have been identified share the common characteristic of being transmembrane receptors, often of the TNFR superfamily, that bind transmembrane ligands on the APC. The most important costimulator, CD28, binds the ligands CD80 (B7.1) and CD86 (B7.2), both of which are expressed on activated APCs. Costimulation results in the clonal expansion

TABLE 17.1 **Properties of Cytotoxic T Lymphocyte (CTL) Populations**

Marker	Naïve CD8 T Cell	Effector CTL	Effector Memory (T_{EM})	Central Memory (T_{CM})	Tissue-Resident Memory (T_{RM})
CD69	–	++	–	–	++
CD62L	++	–	–	++	–
CD44	+	+++	+++	+++	+++
CCR7	+	–	–	+	–
IL7R (CD127)	++	+	+	++	+
IL2Rβ (CD122)	+	+	++	+++	+
Main tissue distribution	Lymph nodes, spleen, blood	Lymph nodes, spleen, blood, nonlymphoid tissue (*e.g.*, lung, liver)	Nonlymphoid tissue (*e.g.*, lung, liver), spleen	Lymph nodes, spleen, blood	Nonlymphoid tissue (*e.g.*, lung, liver, skin)
Cytotoxic function	–	+++	++	–	++
IFN-γ	–	+++	+++	+	+++

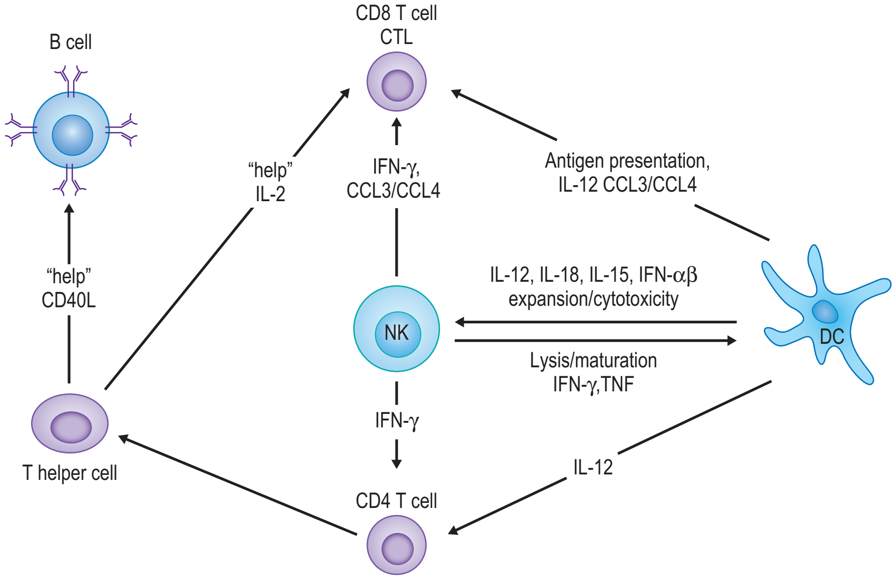


FIG 17.2 Cellular Interactions During an Immune Response in the Lymph Node. After their encounter with antigen, dendritic cells (DCs) move to the draining lymph node, where they initiate an antigen-specific immune response. In the very early stages, natural killer (NK) cells are recruited to the lymph node and modulate various aspects of this response. NK cells provide cytokines that induce the maturation of DCs, which enables these cells to efficiently present the antigen to T cells in the context of costimulatory signals. NK cells also have the ability to eliminate immature DCs and to provide "early" interferon (IFN)-γ for the initiation of a T-helper 1 (Th1)-type CD4 T-cell response. Finally, NK cells produce a range of chemokines, most notably chemokine ligand 3 (CCL3) and CCL4, which are crucial for the recruitment of CD8 T cells into the immune response.

of CTLs with the selected antigen specificity. The expression of CD80/86 is tightly regulated. High-level expression occurs only after an APC receives activation signals, such as inflammatory cytokines, or components of the pathogens such as lipopolysaccharide. Costimulation via CD28 is most critical during the initiation of the immune response, as it promotes IL-2 production, which, in turn, supports the development of effector T cells. Naïve T cells that receive TCR stimulation in the absence of costimulatory signals can become nonresponsive to antigen, a state termed *anergy.*

Cross-Presentation and Priming

After it was established that only direct interaction with an APC and appropriate costimulation led to full CTL activity, a problem arose in explaining the mechanism by which antigens in non-APCs were recognized by CTLs. This dilemma was resolved by the discovery that there are two distinct mechanisms by which CTLs encounter peptide–MHC class I complexes.[7] If the APC expresses the antigen, for example, if a virus infects it, then the APC can process the antigen via the endogenous MHC class I pathway for presentation. The more intriguing situation arises when the APC does not express the antigen. In this case, the APC acquires and processes the antigen via a process termed *cross-presentation* (Chapter 6).[7] Cross-presentation is initiated by the capture of foreign or exogenous antigens by phagosomes. The antigens are then processed by an unusual mechanism that directs the peptides to the MHC class I pathway and presentation on the cell surface. An encounter of a CTL with an antigen processed in this pathway is termed *cross-priming.*[7] This strategy assists CTLs in viral control in two ways. First, cross-priming alleviates the need for DCs to be infected for detection and therefore prevents viruses escaping immune detection by avoiding infection of DCs. Second, viruses often attempt to evade CTLs by downregulating the MHC class I pathway. "Cross-dressing" is a process related to cross-priming but involves the direct exchange of preformed peptide–MHC complexes to uninfected cells. This mechanism acts to significantly amplify the magnitude of priming of T cells, and like cross-presentation, potentially circumvents immune evasion strategies used by pathogens.[7]

The Contraction of Effector Populations

After activation of CTLs in secondary lymphoid organs, an immune response is characterized by the rapid proliferation of antigen-specific cells and their acquisition of effector functions. CTLs proliferate at one of the fastest rates known for mammalian cells, with a cell cycle time of approximately 6 hours. Infection leads to a dramatic increase in the numbers of pathogen-specific CTLs, from almost undetectable initial levels to several million cells in the course of a single week. The magnitude of this response is dependent on the kind of infection and the dose of the antigen and is controlled by the number of precursors recruited into the response.[8] The expansion phase is followed by a contraction of the CTL population that is independent of both the magnitude of the response and clearance of the antigen. This phase is essential to prevent nonspecific tissue damage through uncontrolled cytokine release and cytolytic activity. Contraction also preserves the flexibility of the T-cell response to new infections while memory of previously encountered antigens is maintained. Typically, less than 5% of the expanded antigen-specific population survives in the long-lived memory pool.

The Long-Term Maintenance of Memory Cells

The production of long-lasting antigen-independent memory cells is essential for a rapid response should reinfection occur. CTL memory provides a more vigorous response than the primary challenge for both quantitative and qualitative reasons.[8] Quantitatively, because of the substantial clonal expansion during a primary infection, the precursor frequency of antigen-specific CTLs is vastly higher in immune individuals than in naïve subjects, thus allowing for a stronger response. Qualitatively, memory CTLs exhibit striking efficiency in elaborating the effector functions associated with the rapid production of IFN-γ. This enhanced response is the result of reprogramming of gene expression profiles by epigenetic changes in DNA methylation or chromatin structure.

The CTL memory compartment is composed of three cell types: effector memory (T_{EM}) cells, central memory (T_{CM}) cells, and tissue-resident memory (T_{RM}) cells. These subsets differ in their surface molecule expression and in their ability to exhibit effector functions (see Table 17.1). Like their naïve counterparts, T_{CM} cells express high levels of CD62L and CCR7 and reside primarily in secondary lymphoid organs. T_{CM} cells are capable of prolonged homeostatic self-renewal in the absence of antigen. T_{EM} cells, in contrast, are characterized by low expression of CD62L and CCR7 and are distributed throughout the body, including peripheral tissues, such as the lung and the gut, where they can immediately confront invading pathogens. T_{RM} cells are the most recently identified memory population that resides in peripheral tissues long after an infection is cleared, providing potent early response to reinfection of the same tissue.[5] CD4 T-cell help and cytokines, including IL-15 and IL-7 and their receptors, have been identified as crucial for the survival and maintenance of the memory T-cell pool.[5,8]

CD4 T-CELL HELP

The final player in the initial activation of CTL is the "help" provided by CD4 T cells specific for an antigen linked to the CTL epitope (Chapter 16). The processes by which help is provided are poorly understood. It is likely that cytokines, such as IL-2 and IL-21, are involved and that the CD4 T cells influence both DCs and CTLs.[9,10] These cytokines, provided by the CD4 T cells, promote the survival, proliferation, and programming of memory CTLs. CD4 T-cell–deficient mice have been developed to study these CTL responses. Interestingly, "helpless" CD8 T cells resemble CTLs in chronic infections and cancer, in which the targets are not cleared despite a robust CTL response. PD-1 (*p*rogrammed *d*eath 1), an inhibitory member of the TNFR family, is expressed on both helpless CTLs and on CTL cells during chronic infections. Blocking the interaction of PD-1 with its ligands greatly enhances the numbers and functions of impaired CTLs.[11,12]

DETECTION AND ANALYSIS OF CTL FUNCTION

Much progress in our understanding of the generation of CTLs in an immune response has resulted from the development of several accurate and sensitive assays for CTL function. Traditional CTL assays were performed on bulk populations of effector cells. In these assays, target cells are labeled, often with ^{51}Cr or a nonradioactive dye, and then pulsed with peptide-antigen. Peptide-specific CTLs are incubated with the target cells. Lytic activity is then measured by the release of label into the culture supernatant or the loss of target cell viability directly. Although such an approach provides a powerful quantitative assay for CTL activity, it has the disadvantage of requiring prestimulation of the CTL population for 1–2 weeks to expand the numbers of antigen-specific CTLs to detectable levels. In mouse models, this limitation was overcome by the development of animals that were monoclonal for the TCR. These mouse strains express transgenic TCRs that recognize either MHC class I (CD8)– or MHC class II (CD4)–specific epitopes. Such T cells, which are specific for a single peptide, have proven extremely useful in the study of CTL responses, as antigen-specific cells can be easily detected.

TCR transgenic mouse models do, however, have some limitations, as they do not recapitulate the diversity of the normal immune response and represent an approach that cannot be used in human studies. The development of labeled tetramers of MHC class I–peptide complexes that bind to the endogenous TCR has made a major contribution to overcoming these previous limitations, as it allows for the detection of antigen-specific CTLs within a polyclonal population.[13] This technique is simple and broadly applicable. It can accurately detect rare antigen-specific populations from patient material or animal tissue (Fig. 17.3). The ability to prospectively identify antigen-specific CTLs has been combined with single-cell functional assays, such as ELISpot or intracellular cytokine assays. The finding that the lysosomal compartment proteins lysosomal-associated membrane protein 1 (LAMP-1/2; CD107a/b) can be detected on the cell surface of degranulating CTLs provides another powerful flow cytometry–based single-cell assay to investigate CTL lytic function.

NATURAL KILLER CELLS

Properties of NK Cells

NK cells are a lymphoid lineage, which, because of the absence of antigen–receptor rearrangements, belongs to the innate immune response. The development of B and T lymphocytes has been extensively studied, and anatomical sites and major mediators are well characterized, whereas NK cells have proven more elusive. Part of the problem lies in their relative scarcity, amounting to less than 2% of cells in bone marrow or spleen and to approximately 5% of the lymphocytes in blood. The lack of appropriate markers to distinguish NK cells, coupled with phenotypic heterogeneity between species and within mouse strains, has also complicated their analysis. In recent years, it has become apparent that NK cells develop in bone marrow from a lymphoid progenitor.[14] IL-15 is essential for most aspects of NK-cell biology, including differentiation, survival, proliferation, and activation *in vivo.* The IL-15R consists of three components: the IL-2/15β chain (CD122), a unique IL-15Rα chain, and the common γ chain, which contains the intracellular signaling component of the receptor.[15] The importance of this receptor complex is emphasized by the lack of NK cells in patients with X-linked severe combined immunodeficiency (SCID) and in dogs and mice that have mutations in the common γ-chain or the *IL2B* gene. The IL-15/IL-15R complex does not function in a manner similar to other cytokines that are produced as soluble ligands and bind to their receptor in a paracrine or autocrine manner. IL-15 is virtually undetectable in body fluids or cell culture supernatants despite the ubiquitous distribution of mRNA. The solution to this paradoxical observation is that IL-15 function requires the presence of IL-15Rα in the same cell. The IL-15/

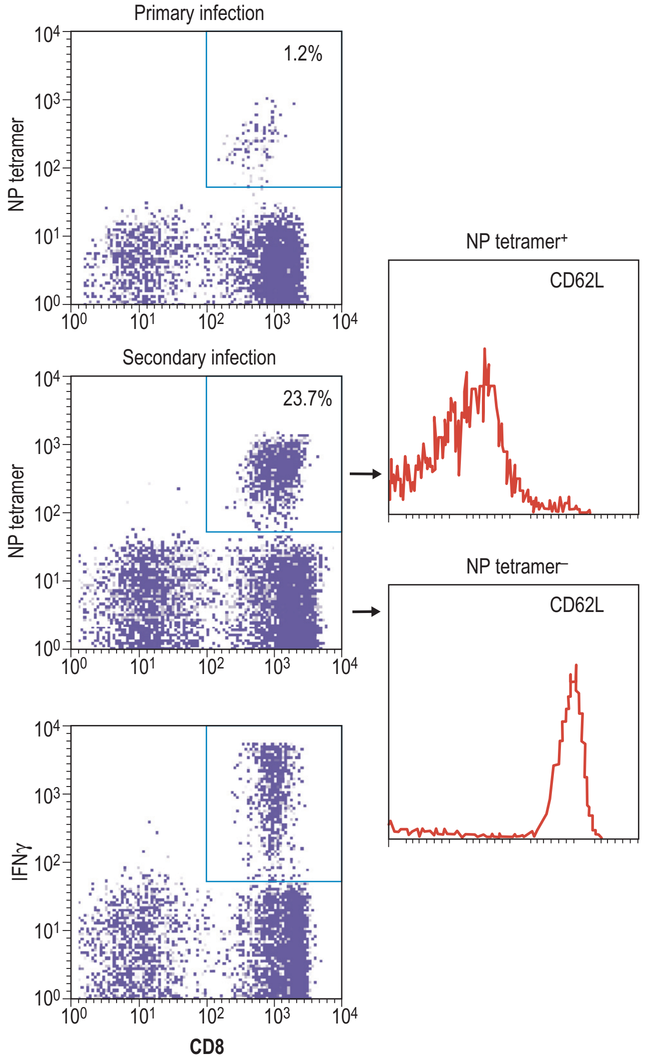


FIG 17.3 **Monitoring a Virus-Specific Cytotoxic T Lymphocyte (CTL) Response.** A fluorescent-labeled tetramer complex comprising a major histocompatibility complex (MHC) class I molecule and a virus-specific nucleoprotein (NP) antigen peptide is used to detect NP-specific CD8 T cells in the spleens of mice infected with influenza virus. During the primary response, virus-specific CTLs expand in the draining lymph node of the lung and later exit to the lung and spleen *(upper left).* After the infection is resolved, a small population of virus-specific memory CTL resides in the spleen. A second infection with influenza results in very high numbers of virus-specific CTLs in the spleen *(middle left)* and the lung. Specific CTLs can be re-stimulated with the NP peptide *in vitro,* inducing interferon (IFN)-γ secretion that is detected by intracellular staining with a labeled anti-IFN-γ antibody *(lower left).* Staining with an antibody to CD62L shows the typical profile of effector CD8 T cells in the NP-tetramer–positive CTLs *(upper right)* and of naïve CD8 T cells in the tetramer negative population *(lower right).*

TABLE 17.2 Human and Mouse Natural Killer (NK) Cell Receptors (Partial List)

Receptor	Species (H, Human; M, Mouse)	Function	Ligands (H, Human; M, Mouse)
LILRB1	H, M	Inhibitory	HLA-class I
KIR2DL1, 2, 3	H	Inhibitory	HLA-C
KIR2DL4	H	Activating	HLA-G
KIR2DL5	H	Inhibitory	?
KIR3DL1	H	Inhibitory	HLA-Bw4
KIR3DL2	H	Inhibitory	HLA-A3, A11
KIR3DL3	H	Inhibitory	?
KIR2DS1, 2, 3, 4, 5	H	Activating	HLA-class I
KIR3DS1	H	Activating	HLA-Bw4
CD94/NKG2A	H, M	Inhibitory	H; HLA-E M; Qa-1b
CD94/NKG2C, E	H, M	Activating	H; HLA-E, M; Qa-1b
NKG2D	H, M	Activating	H; MICA/B, ULBP1–4, M; H60, MULT1, RAE1
CD16 (FcγRIII)	H, M	Activating	Immune complexes
CD27	H, M	Activating	CD70
CD244 (2B4)	H, M	Activating/ Inhibitory	CD48
Ly49A-C, E-G, I-O	M	Inhibitory	MHC-class I
Ly49D	M	Activating	H-2D^d
Ly49H	M	Activating	MCMV m157
Ly49P	M	Activating	H-2D^k/MCMV m04
KLRG1	H, M	Inhibitory	E-, R-, N-cadherins
NKR-P1A	H	Inhibitory	LLT1 (CLEC2D)
NKR-P1A, B, C, E, F	M	Activating/ Inhibitory	Clr family
NKR-P1B, D	M	Inhibitory	Clr-b, Clr-g
PILRα/PILRβ	M	Activating/ Inhibitory	O-glycosylated CD99
NKp46	H, M	Activating	Viral hemagglutinin, HSPG
DNAM-1 (CD226)	H, M	Activating	CD112, CD155

IL-15Rα complex is then presented on the cell surface in *trans* to the NK cell expressing the IL-15Rβγ complex.[15] In this manner, IL-15 is thought to maintain the homeostatic frequency of NK cells in the body. Once NK cells have acquired IL-15 responsiveness in bone marrow, they proceed through an ordered differentiation process that results in an expansion phase and the acquisition of a panel of activating and inhibitory receptors (Table 17.2).[1] It was previously assumed that these bone marrow–derived NK cells were fully functional, but recent data suggest that multiple mature NK-cell subsets exist, with varying levels of effector function.[14]

Tissue Distribution and Diversity of NK Cells

In line with their surveillance function, NK cells are found at many sites in the body. In mice, where tissue distribution has been thoroughly studied, NK cells are found in bone marrow, peripheral blood, the thymus, lymph nodes, and the spleen. The broad range of tissue distribution suggests that there is diversity in the function of mature NK cells. Furthermore, NK cells are also found in organs, such as the liver and small intestine, and in the decidual lining of the uterus along with a closely related innate lymphocyte population, named *group 1 innate lymphoid cells* (ILC1; Chapter 2), which likely have a role distinct from that of NK cells in host defense. Human NK cells can be divided into subsets based on the expression of CD16 (FcγRIII) and the

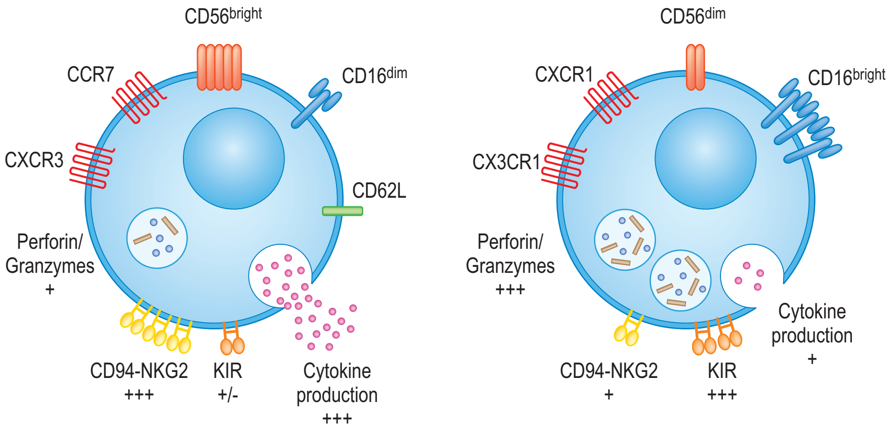


FIG 17.4 Schematic Representation of Human Natural Killer (NK) Cells. The human NK-cell subsets show distinct receptor expression and effector functions. CD56^{bright} NK cells produce high levels of cytokines and have patterns of chemokine and homing receptor expression that distinguish them from CD56^{dim} NK cells. CD56^{dim} NK cells express high levels of killer immunoglobulin-like receptor (KIR) and cytotoxic activity. The relative level of receptors and effector molecules is indicated on an arbitrary scale, with +/– being weak and +++ being strong expression.

adhesion molecule CD56 (Fig. 17.4).[16] CD56^{dim} NK cells represent less than 90% of peripheral blood NK cells, are CD16$^+$KIR$^+$, and display greater cytotoxicity. In contrast, their less mature counterparts, CD56^{bright} NK cells, show greater proliferative potential and cytokine production and are the principal NK cell population in secondary lymphoid organs. CD94$^+$CD161$^+$ NK cell progenitors can be found in cord blood and lymph nodes and develop into CD56^{dim} cells via CD56^{bright} intermediaries; this differentiation is also dependent on IL-15.[16] One important limitation in understanding the functional role played by NK-cell subsets has been the inability to identify populations corresponding to the CD56^{bright} and CD56^{dim} populations in rodents, as this molecule is not expressed in mouse NK cells. NK-cell heterogeneity, based on the differential expression of CD11b (integrin αM) and CD27, has been dissected in mice. Developing NK cells in bone marrow are CD27$^+$CD11b^{dim}, whereas these cells differentiate into CD27$^+$CD11b^{bright} and CD27$^-$CD11b^{bright} stages in the spleen. Interestingly, CD27$^+$ cells are found in lymph nodes, whereas the CD27$^-$ NK cells are localized predominantly in peripheral blood and the lung.[1,14] Comparative studies have shown that human CD56^{dim} NK cells are CD27$^-$, whereas CD56^{bright} cells are CD27^{++}, suggesting that the CD27 marker now enables direct comparisons of NK-cell subsets in mice and humans.

CYTOKINE REGULATION OF NK-CELL ACTIVATION, FUNCTION, AND HOMEOSTASIS

NK cells rely on IL-15 for their development and homeostatic maintenance in the peripheral organs. However, cultured NK cells also respond to other cytokines, including IL-2, -4, -12, -18, and -21, with their role being more important for function than homeostasis.[17] IL-2R shares its β and γ subunits with the constitutively expressed IL-15R (Chapter 9). The dimeric IL-2Rβγ is able to respond to the high levels of IL-2 supplied *in vitro* to induce proliferation of NK cells that were historically called *lymphokine-activated-killer* (LAK) cells. IL-2 activation will also induce the IL-2Rα chain to complete the high-affinity trimeric receptor. IL-4 activates human NK cells and promotes the proliferation of a fraction of NK cells characterized by their ability to produce IL-13.

The cytokines IL-12 and IL-18 have profound effects on NK-cell function.[18] IL-12 and IL-18 are produced by macrophages and DCs during inflammatory immune activation, such as viral infection. Although NK cells are present in IL-12– or IL-18–deficient mice and only marginally reduced in double mutant animals, cytotoxic activity and the ability to respond to infections, such as mouse CMV (MCMV) infection, is impaired. Cultivation of NK cells in IL-12 and/or IL-18 induces short-term activation, cytotoxicity, and IFN-γ production, whereas longer cultures produce more specialized cytokine-producing cells.[17]

Another NK cell–activating cytokine is IL-21.[19] IL-21 is a member of the common γ chain family of cytokines produced by CD4 T cells and, in contrast to IL-2 or IL-15, does not promote proliferation. Instead, it enhances maturation and effector function. Mouse NK cells treated with IL-21 display a broad-spectrum increase in cytotoxic function and produce cytokines, including IFN-γ and IL-10.[17] Of note, mice treated with IL-21 show a marked increase in NK cell–mediated tumor rejection, highlighting the potential for use of this cytokine as an anticancer therapeutic.[19]

NK-CELL RECEPTORS

NK cells differ from CTLs in that they do not require the expression of MHC class I to recognize target cells. In fact, the reintroduction of allogeneic MHC class I molecules into previously susceptible cell lines confers resistance to NK cell–mediated killing. These observations led to the missing-self hypothesis, which proposes that NK cells survey tissues for the usually ubiquitous MHC class I expression and react against cells that do not express it.[1]

The missing-self hypothesis, although altered over time to encompass other observations, has been extremely useful in providing a predictive framework by which to investigate NK-cell receptors and the recognition of target cells (Fig. 17.5).

In the past two decades, a large number of NK-cell receptors and their ligands have been identified (see Table 17.2).[20] Although a number of these NK-cell receptor ligands are MHC class I

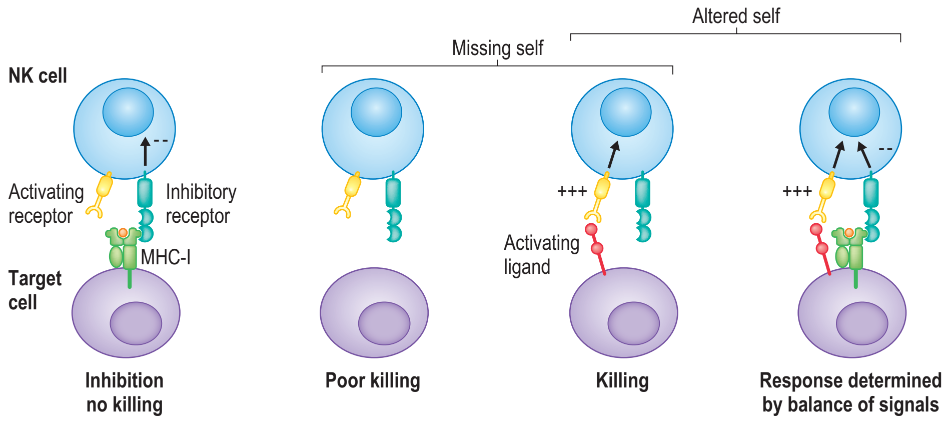


FIG 17.5 **Natural Killer (NK)–Cell Recognition of Target Cells.** NK cells express inhibitory receptors for major histocompatibility complex (MHC) class I and activating receptors for a variety of cellular, viral, and stress-induced ligands that alter the outcome of encounter with a target cell. The missing-self hypothesis initially predicted that NK cells would be activated to kill in the absence of MHC class I inhibition. However, an activating signal is also required. This activating signal can be provided by cellular ligands or by viral or stress-induced proteins, termed "altered self." In the presence of both inhibitory and activating signals, the outcome is determined by quantitative differences in signal strength between the two.

KEY CONCEPTS

Natural Killer (NK)-Cell Receptors

- Inhibitory receptors:
  - Recognize mostly major histocompatibility complex (MHC) class I ligands with high affinity
  - Signal via immunoreceptor tyrosine-based inhibitory motifs (ITIMs)
  - Recruit phosphatases (SHP and SHIP) to prevent a cytotoxic response
  - Required for NK-cell licensing
- Activating receptors:
  - Do not bind MHC class I molecules with high affinity
  - Ligands include viral molecules and stress-induced proteins
  - Signal via immunoreceptor tyrosine-based activation motifs (ITAMs)
  - Use several signaling adaptors, including DAP12

molecules themselves, as predicted by the missing-self hypothesis, there are many other classes of ligands. NK-cell receptors are classified into either activating or inhibitory types. This distinction was initially based on the *in vitro* properties of cross-linking these receptors but is increasingly regarded as the expression of the biochemical signal-transduction pathways activated by the receptor. Interestingly, although this strategy of target recognition is conserved in all mammals tested, in rodents and humans, the receptors have evolved from two independent gene families, killer immunoglobulin-like receptors (KIRs) in humans and the Ly49 family in mice (see Table 17.2).

NK-Cell Receptor Signaling

The signals derived from NK receptors are defined as inhibitory or activating in terms of their effect on NK-cell function. All extensively characterized inhibitory receptors carry an immunoreceptor tyrosine-based inhibitory motif (ITIM) in their intracellular domain. Ligation of ITIM-containing receptors causes tyrosine phosphorylation and the recruitment of a variety of phosphatases, including SHP and SHIP, that act to dampen downstream signaling pathways and NK-cell effector functions.[20] In contrast, activating receptors use immunoreceptor tyrosine-based activation motifs (ITAMs) to transduce stimulatory signals. Engagement of an ITAM-containing receptor results in tyrosine phosphorylation and recruitment of adaptor molecules, including FcεR1γ, CD3ζ, or DAP12/DAP10. The best-characterized activating receptor is CD16, an Fc receptor that binds IgG and is responsible for the antibody-dependent cellular cytotoxicity (ADCC) of human NK cells. CD16 recruits FcεR1γ and CD3ζ, which, in turn, attract the tyrosine kinases syk and ZAP70. These molecules then promote effector functions via multiple signal-transduction pathways.

NK RECEPTORS THAT RECOGNIZE MHC-I MOLECULES

NK cells recognize a wide variety of MHC class I molecules of both classic and nonconventional types. The receptors providing this recognition fit broadly into the immunoglobulin-like and lectin-like superfamilies. They show significant differences between mice and humans.

Killer Cell Immunoglobulin-Like Receptors in Humans

The *KIR* genes are a family of 15 genes that are physically linked on chromosome 19.[21] The locus shows a high degree of variation in humans, with both the number of *KIR* genes varying between individuals and extensive allelic polymorphisms. The KIR family was originally called *killer inhibitory receptors,* but subsequent studies have demonstrated that activating receptors are also encoded by the KIR locus. As expected, the inhibitory KIRs contain the ITIM, whereas all activating KIR utilize DAP12 to transduce signals from an ITAM.[20] Surface expression of KIR is acquired with maturation, driven by IL-15, and as such, NK cells lacking KIR or NK cells with combinations of multiple activating and inhibitory KIRs can be found in human blood. Similar to most other NK-cell receptors, some T cells can also express KIRs after activation, in particular in response to IL-15.

KIRs recognize the human MHC class I molecules human leukocyte antigen-A (HLA-A), -B, and -C (see Table 17.2).[20,21] The specificity of the inhibitory KIRs has been extensively characterized, with, for example, the different isoforms of KIR2D recognizing all known alleles of HLA-C. The ligands of the activating KIRs are less clear; however, they do not bind with high affinity to HLA molecules. KIR2DL4 is the most evolutionarily distinct member of the family and appears to be expressed in all activated NK cells in culture and on the CD56$^{\text{bright}}$ subset in peripheral blood. KIR2DL4 also has some distinct structural features and may act as an unconventional activating receptor binding HLA-G.

There are epidemiological data implicating particular KIR in a variety of autoimmune pathologies and viral responses.[22] For example, individuals with KIR2DS2 and some HLA-C alleles are predisposed to rheumatoid arthritis with vascular complications. Conversely, individuals with human immunodeficiency virus (HIV) infection who are homozygous for HLA-Bw4 progress to acquired immunodeficiency syndrome (AIDS) more slowly than those with other haplotypes, especially when they have the *KIR3DS1* gene. The mechanism by which this occurs is unknown, but the activity suggests that KIRs recognize an HIV-associated peptide in the context of HLA-Bw4.[23]

The Ly49 Family in Rodents

In a remarkable example of the power of selection to shape the evolution of the immune system, a multigenic locus with functional properties almost identical to those of the *KIR* genes evolved independently in rodents.[21] Mice, which have only two *KIR* genes of unknown function, have an analogous cluster of type II transmembrane spanning lectin-like genes called the Ly49 family on chromosome 6.[20] This family consists of more than 20 members and is highly polymorphic between mouse strains. Like KIR molecules, Ly49 receptors are also highly variegated in their expression in NK cells. *Ly49* genes encode activating and inhibitory receptors that bind to MHC class I molecules and signal via ITAMs and ITIMs, respectively (see Table 17.2).[20] Ligand-binding studies have revealed that inhibitory Ly49 receptors, such as A, C, and I, function to prevent self-reactivity by NK cells by binding to MHC class I molecules. The function of the activating receptors has proven more difficult to elucidate. Ly49D is known to have high affinity for the MHC class I H-2D^d allele and is involved in rejecting bone marrow allografts expressing H-2D^d, but the function of Ly49D in the normal immune response is unclear. For one activating receptor, Ly49H, the physiological function is known, as it recognizes the m157 molecule of MCMV and is important in early viral control.[24]

CD94 and NKG2 Family

Unlike the KIR or Ly49 families, the CD94/NKG2 complex is present both in rodents and humans. CD94/NKG2 receptors recognize nonclassic MHC class I ligands, such as HLA-E in humans and Qa1^b in mice.[22] A single CD94 gene is physically linked to four *NKG2* (A, C, E, and a truncated F) genes in humans. CD94 is found on the cell surface, either alone or with activating (C, E) or inhibitory (A) forms of NKG2. Interestingly, both the activating and the inhibitory complexes recognize HLA-E, which present predominantly the leader peptides of other HLA molecules, but not HLA-E itself. This system may provide a mechanism for NK cells to monitor the expression of multiple MHC class I proteins using HLA-E as a surrogate, with a prominent example being the requirement of CD94 for the control of mouse ectromelia virus.[25]

NKG2D

NKG2D, which is only distantly related to the NKG2 family, is a single, nonpolymorphic gene that is expressed on all NK cells from an early stage.[26] In the mouse, NKG2D signals through both DAP12 and another adaptor, DAP10, whereas human NK cells use only DAP10. Activation of NKG2D using a specific antibody results in enhanced cytotoxicity and cytokine secretion. The ligands of NKG2D are a family of molecules with structural similarity to MHC class I proteins, including MICA/B in humans and RAE-1 in mice. These ligands represent a diverse array of sequences, yet all bind with high affinity to NKG2D.[26] Interestingly, the MIC family in humans is highly polymorphic, suggesting that in this receptor–ligand interaction, the diversity comes from ligands rather than receptors. Transfection of otherwise resistant tumor cells with NKG2D ligands restores the susceptibility of these cells to NK-cell cytotoxic function. The key to understanding NKG2D function lies in the fact that the ligands are inducible and provide a mechanism for NK cells to detect stressed tissue, such as virally infected or malignant cells, a phenomenon that has been termed *altered self* (see Fig. 17.5).[26] Conversely, certain cancers have been found to mutate or lose NKG2D ligand expression to evade NK cell immunity.

NK CELL RECEPTORS THAT RECOGNIZE NON-MHC I MOLECULES

Beyond the multiple systems of activating and inhibitory receptors that NK cells have evolved to recognize MHC class I molecules and their structural variants, they also have several other receptor families that bind non–MHC class I ligands.[20,22] These include the NKR-P1 family, which is a polymorphic multigene family in rodents consisting of activating (A, C, F) and inhibitory (D) forms, but consists of only a single member in humans, whose activity is inhibitory. The ligands for some family members have been recently reported and are themselves lectin family receptors, including Clr-b and Clr-g in mice and LLT1 in humans.[20] Another NK receptor that is conserved between species is CD244 (2B4), a pan-NK-cell–expressed molecule whose ligand is CD48. Based on antibody cross-linking studies, human CD244 is designated an activating receptor. However, analysis of CD244-deficient mice suggests that the mouse homolog is inhibitory. A number of additional activating receptors exist, including the natural cytotoxicity receptors NKp30, NKp44, and NKp46. NKp30 and NKp46 are broadly expressed on human NK cells, and NKp46 is expressed on all mouse NK cells.[27] NKp44 is specifically expressed on activated human NK cells. These receptors can be activated by antibody cross-linking, and although their ligands are unknown, there is evidence that NKp46 binds hemagglutinin on influenza virus–infected cells, to *Fusobacterium nucleatum* and yet to be determined ligands on tumors, indicating that NKp46 may have a direct role in NK-cell activation to various pathogens.[27]

NK-CELL LICENSING AND SELF-TOLERANCE

The plethora of MHC class I–binding inhibitory receptors, as proposed by the missing-self hypothesis, explains the influence of

MHC class I on NK-cell lytic function against tumors. However, how self-tolerance is achieved has been less clear. The initial theory to explain self-tolerance was the "at least one receptor" model, which proposed that NK cells must express at least one self–MHC class I inhibitory receptor. A second model suggested that the receptor repertoire is shaped by selection by the specific MHC haplotype and the presence of self-ligands. The observation that NK cells are not autoreactive in the absence of any inhibitory ligands (MHC class I–deficient mice) and are actually poor killers suggests that the situation is more complex than these models allow for. A concept termed *licensing* was proposed to account for these observations.[28] Under this model, NK cells are initially unresponsive or "unlicensed" and acquire functional competency through binding of at least one inhibitory receptor before they can be activated and display cytotoxic function. An alternative "disarming" model proposes that all NK cells are initially responsive but that chronic stimulation by normal cells renders these cells unresponsive, or "anergic," unless the stimulation is opposed by MHC class I–specific inhibitory receptors. More recently a third "rheostat" model has also been proposed, whereby the number and affinity of inhibitory receptors it expresses tune NK-cell reactivity in a quantitative manner.[28] Regardless of the exact tolerance mechanism, it is interesting that unlike with T cells, there is no evidence for clonal deletion of autoreactive NK cells.

SPECIFIC NK-CELL FUNCTIONS

The ability to separate NK cells from T cells both phenotypically and genetically has greatly enhanced the understanding of NK-cell functions. Although some cytotoxic and immunomodulatory capacities overlap with those of T cells, it is also apparent that some functions of NK cells are unique. Specific examples of NK-cell functions are discussed below.

Control of Viral Infections

NK-cell activity rises early in the course of viral infection, partly driven by the release of IL-12, IL-18, and IFN-α/β, which stimulates activation (Chapter 25). The evidence that NK cells are essential for host defense against viruses comes directly from patients and mice lacking NK-cell function and indirectly from viral strategies to avoid NK-cell recognition. Human patients with selective deficiencies in NK cells show a pronounced susceptibility to recurrent severe infections, especially with HSV and CMV.

A powerful example of the role of NK cell–activating receptors in viral control is NK cell–mediated resistance to MCMV. Mouse strains that lack Ly49H are highly susceptible to MCMV, leading to uncontrolled viral replication and death.[24] Importantly, this protection is mediated by the recognition of the m157 protein of MCMV by Ly49H. The rapid accumulation of Ly49H$^+$ NK cells during MCMV infection is the first example of clonal expansion of NK cells in a manner similar to that of B and T cells.[24] Surprisingly, these studies also found compelling evidence for the maintenance of a long-lived NK-cell "memory"-like population (described below), blurring the line between innate (NK cells) and adaptive (B and T cells) immune cells.

As mentioned earlier, there is also evidence that NK cells have direct effects on the progression of HIV infection.[23] NK cells are able to lyse HIV-infected target cells either directly or by ADCC. Despite this ability, NK-cell responses are impaired in patients with HIV infection, as infected T-cell blasts selectively downregulate some *HLA* genes to avoid CTL activity but remain resistant to NK-cell cytotoxicity. These findings are supported by studies that show that high-risk, but uninfected, individuals appear to have increased NK-cell activity and that the combination of the expression of HLA-Bw4 and the *KIR3DS1* gene is associated with delayed progression to AIDS. Finally, HIV viremia induces several functional abnormalities on NK cells, suggesting that this complex virus and NK cells interact at multiple levels.[23]

Control of Malignant Cells

NK cells were named as such during the early 1970s for their potent ability to kill leukemia cell lines *in vitro* and since then have been the focus therapies to boost cancer lysis in humans. Evading the immune system is a hallmark of cancer, but the contribution of NK cells in tumor immune surveillance has been difficult to assay as specific NK cell–deficient mouse models have only recently been generated. There is abundant evidence that NK cells can reduce tumor burden in animal models, and that the administration of cytokines that enhance NK-cell function or number, including IL-2, -12, -15, and -21, or those that induce IFN-α production, are protective against metastasis.[18] In light of this, many clinical trials have been conducted to assess either cytokine treatment or the injection of *ex vivo* treated NK cells.[18] Unfortunately, the high doses of IL-2 and IL-15 required for efficacy are relatively toxic, and transferred NK cells have proved difficult to target to tumors. Despite this, some successes have been demonstrated for melanoma, leukemia, and lung and hepatic cancers. Today, numerous immunotherapy trials are ongoing and aim to exploit and boost NK cell–mediated tumor killing, including antibodies against inhibitory receptors on NK cells, modified IL-2/IL-15 reagents, bispecific or trispecific and killer engaging (BiKE or TRiKE) antibodies that bind NK cells and tumor antigens, tumor-specific antibodies, *in vitro* expanded patient-derived NK cells, and irradiated NK cell lines, to name a few.[18]

Role of NK Cells in Hematopoietic Stem Cell Transplantation

CLINICAL PEARLS

Exploiting Natural Killer (NK) Cells in Leukemia Therapy

- Hematopoietic stem cell transplantation requires donor and recipient human leukocyte antigen (HLA) matching to reduce graft-versus-host disease (GvHD) mediated by transplanted cytotoxic T lymphocytes (CTLs).
- Haploidentical donors and recipients (those that share only one HLA haplotype) represent 50% of unrelated transplants and undergo a stronger conditioning regimen to avoid GvHD.
- Alloreactive NK cells are present after haploidentical transplant and provide a potent graft-versus-leukemia (GvL) effect in animal models.
- Transplantation from NK cell–alloreactive donors controls leukemia relapse and improves engraftment without causing GvHD.

In humans, allogeneic bone marrow transplantation can cure leukemia through the reaction of donor CTLs in the graft against the residual leukemic cells. These transferred T cells also mediate GvHD. The need to prevent GvHD as a result of strong immunosuppression is the major cause of transplantation failure because of infection and cancer relapse. It has been proposed that obtaining the transplant from a haploidentical donor (identical at one HLA haplotype and fully mismatched on the other, for example, a parent) provides allogeneic NK cells with an HLA haplotype that supplies more KIR ligand than a matched recipient would provide. Hence it would yield a stronger graft-versus-leukemia (GvL) effect.[29] Indeed, mice treated with alloreactive NK cells tolerate 30 times the lethal dose of mismatched bone marrow cells without developing GvHD, and alloreactive NK cells eradicated human acute myeloid leukemia (AML) transplanted into immune-deficient mice. Retrospective studies on patients with AML and acute lymphoblastic leukemia (ALL) who received haploidentical grafts revealed that transplants with alloreactive NK cells showed better engraftment and GvHD protection and less relapse.[29]

These studies suggest that therapeutic treatment with alloreactive NK cells will be effective in eliminating residual cancer cells following frontline treatments or in preventing cancer relapse. Feasibility studies have shown that the production of clinical-grade cultured human NK cells is possible and that the transferred cells persist for some time in the patient. An alternative strategy is the use of monoclonal antibodies (mAbs) directed against a particular inhibitory KIR. Preclinical studies suggest that blockade of the inhibitory KIR enhances antitumor activities of the endogenous NK cells, providing a rationale for ongoing clinical trials to test this approach.[18,29]

NK-Cell Memory

ON THE HORIZON

Clinical Trials of Transferred Alloreactive Natural Killer (NK) Cells

- Adoptive transfer of cultured clinical grade NK cells has proven feasible and well tolerated by patients and is a promising therapy to limit relapse in patients with cancer.
- A treatment regimen comprising preconditioning with chemotherapeutic reagents followed by alloreactive NK cells and interleukin-2 (IL-2) administration has been examined.
- Therapeutic blockade of inhibitory NK-cell receptors using anti–killer immunoglobulin-like receptor (KIR) monoclonal antibodies is undergoing clinical trials for their efficacy in boosting NK cell–mediated killing of cancer cells.

During the last decade several studies have described the ability of NK cells to generate a "memory" response to pathogens and antigens. The original documentation of NK-cell memory was in a cutaneous contact hypersensitivity (CHS) model. CHS involves the sensitization by skin exposure to the haptens 2,4-dinitrofluorobenzene (DNFB) and oxazolone (OXA), followed by rechallenge at a previously unexposed site, such as the ear, several days later. NK cells accumulated in the ears of T/B cell– deficient mice upon DNFB rechallenge, and this was dependent on the adhesion molecules of the selectin family.[24] The concept of NK-cell memory has since been extended to viral responses, including MCMV, influenza, vaccinia virus, HSV, HIV, and vesicular stomatitis virus where NK cells from mice preexposed to the virus were more efficiently activated and prolonged survival from lethal infection compared with naïve NK cells. Although the interaction between Ly49H and the viral protein m157 is required for the memory response elicited by MCMV infection, the NK cell receptor–ligands combinations involved in the other viral models are not established. In most settings, the NK-cell memory response is specific to the antigen (hapten or virus), and the memory NK cells or their precursors are found within the CXCR6$^+$CD49a$^+$ liver ILC1 pool.[14] These liver-resident ILC1 are transcriptionally and phenotypically distinct from conventional NK cells, and their role in mammalian immunity is still to be elucidated. Interestingly, ILC1 do not migrate in steady state. Thus how they are activated in a hapten-specific manner during CHS remains to be determined.[14] Furthermore, inflammation, in general, and especially the cytokines IL-12, IL-18, type I IFN, and the NLRP3 inflammasome, are critical for the generation of hapten-dependent memory NK-cell development.[24]

INTERACTIONS OF CTL AND NK CELLS IN THE IMMUNE RESPONSE

Although studies of CTLs and NK cells in isolation have greatly advanced our understanding of their functions, it is obvious that these immune cells function in a system that depends on numerous interactions between the various cell types at multiple levels. In the case of cytotoxic cells, these immunomodulatory interactions are becoming increasingly appreciated (see Fig. 17.2). In particular, it has become apparent that NK cells and DCs interact specifically to promote some outcomes, such as maturation and priming of NK cells. CTL activation by mature DCs is influenced directly by NK cell–derived IFN-γ and indirectly through the role of NK cells in promoting a Th1 response in CD4 T cells (Chapter 16).[30]

Increasing evidence has emerged to implicate DC–NK cell cross-talk in various aspects of the immune response. DCs produce a variety of cytokines that modulate NK-cell behavior, including IL-12, IL-15, IL-18, and IFN-α.[30] The outcome of these interactions for the DC depends on its maturity; immature DCs are killed by NK cells, whereas mature DCs are resistant to lysis. The interactions of mature DCs and NK cells would be expected to occur at the site of infection, where DCs provide inflammatory stimuli for NK cells. Type I IFN production by DCs promotes MHC-I upregulation by CTLs and their protection from NK-cell killing during viral infections. The other site of encounter is the lymph node, where, during an immune reaction, NK cells are recruited by chemokines and interact with mature DCs and CD4 T cells to induce a Th1 response. This process requires IFN-γ production from NK cells (see Fig. 17.2).[30]

NK cell–CTL interactions are also important in generating an immune response to tumors that are, in general, poorly immunogenic.[31] DCs that are matured *in vitro* by NK cells show strongly enhanced ability to induce Th1 and CTL responses, including cytokine secretion. IFN-γ, in particular, is very important in the rejection of primary tumors and the formation of CTL memory to tumors. It is also likely that killing by NK cells and CTLs provides DCs with increased access to tumor antigens and promotes further adaptive immunity. Using DCs to harness the helper function of NK cells as well as the cytotoxic functions of both CTLs and NK cells offers therapeutic promise and is currently being tested in a variety of cancers.

EVASION OF THE CYTOTOXIC RESPONSE

Viruses

As one principal function of CTL and NK cells is the control of viral infections, it is not surprising that viruses have strategies to interfere with the host response (Chapter 25). The multiplicity of these evasion strategies indicates that this is an essential step for long-term viral persistence.

These strategies include:

Latency. This involves minimizing viral gene expression and thereby avoiding detection. Examples include HSV in neurons, HIV in T cells, and EBV in B cells.

Antigenic variation. Viruses possess the ability to rapidly mutate its genome and produce escape variants that are no longer visible to CTLs. Such mutations were shown for MCMV infection in mouse and HIV infection in humans.

Infection of immune nonaccessible sites. Such inaccessible sites include infection of the central nervous system by HSV or rubella virus.

Production of viral defense molecules (immunoevasins). Many viruses, including adenovirus and EBV, interfere with cytotoxic activity by producing proteins that either hinder Fas or TNF-mediated killing or inhibit the function of antiviral cytokines, such as IFN-α. A number of viruses, including EBV, produce homologs of antiapoptotic molecules, such as Bcl2, to inhibit killing by CTLs. Various members of the poxvirus family have evolved homologs of the naturally occurring IL-18–binding protein that inhibits IL-18 activity and NK-cell function.[32]

Modulation of molecules involved in recognition. A widely utilized viral strategy to evade the cytotoxic response is to interfere with antigen processing, presentation, or the expression of other molecules required for CTL recognition (Chapter 6).[33] Several viruses, including adenovirus, downregulate MHC class I expression on the cell surface. This can be achieved by a number of mechanisms. For example, adenovirus type 2 E3 protein forms a complex with MHC class I to prevent antigens from being processed; MCMV gpm152 protein causes retention of the MHC class I molecules in the Golgi compartment; and CMV proteins US2 and US11 promote the rapid degradation of newly synthesized MHC class I complexes. An alternative approach is to interfere with antigen processing, either inhibiting the expression of the TAP protein, as is the case for HSV, or producing proteins that are resistant to antigen digestion by the proteasome, such as the EBNA-1 protein of EBV. This inhibition strategy is not restricted to MHC class I, as both human CMV and MCMV express proteins that inhibit the cell surface expression of NKG2D ligands.

Tumor Cells

Part of the evidence that CTLs and NK cells function to control malignant cells comes from the effort tumor cells will go to avoid cytotoxic activity (Chapter 77). Conversely, promoting the cytotoxic response either through specific tumor antigens, blocking immune checkpoint inhibitors, or through polyclonal stimulation remains one of the most actively pursued strategies in cancer therapy.

Tumors evade cytotoxic function in a number of ways:

Downregulation or loss of MHC class I expression. This strategy is common in solid tumors, including metastatic melanoma and breast cancer, where MHC class I downregulation accounts for up to 50% of samples. MHC class I loss is also induced by mutations in the gene encoding β_2-microglobulin or transcription factors that regulate its expression. MHC class I downregulation is associated with changes in the regulatory mechanisms controlling antigen presentation and can often be corrected by treatment with cytokines, such as IFN-γ.

Induction of immune checkpoint ligands. One mechanism to suppress the infiltrating antitumor CLT/NK cell response is upregulation of ligands to inhibitory receptors expressed on CTLs/NK cells. Two inhibitory receptors, CTLA-4 and PD-1, are expressed on CTLs and to a less extent on NK cells, and blocking of antibodies against these has revolutionized cancer immunotherapy.[12] In particular, IFN-γ upregulates PD-1 ligands (PD-1L and PD-2L) on tumor cells, and their binding to PD-1 potently inhibits tumor-specific CTL activity. Anti-PD-1 monotherapy offers a high rate of sustained tumor regression in metastatic melanoma and is in clinical trials in lung and breast cancers.

Antigenic mutation. Tumors can also avoid CTL activity by antigenic loss. This strategy takes the form of silencing or mutating epitopes that are particularly immunogenic to CTL.[31]

Broad-spectrum tumor-derived immune suppression. Tumors express a variety of membrane-bound and soluble factors that can suppress the immune response, including FasL, which protects the tumor by inducing apoptosis in activated Fas-expressing CTLs. This model is not universally accepted, and a role for FasL in inducing the expression of inflammatory cytokines is also possible. Tumors can also express TGF-β, which acts on CTLs and NK cells to inhibit their metabolism, proliferation, and expression of effector molecules, such as perforin and granzymes. TGF–β also acts on NK cells to downregulate the expression of NKG2D. Pharmaceutical inhibitors of TGF-β signaling have been developed, and one of them, galunisertib, is being tested in patients with cancer.[18] In addition, tumors can produce soluble decoy ligands, such as MIC, which suppresses NKG2D function. CD73-mediated adenosine production also plays important immunosuppressive roles in the tumor microenvironment. CD73 is a surface receptor expressed on many tumors and catalyzes extracellular adenosine monophosphate (AMP) into adenosine that can bind its receptor (A2AR) on NK cells/CTLs to suppress their activity. As such, A2AR and CD73 antagonists are being developed, and their trials in cancer are ongoing.[18]

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Sun JC, Lanier LL. NK cell development, homeostasis and function: parallels with CD8(+) T cells. Nat Rev Immunol 2011;11:645–57.
2. Voskoboinik I, Whisstock JC, Trapani JA. Perforin and granzymes: function, dysfunction and human pathology. Nat Rev Immunol 2015;15(6):388–400.
3. Strasser A, Jost PJ, Nagata S. The many roles of FAS receptor signaling in the immune system. Immunity 2009;30:180–92.
4. Johnstone RW, Frew AJ, Smyth MJ. The TRAIL apoptotic pathway in cancer onset, progression and therapy. Nat Rev Cancer 2008;8:782–98.
5. Mueller SN, Gebhardt T, Carbone FR, et al. Memory T cell subsets, migration patterns, and tissue residence. Annu Rev Immunol 2013;31:137–61.
6. Dustin ML. The immunological synapse. Cancer Immunol Res 2014;2:1023–33.
7. Norbury CC. Defining cross presentation for a wider audience. Curr Opin Immunol 2016;40:110–16.
8. Chang JT, Wherry EJ, Goldrath AW. Molecular regulation of effector and memory T cell differentiation. Nat Immunol 2014;15:1104–15.
9. Cox MA, Harrington LE, Zajac AJ. Cytokines and the inception of CD8 T cell responses. Trends Immunol 2011;32:180–6.
10. Greyer M, Whitney PG, Stock AT, et al. T cell help amplifies innate signals in CD8(+) DCs for optimal CD8(+) T cell priming. Cell Rep 2016;14:586–97.

11. Attanasio J, Wherry EJ. Costimulatory and coinhibitory receptor pathways in infectious disease. Immunity 2016;44:1052–68.
12. Sharma P, Allison JP. Immune checkpoint targeting in cancer therapy: toward combination strategies with curative potential. Cell 2015;161(2):205–14.
13. Reguzova AY, Karpenko LI, Mechetina LV, et al. Peptide-MHC multimer-based monitoring of CD8 T-cells in HIV-1 infection and AIDS vaccine development. Expert Rev Vaccines 2015;14:69–84.
14. Seillet C, Belz GT, Huntington ND. Development, homeostasis, and heterogeneity of NK Cells and ILC1. Curr Top Microbiol Immunol 2016;395:37–61.
15. Huntington ND. The unconventional expression of IL-15 and its role in NK cell homeostasis. Immunol Cell Biol 2014;92:210–13.
16. Yu J, Freud AG, Caligiuri MA. Location and cellular stages of natural killer cell development. Trends Immunol 2013;34:573–82.
17. Brady J, Carotta S, Thong RP, et al. The interactions of multiple cytokines control NK cell maturation. J Immunol 2010;185:6679–88.
18. Guillerey C, Huntington ND, Smyth MJ. Targeted natural killer cells in cancer immunotherapy. Nat Immunol 2016;17:1025–36.
19. Croce M, Rigo V, Ferrini S. IL-21: a pleiotropic cytokine with potential applications in oncology. J Immunol Res 2015;2015:696578.
20. Pegram HJ, Andrews DM, Smyth MJ, et al. Activating and inhibitory receptors of natural killer cells. Immunol Cell Biol 2011;89:216–24.
21. Guethlein LA, Norman PJ, Hilton HH, et al. Co-evolution of MHC class I and variable NK cell receptors in placental mammals. Immunol Rev 2015;267:259–82.
22. Das J, Khakoo SI. NK cells: tuned by peptide? Immunol Rev 2015;267:214–27.
23. Scully E, Alter G. NK Cells in HIV Disease. Curr HIV/AIDS Rep 2016;13:85–94.
24. Cerwenka A, Lanier LL. Natural killer cell memory in infection, inflammation and cancer. Nat Rev Immunol 2016;16:112–23.
25. Fang M, Orr MT, Spee P, et al. CD94 is essential for NK cell-mediated resistance to a lethal viral disease. Immunity 2011;34:579–89.
26. Raulet DH, Gasser S, Gowen BG, et al. Regulation of ligands for the NKG2D activating receptor. Annu Rev Immunol 2013;31:413–41.
27. Kruse PH, Matta J, Ugolini S, et al. Natural cytotoxicity receptors and their ligands. Immunol Cell Biol 2014;92:221–9.
28. Kadri N, Thanh TL, Hoglund P. Selection, tuning, and adaptation in mouse NK cell education. Immunol Rev 2015;267:167–77.
29. Moretta L, Montaldo E, Vacca P, et al. Human natural killer cells: origin, receptors, function, and clinical applications. Int Arch Allergy Immunol 2014;164:253–64.
30. Chijioke O, Munz C. Dendritic cell derived cytokines in human natural killer cell differentiation and activation. Front Immunol 2013;4:365.
31. Vesely MD, Kershaw MH, Schreiber RD, et al. Natural innate and adaptive immunity to cancer. Annu Rev Immunol 2011;29:235–71.
32. Alzhanova D, Fruh K. Modulation of the host immune response by cowpox virus. Microbes Infect 2010;12:900–9.
33. Christiaansen A, Varga SM, Spencer JV. Viral manipulation of the host immune response. Curr Opin Immunol 2015;36:54–60.

MULTIPLE-CHOICE QUESTIONS

1. How do cytotoxic T cells recognize target cells?
   A. They cannot, only natural killer cells can recognize tumor cells
   B. Through detection via stress-induced ligands
   C. Recognition of major histocompatibility complex (MHC) class I–peptide complexes specific for the T-cell receptor
   D. Through sensing the inflammatory environment

2. How is the immunity provided by memory cells superior to that elicited by primary challenge?
   A. The memory compartment has a higher frequency of cells specific for the given antigen.
   B. Memory cells are intrinsically able to respond faster to rechallenge with the same infection.
   C. Memory cells can remain at the site of previous infection.
   D. All of the above.

3. How do tumor cells evade cytotoxic T cells?
   A. Avoid them by residing away from the bloodstream
   B. Express ligands that interact with inhibitory receptors on cytotoxic T cells
   C. Cytotoxic T cells cannot kill tumor cells
   D. Proliferate more rapidly than the tumor cells

4. Natural killer cells selectively target:
   A. The tumor microenvironment
   B. Immune cells activated by infection
   C. Cells that have lost MHC class I expression on the cell surface
   D. Memory cytotoxic T cells

5. Which is NOT a pathway that cytotoxic T cells and natural killer cells use on their targets?
   A. Cellular senescence
   B. Fas pathway
   C. Perforin pores
   D. Granzyme activity

18

Regulatory Immune Cells

James B. Wing, Shimon Sakaguchi

The normal mammalian immune system protects the individual from a myriad of potentially pathogenic microorganisms. However, the immune system must also be tightly regulated to prevent it from attacking self-constituents and thus causing autoimmunity. This is partly achieved through central tolerance and the deletion of T cells that recognize self antigens in the thymus. However, this process is incomplete; that is, some self-reactive T cells are present in the periphery of healthy individuals and are capable of causing autoimmunity. This creates a need for peripheral tolerance mechanisms that control the action of such self-reactive cells.

The mechanisms that suppress harmful immune responses are of interest to both basic and clinical immunologists, as the failure of protective immunity can lead to increased susceptibility to infectious diseases, and loss of self-tolerance may trigger an autoimmune disorder. Furthermore, it is often clinically desirable to enhance an immune response to certain self (or quasi-self) antigens, such as tumor antigens, or to induce immune suppression for the purpose of facilitating organ transplant acceptance. Elucidation of the mechanisms responsible for immune regulation and maintenance of self-tolerance is, therefore, one of the primary goals of current medical immunology.

KEY CONCEPTS

Immunological Self-Tolerance

Immunological self-tolerance is actively acquired and maintained throughout life by a series of mechanisms that cooperatively and complementary operate to prevent the maturation and activation of potentially self-reactive lymphocytes.

The mechanisms include:

- Clonal deletion
- Clonal anergy
- Clonal ignorance
- Dominant suppression

One key feature of the adaptive immune response is that once triggered, it shows essentially the same effector activity, whether the target antigen is a microbe or a self antigen, leading to elimination of the microbe or destruction of self-tissue. To prevent self-destructive immune responses while allowing protective immune responses to nonself antigens, the mammalian immune system has evolved various regulatory contrivances that inhibit the initial generation of potentially harmful self-reactive T and B lymphocytes, termed *central tolerance,* or, after lymphocyte generation, downregulate cellular activation and expansion upon encounter with self antigens, termed *peripheral tolerance.* For T cells, central tolerance is established in the thymus, where many potentially dangerous lymphocytes carrying high-affinity T-cell receptors (TCRs) for self antigens are deleted via negative selection during development. This results in the generation of a peripheral T-cell repertoire that is largely self-tolerant. However, there is abundant evidence that some autoreactive T cells escape thymic deletion, and potentially pathogenic self-reactive T cells are, indeed, present in most individuals. Nevertheless, autoimmune diseases only occur infrequently, indicating that autoreactive T cells are somehow controlled in the periphery. Such peripheral mechanisms of self-tolerance include further deletion of self-reactive T cells, seclusion of self antigen from T lymphocytes, low TCR affinity or lack of costimulation in antigen recognition (clonal ignorance), inactivation of autoreactive T lymphocytes upon encounter with antigen without costimulation (clonal anergy), and active suppression of self-reactive lymphocytes by other lymphocytes (peripheral suppression).[1]

There are various mechanisms of peripheral self-tolerance, and this chapter deals with peripheral suppression mediated by regulatory T cells (Tregs) and other suppressive non–T cells. Several types of T cells with regulatory activity have been described, including subpopulations of $\gamma\delta$ T cells, natural killer T cells (NKT cells), and CD8 and CD4 T cells (Table 18.1; Fig. 18.1). Some of these Tregs are constitutively produced as a separate lineage in the immune system, whereas others are induced from naïve T cells as a product of a particular mode of antigen stimulation in a particular cytokine milieu. Although it remains to be determined how each cell population is functionally stable and physiologically important, this abundance and apparent redundancy of Treg populations may not be surprising when one considers how essential it is to maintain immune homeostasis and self-tolerance.

This chapter focuses on CD4 Tregs; in particular thymically produced Tregs (tTregs) that specifically express the transcription factor Foxp3 and maintain high expression of the interleukin-2 (IL-2) receptor α chain, CD25. Foxp3$^+$CD25$^+$CD4$^+$ Tregs have been the subject of the majority of recent Treg studies and may have the broadest implication to our understanding of the mechanism of various immunological disorders. Loss of function or numerical deficiency of tTreg function or number can be a primary cause of autoimmune disease, allergy, and inflammatory disorders, such as inflammatory bowel disease (IBD) in humans. Conversely, since these cells can prevent targeting of tumor tissues by conventional CD4 and CD8 T cells, inactivating them is a key goal in cancer immunotherapy. Because of their natural presence in the immune system, they are also a good target for

TABLE 18.1 Subsets of Thymic and Peripheral Foxp3$^+$ Tregs and Foxp3$^-$ Tregs

	FOXP3$^+$Treg	Tr1 and Th3	Double Negative Tregs	Qa-1–Restricted CD8$^+$Treg	CD8$^+$CD28$^-$Treg	NKT Cell	γδ T Cell
Site of generation	Thymus/ periphery	Periphery	Periphery	Periphery	Periphery	Periphery	Periphery
Marker	Foxp3, CD25, CTLA-4, GITR	IL-10 TGF-β	TCR αβ$^+$CD4$^-$CD8$^-$NK1.1$^-$	Nonclassic MHC Ib Qa-1	Not specified	Invariant TCR chain Va14 (mouse), Va24 (human)	Various subsets Vg5$^+$ (mouse) Vg1$^+$ (human)
Specificity	Peptide plus MHC class II	Peptide plus MHC class II	Peptide plus MHC class I	Peptide plus MHC class Ib	Peptide plus MHC class I	Glycolipids plus CD1d	Glycolipids plus CD1, Peptide plus MHC class Ib
Target cell	T cells, B cells, APCs, NK cells NKT cells	T cell	T cell	T cell	DCs/APCs	T cells, APCs	T cells, APCs, epithelial cells
Suppressive mechanisms	Cell-contact Costimulation modification Cytokine production	IL-10 TGF-β	Perforin, trogocytosis, Fas	Perforin	Induction of ILT3/ILT4 in DCs	IL-10, Th2 cytokines	Lysis, CD95-CD95 ligand pathway, thymosin-b4
Reported suppressive function	Autoimmunity Transplantation Allergy Infection Cancer	Autoimmunity Transplantation Allergy	Autoimmunity Transplantation	Autoimmunity	Autoimmunity Transplantation	Autoimmunity Transplantation Cancer	Autoimmunity Allergy (dermatitis) Infection

APC, antigen-presenting cell; DC, dendritic cell; ILT, immunoglobulin transcript; Tr1, regulatory type 1 cell; Th3, T-helper 3 cell; Treg, regulatory T cell; MHC, major histocompatibility complex.

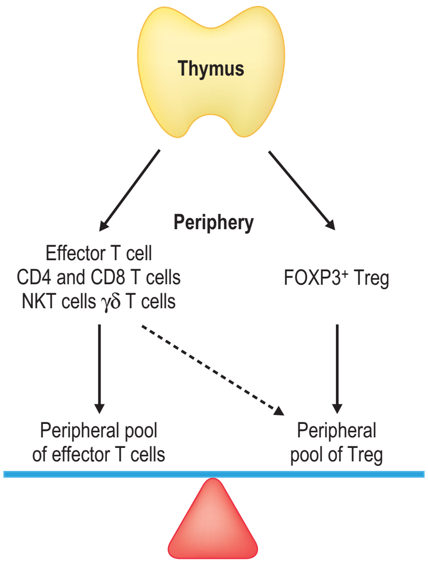


FIG 18.1 Developmental Pathways of Regulatory T Cells (Tregs). Treg cells can develop either in the thymus or the periphery and are vital for maintaining tolerance as a counterbalance to effector T cells. Thymically generated Tregs express Foxp3 and develop within the thymus following to a specialized combination of T-cell receptor (TCR) and costimulatory signals. Extrathymic development of CD4$^+$ Tregs and CD8$^+$ Tregs can ensue from a host of different conditions, such as high concentrations of transforming growth factor-β (TGF-β), interleukin-10 (IL-10), or other particular circumstances surrounding antigen priming. The signals that control differentiation of γδ T cells and natural killer T cells (NKT cells) to cells with regulatory properties are less well defined.

the treatment and prevention of a variety of immunological diseases.[2]

CD4 REGULATORY T CELLS

The nomenclature used to define Tregs has been inconsistent over the years, but recent efforts have been made to standardize the naming of these cells.[3] It is recommended that Foxp3$^+$ Tregs be clearly separated into thymus-derived tTregs, peripherally *in vivo*–induced pTregs and *in vitro*–induced iTregs. Readers should be aware that in the past tTregs have often been referred to as *natural Tregs* (nTregs), whereas both pTregs and iTregs were grouped together as *induced* or *adaptive Tregs.* Where there is no specific identification allowing confirmation of the cells being tTregs or pTregs, the simple term *Tregs* can be used to refer to Foxp3$^+$Tregs. In this chapter, Foxp3$^-$ Treg populations, such as Tr1 and Th3, are referred to collectively as *Foxp3$^-$ Tregs* but have also previously been described as induced or adaptive Tregs.[4]

Foxp3$^-$ CD4 Tregs, such as IL-10–secreting type 1 regulatory T cells (Tr1) and transforming growth factor-β (TGF-β)–secreting T-helper-3 (Th3) cells also differentiate from naïve T cells following antigen stimulation in conjunction with specific conditions in the periphery. Although both Foxp3$^+$ Tregs and Foxp3$^-$ Tregs are suppressive, they differ in their development, phenotype, and function, and they are thus considered to be of separate lineages (see Table 18.1 and Fig. 18.1).[4]

Thymus-Derived Regulatory T Cells

The first report of autoimmune-preventive, thymus-derived T cells in the normal immune system occurred about 40 years ago when thymectomy on day 3 of life (d3Tx) was shown to cause organ-specific autoimmune diseases, such as oophoritis, in otherwise healthy mice, a finding later attributed to the slightly delayed thymic emigration of tTregs in comparison with effector T cells.[5] Subsequent studies showed that the development of autoimmune diseases could be inhibited if the thymectomized animals were reconstituted with CD4$^+$CD8$^-$ thymocytes or CD4$^+$ splenocytes from histocompatible immune-uncompromised animals. Athymic mice transferred with non-Tregs or thymocytes spontaneously develop organ-specific autoimmune diseases, which can be reversed by cotransfer of tTregs from normal adult mice.[6,7] Tregs can suppress the proliferation and cytokine production of conventional CD4 or CD8 T cells *in vitro.* tTregs are thought to arise from T-cell clones with relatively high reactivity to self antigens presented in the thymus.

To pinpoint a specific phenotype for CD4 T cells with regulatory function, surface markers with more restricted expression patterns have been explored. High and stable expression of the IL-2 receptor α chain, CD25 has been found to be a useful and specific surface marker of Foxp3$^+$ Tregs. Between 5% and 10% of CD4$^+$ T cells express CD25 constitutively in the thymus and periphery of mice. Importantly, transfer of CD4 lymphocytes depleted of CD25$^+$ cells induces autoimmunity in athymic nude mice, whereas cotransfer of CD4$^+$CD25$^+$ cells protects the mice from disease induction (Fig. 18.2). Other markers shown to be associated with Foxp3$^+$ Tregs are cytotoxic T lymphocyte antigen-4 (CTLA-4; CD152) and glucocorticoid-induced tumor necrosis factor receptor protein (GITR). However, these and CD25 are not truly specific markers of tTregs, since conventional T cells upregulate GITR, as well as CTLA-4 and CD25, after activation.[6] This dilemma is especially apparent when investigating Foxp3$^+$ Tregs in humans, where a significant proportion of CD4$^+$ T cells in peripheral blood express CD25, yet only 2%–4% of CD4$^+$ T cells, enriched among cells with the highest expression level of CD25 (CD25^{high}), have suppressive properties (Fig. 18.3).[8] The fact that CD25$^+$ Tregs are not a discrete population in humans poses a problem both when obtaining cells for experimental purposes and when evaluating their role in a clinical setting. Therefore finding more specific cell surface markers of CD4$^+$ natural Tregs remains an important goal.

Thymus-Derived Tregs Express the Transcription Factor Foxp3

Specific expression of the transcription factor Foxp3 is closely linked with the development and function of Tregs.[9] The first hint as to the significance of Foxp3 was given by studies of the Scurfy mutant mouse. This mouse strain suffers from a spontaneous X-linked mutation of the *Foxp3* gene, which leads to fatal lymphoproliferative disease associated with multiorgan infiltrates and early death by 3–4 weeks of age in hemizygous males. Similarly, mutations in the human orthologue *FOXP3* are linked to immune dysregulation, polyendocrinopathy, enteropathy, IBD, allergic dermatitis, food allergy, hyperimmunoglobulinemia E, hematological disorders, severe infections and X-linked (IPEX) syndrome, which is an X-linked immunodeficiency syndrome associated with organ-specific autoimmune diseases, such as type 1 diabetes (see Fig. 18.2).[6] Common features of IPEX syndrome and scurfy mice are deficient levels of both tTregs and pTregs.

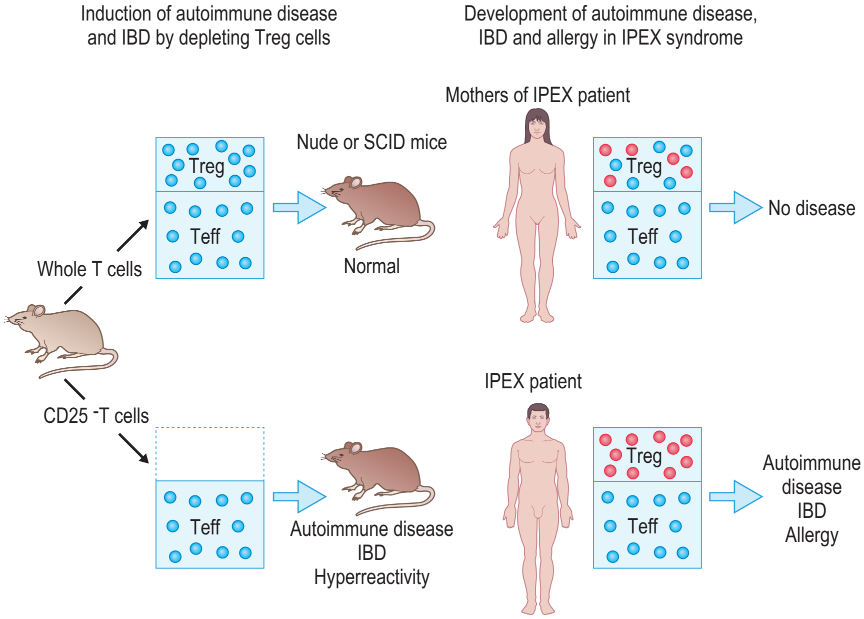


FIG 18.2 Dominant Self-Tolerance in Rodents and Humans. Transfer of T-cell suspensions depleted of Foxp3$^+$CD4$^+$ Tregs induces autoimmune disease and inflammatory bowel disease (IBD) and heightens immune responses to nonself antigens in athymic nude or severe combined immunodeficiency (SCID) mice *(left).* Male children are afflicted by infections and immune dysfunction/polyendocrinopathy/enteropathy/X-linked (IPEX) syndrome. Their mothers with hemizygous defects of the *FOXP3* gene bear defective and normal Tregs as a mosaic because of random inactivation of the X chromosome in each Treg cell. They are completely normal because normal Tregs dominantly control the activation and expansion of effector T (T_{eff}) cells that mediate autoimmune disease, IBD, and allergy *(right).* Blue circles represent intact Treg or T_{eff} cells; red circles represent defective Tregs.

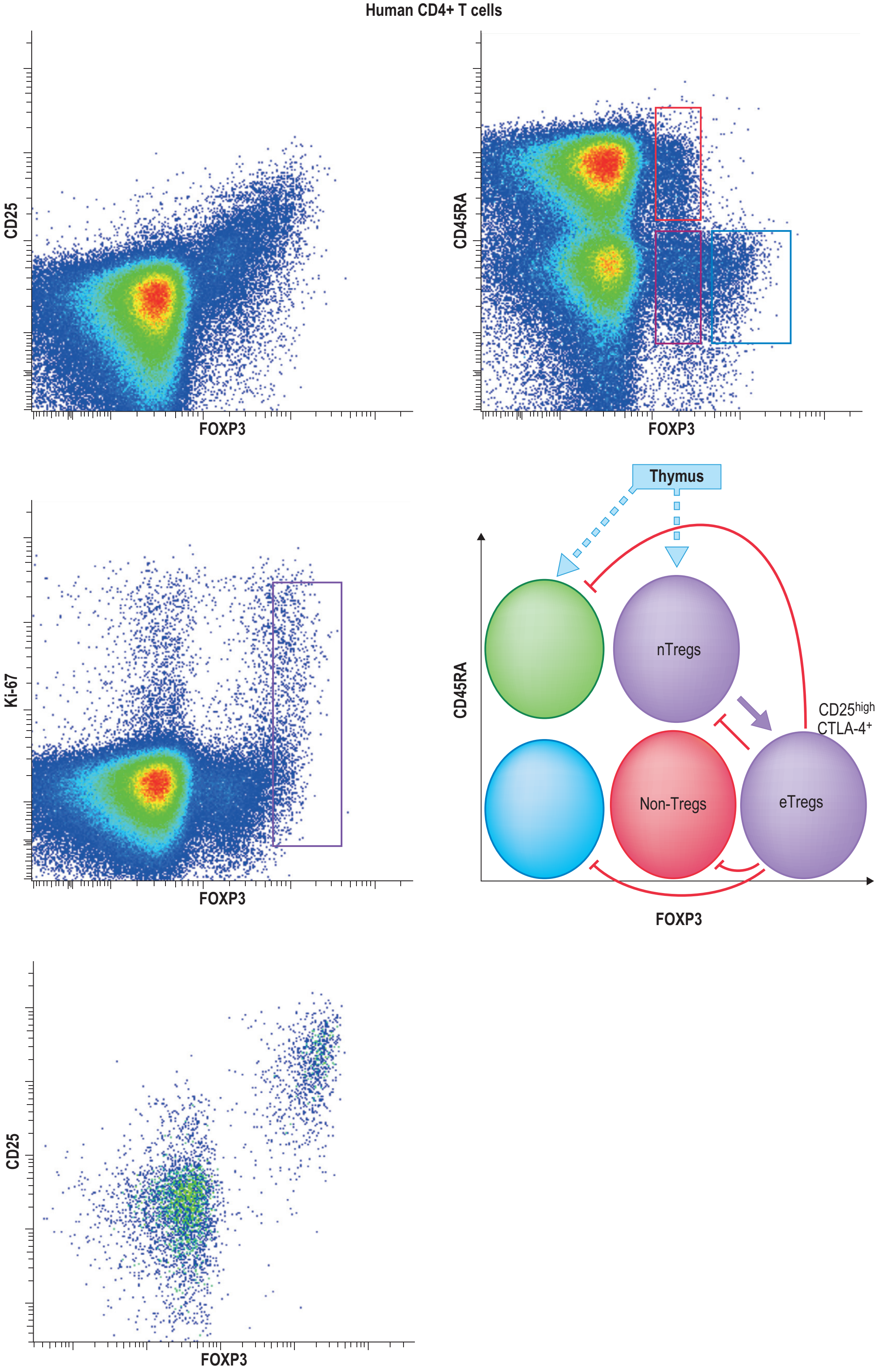


FIG 18.3 **Expression of CD25 and Foxp3 Denotes Tregs in Mice and in Humans.** In mice, between 5% and 10% of CD4$^+$ T cells express CD25, and among these, almost all are Foxp3$^+$. In addition, a few Foxp3-expressing cells reside among CD4$^+$CD25$^-$ T cells. In the peripheral blood of humans, expression of surface marker CD25 is well correlated with intracellular FOXP3, especially when the levels of CD25 and FOXP3 are high *(top left).* CD45RA and intracellular expression of FOXP3 enables the definition of three distinct subsets within FOXP3$^+$CD4$^+$ T cells: CD45RA$^-$ effector Treg cells with high expression of FOXP3 *(blue box),* CD45RA$^+$ naïve Treg cells with low expression of FOXP3 *(red box)*, and CD45RA$^-$ non-Treg cells with low expression of FOXP3 *(top right, purple box;).* Effector FOXP3^{high} Treg cells correspond to actively proliferating Treg cells *in vivo* as they express Ki-67 *(bottom left, purple box).* The thymus produces CD45RA$^+$FOXP3^{low} naïve Tregs as well as naïve CD45RA$^+$ non-Treg cells. Upon antigenic stimulation, naïve Tregs differentiate to CD45RA$^+$FOXP3^{high} effector Tregs, which are potently suppressive and capable of suppressing the proliferation of other Tregs as well as non-Tregs *in vitro (bottom right).*

CD25$^+$CD4$^+$ T cells and CD25$^+$CD4$^+$CD8$^-$ thymocytes specifically express *Foxp3* mRNA in contrast to the cell surface markers used until now. In addition, other thymocytes/T cells, Th1, or Th2 cells scarcely express *Foxp3* even after stimulation.[9] Intranuclear staining of the Foxp3 protein shows that although the majority of Foxp3$^+$ cells in mice reside in the CD4$^+$CD25$^+$ T-cell population, some can also be found in the CD4$^+$CD25$^-$ population (see Fig. 18.3). Importantly, retroviral transduction of *Foxp3* in naïve CD25$^-$ T cells can convert them to regulatory cells with at least some of the suppressive functions of true Tregs. However, although it is essential, it has also become clear that Foxp3 alone is not sufficient to stably maintain the full Treg identity. Another critical factor is the presence of a Treg-type epigenetic pattern in which genes, such as *Foxp3, CTLA-4,* and *GITR,* have stably demethylated CpG residues and a permissive chromatin structure allowing them to be constitutively expressed by Tregs. These epigenetic modifications occur independently of Foxp3 expression and maintain a significant proportion of the *Treg* gene expression even in the absence of Foxp3 itself.[10] However, at the same time, Foxp3 expression is still essential, as shown by the severe diseases seen in humans or mice lacking functional Foxp3. A similar pattern of FOXP3 expression can be observed in humans, with most FOXP3$^+$ cells among CD4$^+$CD25^{high} T cells, and a few being CD25$^-$ or CD25^{low} (see Fig. 18.3). However, in humans, but not in mice, low levels of FOXP3 can be transiently induced by TCR stimulation in conventional T cells. These T cells can be detected directly in blood as CD4$^+$ T cells with a memory phenotype and a weak expression of FOXP3, but with no suppressive function.[8] Human FOXP3$^+$ Tregs with suppressive function can be divided into two subsets: CD45RA$^+$FOXP3^{low} naïve Tregs (nTregs) and CD45RA$^-$FOXP3^{high} effector Tregs (eTregs) (see Fig. 18.3).

In addition to tTregs, there is abundant evidence from mouse studies supporting the peripheral development of T cells with suppressive properties. For example, Foxp3-expressing CD25$^+$CD4$^+$ Tregs, functionally and phenotypically similar to tTregs, can be induced from naïve T cells by *in vitro* or *in vivo* antigenic stimulation in the presence of TGF-β.[11] However, it should be noted that *in vitro*–induced Tregs lack the Treg-type epigenetic pattern and so may not be stable Tregs, making it important to differentiate between *in vivo* peripherally induced pTregs and *in vitro*-induced iTregs.[10] Recent findings indicate that both murine and human thymus-derived Treg cells express Helios, an Ikaros family transcription factor, and Neuropilin-1, whereas most induced Foxp3-expressing CD4 T cells do not. However, some highly activated pTregs and Foxp3$^-$ T cells express these markers to some extent, making them useful, but not of themselves conclusive, to differentiate between tTregs and pTregs.[12] Recent work has demonstrated that tTregs are sufficient for the prevention of widespread autoimmunity but that pTregs have a more specialized function in the prevention of type-2 immune responses in mucosal sites, such as the gastrointestinal tract and the lungs.[11]

CLINICAL RELEVANCE

Infections and Immune Dysfunction/Polyendocrinopathy/Enteropathy/X-Linked (IPEX) Syndrome Is a Result of FOXP3$^+$ Treg Deficiency

When the *Foxp3/FOXP3* gene has a loss-of-function mutation, FOXP3$^+$ regulatory T cells (Tregs) fail to develop, or Foxp3 protein is dysfunctional, and a fatal autoimmune/autoinflammatory disease develops. This monogenic X-linked disease directly demonstrates how crucial FOXP3$^+$ Tregs are for maintaining self-tolerance and immune homeostasis.

Cardinal features of IPEX are:

- Autoimmune diseases (type 1 diabetes, thyroiditis, hemolytic anemia)
- Allergy (dermatitis, hyperimmunoglobulinemia E, food allergy)
- Inflammatory bowel disease

Maintenance of Foxp3$^+$ Tregs

In addition to TCR interaction, it seems that accessory signals, such as costimulation through CD28-B7 or CD40-CD40L, play an important role in the production of Foxp3$^+$ Tregs in the thymus, since animals that lack CD28 or CD40 expression generate only minute numbers of Foxp3$^+$ T cells in the thymus (Table 18.2).[6]

In the periphery, the maintenance of Foxp3$^+$ Tregs requires antigenic priming and cytokines. It is vital that Foxp3$^+$ Tregs encounter specific antigens to remain in the Treg pool. For example, cell transfer experiments in mice thymectomized on day 3 (d3Tx) have demonstrated that Tregs from donors of the same sex are better at protecting against orchitis or oophoritis compared with Tregs from donors of the opposite sex and that Tregs from ovariectomized mice are less competent in preventing oophoritis compared with those from normal females.[5]

IL-2 is vital for the maintenance of natural Tregs, and accordingly, CD25 is not only a marker but also a functionally indispensable molecule for Tregs as a key component of the high-affinity IL-2 receptor. Mice genetically deficient in IL-2 or IL-2 receptor α chain (CD25) or β chain (CD122) develop severe lymphoproliferative disease with lymphocyte infiltration

TABLE 18.2 Signals With Impact on FOXP3$^+$ Treg Induction, Maintenance, and Suppression

	Development	Maintenance/Survival	Suppressive Function
Peptide MHC II interaction	Yes (high affinity)	Yes	Yes, at least initially
CD28	Yes (crucial)	Yes	Not crucial for induction of suppression but high expression on APCs breaks suppression
CD40	Yes	No	No
CTLA-4	No	No	Yes
GITR	No	Modest positive effect	Breaks suppression
TLR ligands	No	Yes	TLR ligands initially break suppression, but this is followed by induction of enhanced suppression
IL-2	Yes (but not crucial)	Yes (crucial)	High levels break suppression
TGF-β	Not required for thymic differentiation but may be involved in peripheral induction	Yes	Yes (not crucial)

Treg, regulatory T cell; MHC II, major histocompatibility complex II; IL-2, interleukin 2; APC, antigen-presenting cell; TGF-β, transforming growth factor-β; CTLA-4, cytotoxic T lymphocyte antigen 4; GITR, glucocorticoid-induced tumor necrosis factor receptor protein; TLR, toll-like receptor.

in multiple organs, resulting in early death.[13] Moreover, genetic deficiencies of CD25 in humans generate a comparable clinical and pathological pattern. It is now known that IL-2–deficient animals have substantially reduced, although not completely depleted, numbers of Foxp3$^+$ T cells and that disease can be prevented by adoptive transfer of normal Foxp3$^+$ Tregs (see Table 18.2).

In addition to antigen recognition and cytokines, Foxp3$^+$ Tregs require appropriate interactions with antigen-presenting cells (APCs) for their function and survival. Several molecules of cell adhesion and costimulation are important for function and homeostasis of Tregs, including CD18/CD11a, GITR, CD28, and CTLA-4.

Effector CD4 T cells differentiate into various subpopulations, such as Th1, Th2, Th17, and Tfh, in response to specific stimulation and cytokine milieus (Chapter 16). It has become clear that Tregs mirror this process to some extent and can also differentiate in response to the same stimuli into a subpopulation specialized to control the matching effector population. As a result, Tregs express transcription factors linked to these fates, such as T-bet (Th1), RORγt (Th17), and BCL6 (Tfh), and this allows Tregs to gain expression of relevant chemokine receptors, such as CXCR5, in the case of T-follicular regulatory (Tfr) cells, enabling these cells to traverse to the germinal center, where they are able to suppress T-follicular helper (Tfh) cells and the germinal center reaction. Similarly, expression of CCR6 and CXCR3 allows a similar process to occur for Th17-Tregs and Th1-Tregs, respectively. This process lets Tregs respond to the environment and specifically travel to and suppress Th1, Th2, Th17, and Tfh-cell responses.[14] Further to the development of Tregs that specifically mirror effector T-cell subsets, it has also been recently demonstrated that Tregs travel to nonlymphoid sites, such as skin, adipose tissue, and skeletal muscles, to suppress inflammation. Tregs from these sites may represent specialized subtypes of Tregs with adapted transcriptional signatures, such as adipose tissue Tregs, expressing the transcription factor peroxisome proliferator-activated receptor-γ (PPAR-γ), which regulates fatty acid metabolism and allows their homeostasis and survival in this environment.[15]

Another important aspect of Treg responsiveness to the environment is that Tregs are affected by the microbiota (Chapter 14). Recently, it has been demonstrated that certain commensal organisms, such as *Clostridium* spp. present in the gut, have large effects on Treg differentiation in these environments. In some cases, this may be indirect via induction of TGF-β production leading to induction of Tregs from naïve T cells, whereas in other cases, microbial metabolites, such as short-chain fatty acids, directly induce Treg proliferation.[11]

Suppressive Function of Foxp3$^+$ Tregs

The standard assay to analyze the suppressive function of Tregs is to coculture purified cell fractions of T cells with Tregs and then measure the proliferative response upon antigenic stimulation in the presence of APCs. Such assays show that freshly isolated Foxp3$^+$CD25$^+$ Tregs are not able to suppress T-cell responses *in vitro* unless they are first stimulated via the TCR. Once activated, they suppress other CD4 and CD8 T cells, irrespective of their antigen specificities.[7] Foxp3$^+$ Tregs can inhibit proliferation and cytokine production of naïve T cells. They can also suppress the function of already differentiated Th1 and Th2 cells *in vitro* and have been shown to reverse ongoing immunopathology, such as colitis, *in vivo*.[11] In addition to suppressing T-cell function, Foxp3$^+$ Tregs are also able to suppress B cells, NK cells, and NKT cells.

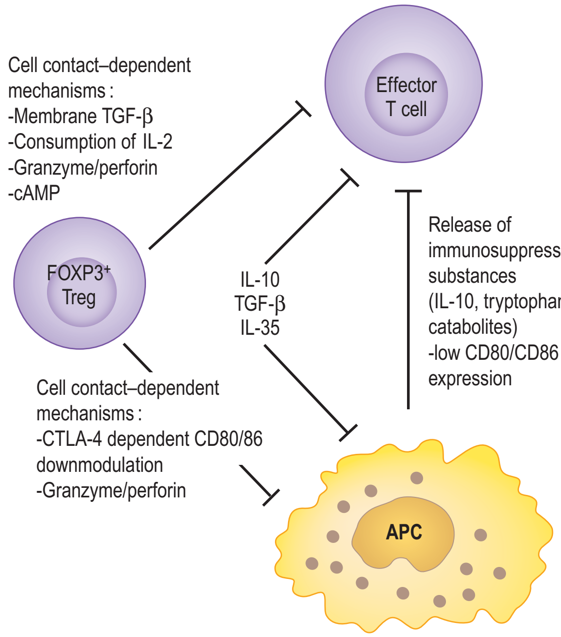

Cell contact–dependent mechanisms:
-Membrane TGF-β
-Consumption of IL-2
-Granzyme/perforin
-cAMP
Effector T cell
FOXP3+ Treg
IL-10
TGF-β
IL-35
Release of immunosuppressive substances
(IL-10, tryptophan catabolites)
-low CD80/CD86 expression
Cell contact–dependent mechanisms:
-CTLA-4 dependent CD80/86 downmodulation
-Granzyme/perforin
APC

FIG 18.4 Proposed Suppressive Mechanisms of Foxp3$^+$ Regulatory T Cells (Tregs). Foxp3$^+$CD25$^+$ Tregs can inhibit many different types of effector cells and can also suppress immune responses at multiple stages, including the initial priming in the lymph node and effector actions at the site of inflammation. The precise mechanism is not known, but numerous theories have been proposed, and it is likely that Foxp3$^+$CD25$^+$ Tregs suppress by several different mechanisms. *In vivo* Tregs can act in a cell contact–dependent manner by competing directly for stimulatory ligands on the antigen-presenting cell (APC), by absorbing essential growth factors, such as interleukin-2 (IL-2), or by directly transmitting an as-yet uncharacterized negative signal to either the T cell or the APC. Alternatively, Tregs can use long-range suppressive mechanisms by means of the cytokines IL-10, IL-35, and transforming growth factor-β (TGF-β).

Foxp3$^+$ Tregs likely suppress by multiple mechanisms, including both soluble and cell surface–bound mediators (Fig. 18.4). *In vitro* studies have shown that Foxp3$^+$ Tregs need direct cell-to-cell contact with responder cells and that suppression does not occur if Tregs are separated from effector T cells by a semipermeable membrane, even if this allows the passage of soluble factors.[7] Moreover, Foxp3$^+$CD25$^+$ Tregs are not prominent producers of either IL-10 or TGF-β *in vitro*. These features are quite different from those of Tr1 or Th3 cells, which primarily rely on soluble immunosuppressive cytokines, such as IL-10 and TGF-β, for their inhibitory function.

Although Foxp3$^+$ Tregs have been shown to suppress effector T cells under APC-free conditions *in vitro*, presumably by absorbing IL-2, it is reasonable that Foxp3$^+$ Tregs *in vivo* control immune responses also by regulating APCs. Indeed, by using two-photon laser-scanning microscopy, it has been shown *in vivo* that although there are limited contacts between Foxp3$^+$ Tregs and effector T cells, stable interactions exist between Foxp3$^+$ Tregs and dendritic cells (DCs) during ongoing suppression in lymph nodes. One way that Foxp3$^+$ Tregs regulate immune responses may be through competition with effector T cells for access to APCs. Foxp3$^+$ Tregs have, relative to conventional T cells, a more activated phenotype (*e.g.*, high expression of adhesion molecules, such as LFA-1) in normal healthy individuals, which gives an advantage when engaging with APCs and results in the prevention of naïve T-cell priming. Interestingly, Foxp3$^+$ Tregs have been shown to modify the function of the APC. APCs cultured with Foxp3$^+$ Tregs downregulate CD80 and CD86 via a CTLA-4–dependent mechanism and become impaired in their ability to stimulate T cells.[16] Additionally, CTLA-4–expressing Tregs can induce production of the enzyme indoleamine dioxygenase (IDO), which catalyzes the amino acid tryptophan to the metabolite kynurenine that represses T-cell responses. Importantly, CTLA-4 expression by Foxp3$^+$ Tregs has been shown to be critical for tolerance *in vivo,* since mice with deletion of this coinhibitory molecule only in Foxp3-expressing cells develop lethal autoimmunity. Other proposed suppressive mechanisms that involve close contact between Foxp3$^+$ Tregs and target cells include surface-bound TGF-β, which is inhibitory at a close range; perforin; and granzyme B, which kills target cells (see Fig. 18.4).[7]

Despite the fact that immunosuppressive cytokines are redundant for suppression *in vitro,* the *in vivo* situation seems somewhat different. Recently, the immune suppressive cytokine IL-35 has been implicated in Treg-mediated suppression, and both TGF-β and IL-10 expressed by Foxp3$^+$ Treg are important in the prevention of IBD in mice. Curiously, adoptive transfer of IL-10–deficient CD25$^+$ Tregs failed to protect against colitis, although it could inhibit the development of gastritis. Furthermore, although perforin- or granzyme B–expressing Foxp3$^+$ Tregs are rare in the spleen, they are abundant in a tumor environment. Taken together, Foxp3$^+$ Tregs can use several different mechanisms of suppression, and one seems to play a more prominent role than another, depending on the local cytokine milieu and the strength and type of immune response.[7]

Tr1 Cells

Some Foxp3$^-$ suppressive T cells have also been identified. The best characterized are IL-10–secreting Tr1 cells, which have been shown to be present and functional in humans following bone marrow transplantation and in response to allergens (see Table 18.1).

Tr1 cells were initially generated *in vitro* from CD4 T cells rendered anergic by chronic stimulation in the presence of IL-10, which is a potent negative regulator of inflammation and lymphoproliferation.[17] T cells obtained in such a fashion produce a unique pattern of cytokines distinct from Th1 or Th2 cells, with IL-10 as their signature cytokine. In addition, Tr1 cells secrete some TGF-β, interferon-γ (IFN-γ), and IL-5, but not IL-4 or IL-2. Tr1 cells can be identified on the basis of surface expression of CD49b and LAG-3, but lacking Foxp3 expression.[17] In addition, antigenic stimulation with immature DCs (*i.e.,* low levels of costimulatory molecules) or with DCs pretreated with IL-10/TGF-β confers on naïve CD4 T cells an anergic and suppressive phenotype both *in vitro* and *in vivo.*[17,18]

Several investigations have shown that Tr1 cells take part in the tolerance process in humans. For example, the presence of Tr1 cells can be correlated with lack of graft-versus-host disease (GvHD) in bone marrow transplantation, and Tr1 cells are also induced after specific immunotherapy in patients with allergies. Experimental data from a murine model shows that Tr1 cells can prevent IBD. Interestingly, Tr1 cells specific for *Escherichia coli* proteins can be isolated from the intestinal mucosa of healthy donors. In pemphigus vulgaris, which is an autoimmune skin disorder with circulating autoantibodies against desmoglein-3 (Chapter 63), Tr1 cells specific for desmoglein-3 can be isolated from pemphigus vulgaris–prone but apparently healthy individuals, whereas patients with pemphigus rarely have such cells. Collectively, Tr1 cells can be induced by autoantigens as well as foreign antigens as a component of the mechanisms that maintain tolerance to both self and nonself.[18]

OTHER SUBSETS OF FOXP3$^-$ REGULATORY T CELLS

Th3 cells were identified in mice during investigations of the mechanisms of oral tolerance. Oral tolerance has presumably evolved to prevent hypersensitivity reactions to food and microbial antigens present in the mucosal flora. It was found that mice fed with myelin basic protein, a neuronal autoantigen, developed T cells that preferentially produced TGF-β together with varying amounts of IL-4 and IL-10. These T cells had suppressive properties, mediated by TGF-β, and could prevent the induction of experimental allergic encephalomyelitis (EAE), the murine equivalent of multiple sclerosis (MS) (Chapter 66). Th3 cells can be identified by their surface expression of latency-associated peptide (LAP), which binds inactive TGF-β. Recently, LAP$^+$ Tregs have been shown to be induced after oral or nasal administration of tolerogenic CD3 monoclonal antibodies. Th3 cells can also be generated *in vitro* if grown in the presence of TGF-β, IL-4, and IL-10, although some of them may become Foxp3$^+$ iTregs. Since the intestinal mucosa has high basal levels of all these cytokines, which are upregulated after antigen administration, it is conceivable that this particular environment drives the formation of Th3 cells whereas a different setting may produce Tregs of another phenotype.[19,20]

In addition to CD4 Tregs, other types of T cells with immunosuppressive properties have also been found; these recognize antigens different from those typically presented to CD4 T cells via MHC class II and may therefore serve to induce tolerance in other settings (see Table 18.1). One example is CD8 T cells with TCRs that recognize antigen presented by the mouse MHC class lb Qa-1 molecule (HLA-E in humans; Chapter 5). Qa-1 has limited polymorphisms and can present both foreign and self-peptides. Since Qa-1 peptide complexes on, for example, CD4 T cells can bind to both the inhibitory CD94-NKG2 receptor complex and the TCR on CD8 T cells, total loss of Qa-1 does not lead to development of spontaneous autoimmunity, since loss of CD94-NKG2 inhibitory signaling is largely compensated for by loss of TCR signaling.[21] However, specific mutation of Qa-1 leading to its loss of CD94-NKG2, but not TCR binding, results in severe lupus-like autoimmunity underpinned by dysregulated Tfh-cell responses. The function of Qa-1–restricted CD8 appears to be largely dependent on perforin expression as perforin-deficient CD8 T cells fail to mediate the suppression of Tfh cells.[21]

Another subset of regulatory CD8 T cells are CD8$^+$CD28$^-$ Tregs, which can suppress immune responses by direct interaction with APCs. These cells are generated by multiple rounds of *in vitro* stimulation with alloantigen and can downregulate costimulatory molecules or upregulate the inhibitory immunoglobulin-like transcript 3 (ILT3) and ILT4 receptors on DCs, leading to the impaired activation of effector T cells.[22]

Double-negative T cells are a class of αβ TCRs expressing cells that lack both CD4 and CD8 and make up around 1% of the TCR αβ cells in mice and humans.[22] Much is still unknown about the action of these cells, but it seems that they may be able to suppress CD4 and CD8 T cells by taking up antigen–MHC complexes from APCs via trogocytosis, capturing plasma membrane fragments from the immunological synapse and then presenting them to other T cells in the absence of costimulatory signaling and possibly a signal-inducing apoptosis, such as Fas. They may also suppress APCs in a manner similar to Foxp3$^+$ Tregs by expression of CTLA-4.[22]

γδ T cells with a regulatory phenotype exist as a subset of the epithelial γδ T cells, which can be found in mice. Mice deficient in γδ T cells do not appropriately regulate responses to various pathogens. This inappropriate regulation manifests as immunopathology in conjunction with the robust development of immunity. γδ T cell–deficient mice also show accelerated autoimmune responses in models of systemic lupus erythematosus (SLE) and spontaneously develop dermatitis when bred on certain genetic backgrounds. Commonly, these conditions are driven by αβ T cells, and γδ T cells will inhibit αβ T cells predominantly in the local environment. In humans, who lack an equivalent population of intraepithelial γδ cells, it is plausible that this immune regulation is provided by other types of suppressive cells.[23]

NKT cells respond to CD1d, the nonclassic class I antigen-presenting molecule, which binds glycolipids rather than peptides. NKT cells can induce either proinflammatory (IFN-γ) or anti-inflammatory (IL-4, IL-10, IL-13) immune responses, but the prerequisites for this choice are ill defined. Nevertheless, under appropriate conditions, NKT cells clearly promote tolerance, which is illustrated in studies of transplantation and oral tolerance (see Table 18.1).[24]

Suppressive Non–T Cells

Recently, interest has also been focused on non–T-cell subsets with suppressive functions. One example of this is B cells with regulatory function (known as *Breg* or *B10 cells*). Breg cells produce the suppressive cytokine IL-10 and have been identified in both mice and humans. Experiments in mice have demonstrated that these cells have suppressive capacity and can influence the development of autoimmune diseases, such as EAE, collagen-induced arthritis, and colitis. Breg cells appear to be induced in the periphery; to date, a clear transcriptional controller providing a clear transcriptional program has not been identified. This, together with the finding that various subsets of Breg cells differentiate from a wide variety of different B-cell populations (marginal zone, plasma blasts, plasma cells, CD5+CD1d^{high} B10 cells), has led to the suggestion that Breg cells may not be a lineage in themselves but, rather, a suppressive state that can be induced in all B cells when given the correct stimuli.[25] The suppressive function of Breg cells has largely been attributed to production of IL-10, TGF-β, and IL-35; this may either be via direct action on effector CD4 and CD8 T cells or indirectly by induction of other suppressive populations, such as Foxp3$^+$Tregs and Tr1 cells.[25]

Myeloid-derived suppressor cells (MDSCs) are broadly divided into polymorphonuclear MDSCs, which are closely related to neutrophils, and monocytic MDSCs, which are related to monocytes. MDSCs are generated in bone marrow and can be differentiated from other myeloid cells by the high levels of expression of NADPH oxidase (Nox2) and nitric oxide synthase 2 (*nos2*), leading to production of reactive oxygen species (ROS) and nitric oxide (NO) and the transcription factor c/EBPβ. MDSCs have decreased phagocytic capacity, which, in combination with their production of suppressive cytokines, such as IL-10 and TGF-β, leads them to have a suppressive effect on T-cell responses. MDSCs are often associated with tumor progression and are believed to play an important role in the establishment of the immunosuppressive tumor environment.[26]

DCs have also been demonstrated to have tolerogenic properties in certain circumstances. It is unclear if tolerogenic DCs are of a stable lineage or, perhaps like Bregs, represent a particular state of differentiation. Antigen presentation by immature DCs may, indeed, be tolerogenic because of lack of costimulation, but antigen presentation by the same cells may be immunogenic once mature and expressing greater levels of costimulatory molecules. Additionally, both pDCs found in the tumor microenvironment and CD103$^+$ conventional DCs in the lamina propria produce the immune suppressive molecule indoleamine 2,3-dioxygenase (IDO), which has been demonstrated to aid in the induction of pTreg cells.[27]

CLINICAL RELEVANCE OF REGULATORY T CELLS

Abundant evidence strongly supports Tregs as key controllers of self-tolerance, and Tregs of various subsets play an active role in the control of almost all types of physiological and pathological immune responses, which also makes them suitable targets for immunotherapy (see Table 18.1; Table 18.3).

THERAPEUTIC PRINCIPLES

Adjustment of the Immune Response by FOXP3$^+$ Regulatory T Cell (Treg)

- Reduction of FOXP3$^+$ Treg suppression or reducing Treg numbers
  - Enhancement of tumor immunity
  - Clearance of infections
  - Improvement of responses to vaccines
- Boosting of FOXP3$^+$ Treg function or increasing Treg numbers
  - Treatment of autoimmunity
  - Treatment of allergic responses
  - Induction of transplantation tolerance
  - Control of excessive immunopathology to foreign antigens (*i.e.*, pathogens)
  - Maintenance of fetomaternal tolerance in pregnancy

TABLE 18.3 Potential Therapeutic Approaches for Treg-Based Therapy

Increase of Treg Numbers or Function	Reduction of Treg Numbers or Function
Ex vivo expansion of pure FOXP3$^+$ Tregs with allo- or autoantigens plus growth factors, such as IL-2, and chemicals, such as rapamycin	Transient reduction of FOXP3$^+$ Treg and/or perturbation of suppression *in vivo* (anti-CD25 antibody, anti-CTLA-4 antibody, or anti-IL-2 antibody)
Ex vivo induction of Treg from conventional T cells by cytokines (IL-10, TGF-β), pharmacological agents or modified DCs	Render effector T cells resistant to suppression (GITR signaling)
In vivo promotion of Tregs, rather than effector T cells, using monoclonal antibody treatment or pharmacological agents (anti-CD3 antibody, anti-CD40L antibody, etc.)	

Treg, regulatory T cell; IL-2, interleukin-2; IL-10, interleukin-10; TGF-β, transforming growth factor-β; DC, dendritic cell; CTLA-4, cytotoxic T lymphocyte antigen-4; GITR, glucocorticoid-induced tumor necrosis factor receptor protein.

Autoimmunity

As discussed above, natural Foxp3$^+$CD4$^+$ Tregs are engaged in active suppression of autoimmune disease, since their depletion results in spontaneous development of autoimmune diseases in rodents. Furthermore, genetic anomaly of Foxp3 function can be a direct cause of autoimmune diseases in humans, as exemplified in IPEX syndrome.[6] A reduction of Foxp3$^+$ Tregs in number or function has been reported in systemic autoimmune diseases, such as SLE, Sjögren syndrome, antineutrophil cytoplasmic antibody (ANCA)–associated vasculitis, Kawasaki disease, systemic sclerosis, psoriasis, autoimmune hepatitis, myasthenia gravis, and IBD.[28] Of note, type II autoimmune polyglandular syndrome, which resembles the systemic disease induced in nude mice reconstituted with splenocytes devoid of CD4$^+$CD25$^+$ T cells, is also characterized by Treg functional deficiency. However, studies of MS and type 1 diabetes did not detect any differences between patients and controls, and conflicting data have been reported in rheumatoid arthritis (RA) with regard to both function and numbers of CD25^{high} Tregs.[28] The accuracy of the identification of Tregs must be taken into account in these human studies, since CD25 and Foxp3 are also expressed to some extent by activated non-Treg T cells in humans.

A general observation is that FOXP3$^+$ Tregs increase in number at the site of inflammation. In the case of RA, synovial fluid from patients with ongoing RA was found to contain increased numbers of Foxp3$^+$CD25$^+$ Tregs compared with the levels in peripheral blood. CD25$^+$ Tregs from the synovial fluid of patients with RA are largely functional, although their numbers or suppressive function is apparently not enough to halt the inflammatory process. In contrast, CD25$^+$ Tregs obtained from the blood of patients with MS reportedly have decreased abilities to suppress proliferation of effector T cells. In summary, reduced levels of FOXP3$^+$CD25$^+$ Tregs in peripheral blood are not a general finding in autoimmune diseases and may not necessarily reflect the actual conditions at the site of inflammation. However, if Treg subsets defined by expression levels of CD45RA and Foxp3 are assessed, cytokine secreting Foxp3^{low} non-Tregs increase in active SLE.[8] Dynamic changes of Treg subsets in various autoimmune states remain to be investigated.

Allergic Disease

CD25$^+$ natural Tregs play an important role in suppressing the development of allergic reactions to innocuous environmental substances. This is best illustrated by IPEX syndrome, which is accompanied not only by organ-specific autoimmune diseases but also by severe eczematous dermatitis, high levels of serum immunoglobulin E (IgE), and sometimes eosinophilia.[6] Indeed, FOXP3$^+$CD25$^+$ Tregs from the blood of healthy donors without allergies suppress both proliferation and production of Th2 cytokines when challenged with specific allergens *in vitro*. If the same experiment is performed with Tregs from individuals with allergies, a marked difference is seen, as these Tregs fail to downregulate Th2-related responses to allergens. IL-10 production by the FOXP3$^+$ Tregs, Tr1 and Breg, has been implicated in the control of allergic responses.[29]

Since the suppressive ability for polyclonal stimuli is retained in patients with allergies, this deficiency is directly related to the allergen the individual is sensitized to and thus probably does not reflect a general Treg deficiency. The inability to suppress Th2 responses induced by birch or grass pollen is aggravated when the allergic reaction is ongoing and effector cells are fully activated, as during spring and summer compared with wintertime. Addition of IL-4 can attenuate the CD25$^+$ Treg-mediated suppression of Th2 clones *in vitro* in the same way as IL-2, which may provide an explanation for the insufficient control of ongoing allergic responses by Tregs. Individuals with or without allergies harbor allergen-specific, IL-4–producing effector T cells, IL-10–producing Tr1 cells, and CD25$^+$ Tregs, but in different proportions. The balance between Th2 and certain Treg populations may, therefore, dictate whether clinical allergy will develop. Indeed, in the setting of curative specific allergy immunotherapy (SIT), allergen-specific IL-10–producing T cells can be induced. Furthermore, children who “grow out” of their allergy to cow’s milk have higher numbers of CD4$^+$CD25$^+$ T cells specific for β-lactoglobulin compared with children with clinically active allergies. This suggests that certain allergies can be cured by the induction or expansion of antigen-specific Tregs and that the balance between Tregs and effector T cells is of importance to prevent allergeis.[30]

Transplantation

The ultimate goal of organ transplantation is to establish tolerance to allogeneic organ grafts as effectively and stably as to self-tissues, but without the need for continuous general immunosuppression (Chapter 81).[31] Needless to say, Tregs have sparked a lot of attention in this area of research. CD25$^+$ Tregs were first shown to suppress GvHD in murine models of allogeneic bone marrow transplantation. Similarly, nude mice transplanted with allogeneic skin rejected the graft when reconstituted with CD4$^+$CD25$^-$ T cells alone but retained the graft when large enough numbers of CD25$^+$ Tregs were cotransferred. In humans, attempts have been made to prevent GvHD in bone marrow transplantation and induce graft tolerance in organ transplantation with the use of purified CD25$^+$ Tregs.[17] Another potential way to promote the induction of Tregs in organ transplantation is to evaluate the effects of various immunosuppressants in terms of altering the balance between Tregs and effector T cells. Different immunosuppressants target different pathways in cell metabolism and can therefore have different effects on cell populations that behave in dissimilar ways, such as effector T cells and Tregs. Dosage and timing of administration, as well as specific drug combinations, seem to be a promising angle of transplantation immunotherapy with the purpose of inducing graft tolerance and preventing graft rejection. Tregs have been shown to home to, and reside within, the graft, which is stably accepted once a dominance of Tregs has become established. Treg-mediated transplantation tolerance is not a systemic phenomenon but, rather, is localized to the graft, and as such, it would not incur the dangers that accompany general immunosuppression.[31]

Tumor Immunity

It is now well known that many of the tumor-associated antigens recognized by a patient’s T cells are normal self-constituents, indicating that antitumor immune responses are within the range of FOXP3$^+$ Treg control. Therefore the presence of natural Tregs in the normal immune system may not only prevent autoimmunity but also hamper immune surveillance of cancer.[32] In fact, FOXP3$^+$ Tregs, in particular, highly suppressive FOXP3^{high} effector Tregs, are abundant in the tumor mass where they likely block any immune response targeting malignant cells. Studies on human malignancies have shown that the frequency of FOXP3$^+$ Tregs is increased in tumors of, for example, metastatic melanoma or pancreatic and lung cancers. Moreover, high levels of FOXP3$^+$

Tregs in the tumor are correlated with a poor prognosis and survival. Treg cells are not only involved in solid tumors but also in hematological malignancies.[32] For example, the architectural pattern of Treg cells in follicular lymphomas is associated with the prognosis of the disease. Whether the elevated levels result from migration of Tregs into the tumor or from an expansion on the site is not clear, but evidence exists in support of both events. For example, ovarian tumor cells and infiltrating macrophages secrete the Treg-recruiting chemokine CCL22, which binds to CCR4 expressed by Tregs, and, in addition, many tumors produce TGF-β, which contributes to the maintenance of FOXP3$^+$ Tregs and may also induce FOXP3 expression in non-Treg cells within the tumor microenvironment. It is now evident that both effector T cells and Tregs must be assessed in monitoring the efficacy of anticancer immunotherapy.[32]

The involvement of Tregs in tumor immunity indicates that antitumor immune responses can be provoked or enhanced by depleting Tregs in a host that is otherwise responding poorly. Experimental mouse models have, indeed, demonstrated that simple depletion of CD25$^+$ Tregs with anti-CD25 antibody results in tumor eradication, and similar effects can be achieved with *in vivo* administration of agonistic anti-GITR, anti-CTLA-4–blocking antibodies or anti-CCR4 cell–depleting antibodies.[32] Depletion of CD25$^+$ T cells also enhances the effect of vaccination with tumor antigens.

Pharmacological agents are another possible way of altering the effector T cell/Tregs ratio. For example, fludarabine was shown to selectively decrease the frequency of CD25$^+$ Tregs in patients receiving chemotherapy. Conversely, previously used regimens, such as administration of exogenous IL-2, are now being reevaluated, since IL-2 may expand Tregs. As expected from the role of Tregs in self-tolerance, a caveat of Treg-based therapies of cancer is the possible development of autoimmunity that may depend on the degree and period of *in vivo* systemic Tregs depletion, as well as the genetic make-up of the host.[32]

In addition to Foxp3-expressing Tregs, MSDCs are believed to have an important role in the establishment and maintenance of the tumor immunosuppressive environment, largely through their production of immunosuppressive cytokines, such as IL-10 and TGF-β, and reactive nitrogen and oxygen species.[26]

Infectious Disease

Immune responses to infectious agents, such as bacteria and viruses, often result in tissue damage, which might be more severe if it were not for the involvement of Tregs. On the downside, in many cases, Tregs can contribute to the development of chronic infections. As previously discussed, Tregs have the potential to directly respond to microbial products and are believed to be engaged in suppressing responses to infectious agents. A number of studies have shown that the outcome of an infection partly hinges on the proper balance between effector T cells and Tregs.[33] Adoptive transfer of Tregs prevents lethal pneumonia in T cell–deficient mice infected with *Pneumocystis jiroveci,* but at the expense of a deficient protective response and microbial clearance. Similarly, Tregs suppress Th1 responses in mice infected with *Helicobacter pylori,* thereby limiting the mucosal inflammation but resulting in a higher bacterial load. Human studies have shown that Tregs from carriers of *H. pylori* suppress responses to *H. pylori* antigens *in vitro* and have increased frequencies of CD25$^{\text{high}}$ T cells in both the stomach and the duodenal mucosa compared with Tregs from healthy controls. Taken together, modulation of the infectious response by Tregs can limit tissue damage but may enhance pathogen survival. This sort of compromise may not always be disadvantageous to the host. For example, in murine *Leishmania major* infection, Tregs prevent complete eradication of the parasites, which results in the persistence of low numbers of microbes that have been proven to be essential for the development of T-cell memory and prevention of reinfection. However, this delicate balance can be tipped in favor of the pathogen, which can be seen in the case of malaria and various viral infections such as human immunodeficiency virus (HIV). For example, HIV-specific CD4 and CD8 T-cell responses are substantially suppressed by Tregs *in vitro* in most individuals with HIV infection. Taken together, future treatments, as well as vaccine design, will need to take Tregs into account, and depending on the pathogen in question, it might be necessary to reduce or enhance the activity of Tregs to achieve a favorable outcome (see Table 18.3).[33]

TRANSLATIONAL RESEARCH

Manipulations of Foxp3$^+$ Tregs in animal models of, for example, autoimmune disease, cancer, transplantation, and infection, have shown a great potential for modulation of immunological diseases by this subset. The findings in animals have already started to make their way into clinical practice and more is likely to come within the next 5–10 years.[2] In cases when immunity needs to be boosted, such as in cancer and during infection, the approach is to identify molecules that inhibit function and differentiation of Tregs as well as molecules that locally deplete them. Particularly, cancer immunotherapy has proven to be successful[34] (Chapter 77). Such examples in current therapy are CTLA-4–specific blocking antibody (ipilimumab; MDX-010/Yervoy; Bristol-Myers Squibb, N.Y.); although CTLA-4 expression is not exclusive to Tregs, it is much more highly expressed and critical to their suppressive function. In addition, recent studies have suggested that in addition to blocking the function of CTLA-4, anti-CTLA-4 antibodies may also deplete Tregs in the tumor environment, but not in the periphery, because of antibody-dependent, cell-mediated cytotoxicity. This depletions specificity to the tumor environment may be attributed to a combination of high levels of CTLA-4 expression and the presence of large numbers of phagocytic cells on the local area.[32] Furthermore, other checkpoint blockade strategies, such as those targeting the cell intrinsic immunosuppressive molecule PD-1 or its ligand PD-L1, have proven to be effective in early clinical trials.[34] Additionally, combination therapies are of great interest. It has been found that anti-CTLA-4 treatment can lead to increased expression of PD-L1 by the tumor itself, which may dampen the effect. Thus adding both anti-CTLA-4 and anti-PD1 or PD-L1 may have greater benefits, as demonstrated by a recent phase I trial of anti-CTLA-4 (ipilimumab) plus anti-PD-1 (nivolumab) in patients with melanoma.[35] Certain combinations may prove particularly effective, since they target different pathways. Again, in this case, CTLA-4 may be acting primarily as a cell extrinsic suppressive molecule on Tregs, and PD-1 may be acting in *cis* to suppress PD-1–expressing CD4 or CD8 effector cells, and as a result, the combination of the two may provide synergistic effects. As might be predicted, these types of immunomodulatory treatments can lead to severe autoimmune adverse events; however, further studies on both CTLA-4 and PD-1 and other potential checkpoints, such as B- and T-lymphocyte attenuator (BTLA), inducible T-cell costimulator (ICOS), and a range of other inhibitor or stimulatory molecules,[34] either as monotherapies or in combination, may help boost efficacy while reducing adverse events.[34]

Conversely, in cases when tolerance should be reinforced (autoimmunity, transplantation, allergy, fetomaternal tolerance) molecules that either mimic or enhance Tregs function and survival are under investigation. In the former case, CTLA-4-Ig (abatacept; Orencia, Bristol-Myers Squibb, N.Y.), a solubilized version of CTLA-4 that, in part, mimics the effect of Tregs on APCs by blocking their expression of B7.1 and B7.2, is currently being used for the treatment of, for example, RA. In addition, cellular therapy, by infusion of *ex vivo* or manipulated expanded Tregs, is in small-scale use in patients with GvHD after hematopoietic stem cell transplantation,[17] and early trials of Treg transfer in recent-onset type 1 diabetes are underway. Further to this, transfer of Tr1 cells in patients with Crohn disease is also under investigation.[17] Cell-based therapies are, indeed, a promising strategy for managing various immunological diseases and immune responses; however, several challenges still remain, particularly those relating to the functional stability of *ex vivo*–expanded or –induced Tregs.[2]

ON THE HORIZON

- It is clear that regulatory T cells (Tregs) in different sites have phenotypic differences, a greater understanding of Treg heterogeneity is critical to clarification of their role in human autoimmunity.
- Specific targeting of tumor-resident Tregs without depletion of peripheral Tregs is critical to enhancing antitumor therapy while avoiding adverse events.
- Expansion and control of Treg subsets, such as T-follicular regulatory (Tfr) or muscle-resident Tregs, may be critical to fine control of specific immunological disorders.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Klein L, Kyewski B, Allen PM, et al. Positive and negative selection of the T cell repertoire: what thymocytes see (and don't see). Nat Rev Immunol 2014;14:377–91.
2. Bluestone JA, Trotta E, Xu D. The therapeutic potential of regulatory T cells for the treatment of autoimmune disease. Expert Opin Ther Targets 2015;1–13.
3. Abbas AK, Benoist C, Bluestone JA, et al. Regulatory T cells: recommendations to simplify the nomenclature. Nat Immunol 2013;14:307–8.
4. Curotto de Lafaille MA, Lafaille JJ. Natural and adaptive foxp3+ regulatory T cells: more of the same or a division of labor? Immunity 2009;30:626–35.
5. Sakaguchi S. Regulatory T cells: history and perspective. Methods Mol Biol 2011;707:3–17.
6. Wing K, Sakaguchi S. Regulatory T cells exert checks and balances on self-tolerance and autoimmunity. Nat Immunol 2010;11:7–13.
7. Yamaguchi T, Wing JB, Sakaguchi S. Two modes of immune suppression by Foxp3(+) regulatory T cells under inflammatory or non-inflammatory conditions. Semin Immunol 2011;23:424–30.
8. Sakaguchi S, Miyara M, Costantino CM, et al. FOXP3+ regulatory T cells in the human immune system. Nat Rev Immunol 2010;10:490–500.
9. Zheng Y, Rudensky AY. Foxp3 in control of the regulatory T cell lineage. Nat Immunol 2007;8:457–62.
10. Kitagawa Y, Ohkura N, Sakaguchi S. Epigenetic control of thymic Treg-cell development. Eur J Immunol 2015;45:11–16.
11. Bollrath J, Powrie FM. Controlling the frontier: regulatory T-cells and intestinal homeostasis. Semin Immunol 2013;25:352–7.
12. Lin X, Chen M, Liu Y, et al. Advances in distinguishing natural from induced Foxp3(+) regulatory T cells. Int J Clin Exp Pathol 2013;6:116–23.
13. Bayer AL, Pugliese A, Malek TR. The IL-2/IL-2R system: from basic science to therapeutic applications to enhance immune regulation. Immunol Res 2013;57:197–209.
14. Wing JB, Sakaguchi S. Multiple Treg suppressive modules and their adaptability. Front Immunol 2012;3:178.
15. Burzyn D, Benoist C, Mathis D. Regulatory T cells in nonlymphoid tissues. Nat Immunol 2013;14:1007–13.
16. Walker LSK. Treg and CTLA-4: two intertwining pathways to immune tolerance. J Autoimmun 2013;45:49–57.
17. Gregori S, Passerini L, Roncarolo MG. Clinical outlook for type-1 and FOXP3(+) T regulatory cell-based therapy. Front Immunol 2015;6:593.
18. Battaglia M, Gregori S, Bacchetta R, et al. Tr1 cells: from discovery to their clinical application. Semin Immunol 2006;18:120–7.
19. Peterson RA. Regulatory T-cells: diverse phenotypes integral to immune homeostasis and suppression. Toxicol Pathol 2012;40:186–204.
20. Weiner HL, da Cunha AP, Quintana F, et al. Oral tolerance. Immunol Rev 2011;241:241–59.
21. Kim HJ, Cantor H. Regulation of self-tolerance by Qa-1-restricted CD8(+) regulatory T cells. Semin Immunol 2011;23:446–52.
22. Ligocki AJ, Niederkorn JY. Advances on non-CD4 + Foxp3+ T regulatory cells: CD8+, type 1, and double negative T regulatory cells in organ transplantation. Transplantation 2015;99:1553–9.
23. Hayday A, Tigelaar R. Immunoregulation in the tissues by gamma/delta T cells. Nat Rev Immunol 2003;3:233–42.
24. Rossjohn J, Pellicci DG, Patel O, et al. Recognition of CD1d-restricted antigens by natural killer T cells. Nat Rev Immunol 2012;12:845–57.
25. Rosser EC, Mauri C. Regulatory B cells: origin, phenotype, and function. Immunity 2015;42:607–12.
26. Kumar V, Patel S, Tcyganov E, et al. The nature of myeloid-derived suppressor cells in the tumor microenvironment. Trends Immunol 2016;37:208–20.
27. Kim SJ, Diamond B. Modulation of tolerogenic dendritic cells and autoimmunity. Semin Cell Dev Biol 2015;41:49–58.
28. Grant CR, Liberal R, Mieli-Vergani G, et al. Regulatory T-cells in autoimmune diseases: challenges, controversies and—yet—unanswered questions. Autoimmun Rev 2015;14:105–16.
29. Akdis CA, Akdis M. Mechanisms of immune tolerance to allergens: role of IL-10 and Tregs. J Clin Invest 2014;124:4678–80.
30. Soyka MB, Holzmann D, Akdis CA. Regulatory cells in allergen-specific immunotherapy. Immunotherapy 2012;4:389–96.
31. Safinia N, Scotta C, Vaikunthanathan T, et al. Regulatory T cells: serious contenders in the promise for immunological tolerance in transplantation. Front Immunol 2015;6:438.
32. Nishikawa H, Sakaguchi S. Regulatory T cells in cancer immunotherapy. Curr Opin Immunol 2014;27:1–7.
33. Berod L, Puttur F, Huehn J, et al. Tregs in infection and vaccinology: heroes or traitors? Microb Biotechnol 2012;5:260–9.
34. Sharma P, Allison JP. The future of immune checkpoint therapy. Science 2015;348:56–61.
35. Wolchok JD, Kluger H, Callahan MK, et al. Nivolumab plus ipilimumab in advanced melanoma. N Engl J Med 2013;369:122–33.

MULTIPLE-CHOICE QUESTIONS

1. What are the symptoms of loss of FOXP3 function?
   A. Inflammatory bowel disease
   B. Allergic disease (allergic dermatitis, food allergy, hyperimmunoglobulinemia E)
   C. Type 1 diabetes
   D. All of the above

2. A cell is Helios$^+$ Neuropilin$^+$ Foxp3$^+$ CD4$^+$. What is it most likely to be?
   A. tTreg
   B. pTreg
   C. iTreg
   D. Tr1

3. What is the primary suppressive mechanism used by Tr1 cells?
   A. Granzyme B
   B. IL-35
   C. CTLA-4
   D. IL-10

4. What extracellular marker is most commonly used to identify Foxp3 expressing Tregs?
   A. CD27
   B. CD69
   C. CD25
   D. CD28

19

Host Defenses in Skin

Hui Xu, Laura Timares, Craig A. Elmets

Skin, which is the largest organ in the body, is the major interface between the individual and the environment. It comprises 12–16% of the body's weight and chiefly functions to protect against environmental threats. It also plays an important homeostatic role in maintaining body temperature, water and electrolyte balance, and vitamin D production. It acts as a barrier against invading pathogens, exogenous chemicals, and the destructive effects of physical agents, such as the sun, wind, and heat. To perform these various functions, skin consists of three self-regenerating specialized compartments that are layered on top of each other (Fig. 19.1). Adnexal tissues, which include hair, nails, sebaceous glands, and eccrine and apocrine sweat glands, are embedded within these tissues to provide specialized protective functions.

KEY CONCEPT

Primary Functions of Skin

- It is the body's largest organ representing 12–15% of body weight, and spans a surface area of ≈ 2 m^2.
- Because of skin's large surface area, it serves as the principal organ for homeostatic thermoregulation and regulation of fluid retention and evaporation.
- It provides a tactile interface with our environment through an array of specialized sensory cells, which are distributed at very high densities at critical points of the skin surface (*e.g.*, finger tips have up to 2500 sensory cells/cm^2).
- It serves as a photosynthetic surface, utilizing the energy of ultraviolet radiation (UVR) to transform 7-dehydrocholesterol to cholecalciferol (vitamin D3), vitamin D$_3$, an essential step in the biosynthetic pathway of vitamin D.
- It provides a strong physical barrier to withstand physical stresses engendered by exposure to the forces of nature. These include shear forces, extreme temperatures, wind, water, and solar radiation. Granular layer keratinocytes are bound together through tight junctions, providing an effective seal that blocks entry of small molecules and microbes.
- It provides a pharmacological barrier by means of enzymes that detoxify or repair damage caused by chemicals and other carcinogens, such as UVR and ionizing radiation.
- It serves as a regenerative organ that continuously replaces older, damaged skin cells, which are eliminated and shed from skin as anucleated squames.
- It provides an immunological barrier by means of a complex integration of innate and adaptive immune mechanisms that serve to protect against entry by microbial pathogens and to neutralize or eliminate potentially harmful exogenous antigens and endogenous neoplastic cells.

The outermost layer of skin is the epidermis. Blood and lymphatic vessels are absent. However, oxygen and nutrients diffuse to epidermal cells from the microvasculature that is housed in the dermal layer that lies under the epidermis. Approximately 95% of cells within the epidermis are keratinocytes. These form a self-renewing stratified squamous epithelium, which differentiates from bottom to top from a regenerating basal layer containing cuboidal cells to the outer most superficial layer, called the *stratum corneum*, composed of flat, anucleate, compact scales. Keratin intermediate filaments are the major proteins produced by keratinocytes. The type of heterodimer keratins that are made during differentiation impacts the cytoskeletal structure and cell morphology.

Keratinocytes bind to one another through specialized adherence junctions called *desmosomes*. Keratin filaments attach to desmosomal proteins to provide structural, tensile strength. During the differentiation process, the composition of the keratin and desmosomal proteins changes. For example, basal layer keratinocytes synthesize keratins 5 and 14, and the desmosomal protein desmoglein 3 is more abundant than desmoglein 1. In the stratum corneum, keratins 1 and 10 and desmoglein 1 are all highly expressed, but keratins 5 and 14 and desmoglein 3 are absent. Finally, the compact tough outer scales (squames) of the stratum corneum are cross-linked with keratin filaments.

The differences in protein localization between the epidermal layers have consequences for immunologically mediated skin diseases.[1] Pemphigus vulgaris, a disease in which autoantibodies are formed against desmogleins 1 and 3, presents with blisters that originate in the suprabasal layer of the epidermis. Conversely, the closely related disease pemphigus foliaceus, in which autoantibodies recognize only desmoglein 1, results in blisters located exclusively in the upper epidermis.

The remaining cells located in the epidermis are pigment-producing melanocytes, neuroendocrine Merkel cells (important for mechanoreception), and Langerhans cells (LCs) (specialized epidermal antigen presenting cells). Melanocytes are derived from the neural crest, and progenitors seed the epidermis early in development. They synthesize melanin pigment in organelles called *melanosomes*, which mature into the melanin-filled granules, which are then transferred and internalized by keratinocytes. Melanin absorbs the damaging effects of ultraviolet radiation (UVR) and protects skin keratinocytes. When stimulated by cytokines, melanocytes may express several immunologically relevant proteins, including intracellular adhesion molecule 1 (ICAM-1; CD54), costimulatory receptor CD40, and major histocompatibility complex (MHC) class I and II molecules[2] (Chapter 5).

Malignant melanomas arise from melanocytes. These tumors evade the host immune response, at least in part, through expression of programmed death ligand 1 (PD-L1). PD-L1 is an immune

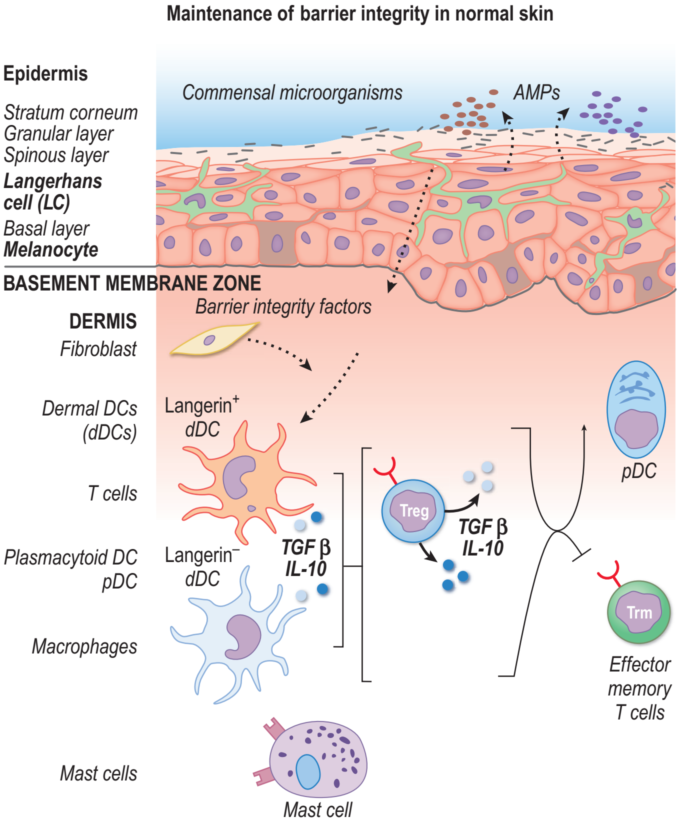


FIG 19.1 Histological Features of Normal Skin. The epidermis is made up of keratinocytes organized into stratified layers. The most superficial layer is the stratum corneum (light pink). The granular layer cells are connected through tight junctions (dark grey). Granular layer cells also secrete antimicrobial peptides (AMPs). The basal layer (tan) contains the proliferative cells that give rise to all the differentiated suprabasal keratinocytes. Basal cells are bound to the basement membrane (black line) and share that scaffold with interspersed melanocytes (brown). Langerhans cells (LCs) (green) are epidermal antigen-presenting cells (APCs) that reside in the suprabasal layer of the epidermis. The dermis contains a number of DC subsets which, under resting conditions, are conditioned by barrier integrity factors. Two of the dermal DC subsets that have been described are langerin$^+$ (orange) and langerin$^-$ (light blue) DCs. Plasmacytoid DCs (pDCs) are also present in the dermis and secrete type I interferons (aqua). Resident memory T cells (T_{RM}) (green) represent previously primed T cells in many lineages (Th1, CTL, Th17, Tc17, Treg) poised to respond when their specific antigens are presented locally, but remain quiescent in uninfected skin. Mast cells (purple) reside near endothelial venules (not shown).

checkpoint molecule that can be found on tumor cells and myeloid suppressor cells. Interaction of PD-L1 with PD-1 receptors on T cells inhibits their antitumor activity. Monoclonal antibodies (mAbs) that block PD-1 interaction with PD-L1 (pembrolizumab and nivolumab) can restore antitumor responses. In clinical trials, these antibodies have shown objective responses and prolonged survival. Ipilimumab, which blocks cytotoxic T lymphocyte antigen-4 (CTLA-4), another immune checkpoint molecule, has comparable efficacy.

The dermis lies beneath the epidermis and serves as a connective tissue layer that provides elasticity and tensile strength to skin. It is filled with a matrix of collagen bundles, elastic fibers, glycoproteins, proteoglycans, and glycosaminoglycans, all of which are produced by dermal fibroblasts. The production of collagen is a dynamic process that involves continual remodeling. Fibroblasts must balance ongoing synthesis of matrix components with the production of matrix metalloproteinases (MMPs) that degrade the matrix. Scleroderma and morphea (localized scleroderma) (Chapter 55) are inflammatory disorders that lead to the overproduction of collagen in the dermis, thus hardening skin.

Other cell types normally found in the dermis are mast cells and diverse sets of tissue macrophages and dendritic cells (DCs).[3] The dermis supports the microvasculature of skin. Endothelial cells of dermal arterioles, capillaries, and postcapillary venules behave differently from those in larger vessels. They express E- and P-selectins, which enable circulating skin-homing T cells to bind through their homing receptors cutaneous lymphocyte-associated antigen (CLA) and P-selectin glycoprotein ligand (PSGL), respectively.[4]

Both homeostatic T-cell trafficking and recruitment of inflammatory leukocytes into skin from the circulation require dermal endothelial cell-specific expression of E- and P-selectins. The egress of leukocytes from blood occurs primarily through postcapillary venules. Egress is regulated, in part, by endothelial responses to cutaneous cytokines and chemokines.

New vessels can also develop in skin. The new vessels can worsen immunologically mediated dermatological diseases and promote tumor development.

The dermis is separated from the epidermis by the basement membrane. This membrane creates a scaffold to which basal keratinocytes and melanocytes are attached. It acts as a selective diffusion barrier that permits passage of necessary small molecule nutrients while retarding entry of macromolecules and inflammatory cells. In patients with systemic lupus erythematosus (SLE) (Chapter 51), antigen–antibody complexes tend to accumulate at the basement membrane. Skin-specific proteins in the dermal–epidermal junction (DEJ) are targets of autoimmune blistering diseases, such as bullous pemphigoid and epidermolysis bullosa acquisita (Chapter 63).

The hypodermis lies beneath the dermis and attaches skin to underlying muscle. It is composed mainly of adipose tissue, which cushions skin impacts. Adipose cells produce leptin, which is implicated in a variety of inflammatory skin diseases, including psoriasis (Chapter 64).

The protective functions of skin are supported by three major barriers: *(i)* physical, *(ii)* pharmacological or detoxifying, and *(iii)* immunological. The physical barrier includes the hair and stratum corneum. The stratum corneum provides an impermeable hydrophobic cover of protein-filled squames that hinders the entry of invading microorganisms or toxic chemicals into the body. It also reflects and scatters UVR to impede the damaging energy from reaching the deeper, regenerating layers of skin. The pharmacological barrier includes detoxifying and repair enzymes synthesized by epidermal cells. Some metabolize chemicals and xenobiotics, whereas others repair DNA damage resulting from UVR or environmental mutagens.

The immunological barrier includes cells and molecules that are unique to skin and are historically referred to as "skin associated lymphoid tissue."[5] They include elements of both innate immunity and acquired immunity. Together, these elements

TABLE 19.1 Immunodermatological Disorders That Affect the Skin

- Papulosquamous disorders
  - Psoriasis
  - Lichen planus
  - Cutaneous graft-versus-host disease
  - Acute, subacute, and discoid lupus erythematosus
- Eczematous disorders
  - Atopic dermatitis
  - Allergic contact dermatitis
- Urticarial disorders
  - Urticaria and angioedema
  - Erythema multiforme
  - Stevens Johnson syndrome
  - Toxic epidermal necrolysis
  - Cryopyrin-associated periodic syndromes
  - Muckle-Wells syndrome
  - Familial cold urticaria
  - Neonatal-onset multisystem inflammatory disease (NOMID)
  - Deficiency of the interleukin-1–receptor antagonist (DIRA)
  - Tumor necrosis factor receptor–associated periodic syndrome (TRAPS)
- Purpuric disorders
  - Leukocytoclastic vasculitis
  - Medium vessel vasculitides (polyarteritis nodosa, Wegener granulomatosis, Churg-Strauss vasculitis)
- Vesiculobullous diseases
  - Pemphigus
  - Bullous pemphigoid
  - Paraneoplastic pemphigus
  - Epidermolysis bullosa acquisita
  - Dermatitis herpetiformis
  - Linear immunoglobulin A (IgA) bullous dermatoses
  - Pemphigus gestationis
- Pigmentary disorders
  - Vitiligo
- Hair disorders
  - Alopecia areata
- Autoimmune disorders
  - Dermatomyositis
  - Mixed connective tissue disease
  - Scleroderma and morphea
- Photodermatoses
  - Polymorphous light eruption
  - Solar urticaria
  - Chronic actinic dermatitis
  - Photoallergic contact dermatitis
- Allergic drug eruptions
- Disorders of subcutaneous tissues
  - Erythema nodosum
  - Erythema induratum
- Immunodeficiencies
  - Ataxia telangiectasia
  - Chronic mucocutaneous candidiasis
  - Chronic granulomatous disease
  - Hyperimmunoglobulin E (IgE) syndrome (HIES)
  - Leukocyte adhesion molecule deficiency
  - Severe combined immunodeficiency
  - Warts–hypogammaglobulinemia–infections–myelokathexis syndrome (WHIM syndrome)
  - Wiskott-Aldrich syndrome
  - Organ transplant recipients on immunosuppressive medications

promote both local and systemic responses to protect against invading pathogens and neoplastic cells. Locally they act to enhance protective mechanisms in affected skin. Systemically they alert the rest of the body to build and expand host defenses. Although these protective measures work exceedingly well, they do not work perfectly. Disturbances in these protective mechanisms can result in increased infections and malignancies when deficient or in immunologically mediated skin diseases when excessive (Table 19.1).

INNATE IMMUNITY AND SKIN

Pattern Recognition Receptors

By rapidly sensing and responding to microbial pathogens and environmental toxins, epidermal keratinocytes act as innate first responders (Chapter 3). Keratinocyte pattern recognition receptors (PRRs) recognize highly conserved sequences in macromolecular components of microbial pathogens and host-derived danger indicators. The identification of pathogen-associated molecular

KEY CONCEPTS

Immunological Defenses of Skin

- The innate defenses of skin provide the first line of defense against environmental antigens and invading pathogens.
- In response to pathogen-associated molecular patterns (PAMPs) and damage-associated molecular patterns (DAMPs), granular layer keratinocytes produce cutaneous microbicidal peptides β defensin and cathelicidin, and proinflammatory mediators, such as interleukin (IL)-1α, IL-1β, IL-6, and tumor necrosis factor-α (TNF-α).
- Tricellulin molecules are specialized tight junction proteins that regulate sampling of the external microbial environment by protruding dendrites of Langerhans cells (LCs) residing in the suprabasal layer of the epidermis.
- At initial skin breach, mediators derived from a variety of cells activate neighboring keratinocytes and LCs to perpetuate a proinflammatory signal cascade throughout the epidermal layers and lead to the activation of dermal cells and dermal microvasculature.
- Immune cells that are resident in healthy skin include epidermal LCs; dermal leukocytes, such as mast cells; macrophages; and dendritic cells (DCs).
- LCs internalize microbes and produce IL-1β, IL-6, and other cytokines; become activated in response to a combination of signals (Toll-like receptors [TLRs], NOD-like receptors [NLRs], IL-1β, and TNF-α); and migrate from skin to reach regional lymph nodes (LNs), where they mature into potent antigen-presenting cells (APCs).
- In LNs, LCs activate T cells and deliver antigens and other information from affected skin to LN resident DCs to initiate and amplify skin-specific immunity.
- Keratinocyte-derived proinflammatory mediators and chemokines (CXCL8/IL-8) act on post endothelial venules to promote extravasation of inflammatory cells from the circulation into the dermis, including neutrophils, monocytes/macrophages, plasmacytoid DCs (pDCs), natural killer (NK) cells, and memory T cells.
- Inflammatory leukocytes enter the epidermis and attach to keratinocytes through interactions with CD54 (intercellular adhesion molecule 1 [ICAM-1]) proteins on activated keratinocytes to clear affected epidermis of invading pathogens and cellular debris.

patterns (PAMPs) and danger-associated molecular patterns (DAMPs) is principally undertaken by an array of PPRs. Membrane-bound Toll-like receptors (TLRs) are present at the cell surface (TLRs 1, 2, 4–6) and intracellular endosomes (TLRs 3, 7–9). Also present are cytosolic nucleotide-binding domain, leucine-rich repeat containing receptors (NLRs) (formerly termed *NOD-like receptors*).[6] TLR and NLR expression profiles are differentially expressed (often with partial overlap) on epidermal and leukocyte cell subsets.

Of the 10 TLRs that humans express, messenger RNA (mRNA) for all have been detected in keratinocytes. However, receptor proteins and activities have been detected for only eight (TLR1-7 and 9).[6] Human melanocytes constitutively express mRNA for TLRs 1–4, 6, 7, and 9, and abundant functional protein is detected for TLRs 2–4, 7, and 9. Human and murine DC subsets also show differential TLR expression profiles. For example, conventional DCs (DCs) express combinations of TLRs 1–5 and 8, whereas TLR expression is restricted to virus-detecting endosomal TLRs 7 and 9 in plasmacytoid DCs (pDCs). The latter is consistent with their primary function as major interferon-α (IFN-α) producers.

NLRs have over 20 members in humans and over 30 in mice, reflecting the broad spectrum of PAMPs and DAMPs they can recognize. Structurally, NLRs can be divided into four subfamilies: *(i)* NOD, caspase recruitment domain (CARD)-containing NLRCs; *(ii)* NALP, pyrin domain (PYD)-containing NLRPs; *(iii)* NLRB, NLR apoptosis inhibitory protein NAIP; and *(iv)* NLRA, the transcription factor CIITA. Primary human keratinocytes express mRNA transcripts for 11 of 14 NALP family genes.[7] NOD1 is expressed in many tissues, including skin keratinocytes. NLRs act independently, but in synergy with TLR signals, to trigger fast-acting potent innate immune responses against microbes. They can either promote nuclear factor-κB (NF-κB) activation or form inflammasomes that generate interleukin-1 (IL-1) family proinflammatory mediators.

TLR ligands are derived from conserved microbial products, such as lipopolysaccharides (LPS) from gram-negative bacteria, lipoteichoic acid and peptidoglycans from gram-positive bacteria, mannans of yeast and fungi, and nucleic acids from viral and bacterial pathogens. Endogenous DNA derived from stressed host cells can be recognized by TLRs. DAMPs are recognized by NLRs. The set of TLRs and NLRs that are triggered determines what types of immunomodulatory or inflammatory responses will be elaborated. Acting as adjuvants, they can also influence the magnitude of the response. In the clinic, a drug known as a TLR7 agonist (imiquimod) is used as an effective topical agent for eliminating human papillomavirus (HPV)–induced genital warts and nonmelanoma skin tumors.

Keratinocytes, DCs, and monocytes also express the PRR dectin-1. Dectin-1 is a membrane-associated glycoprotein that recognizes β-glucan, a polysaccharide component of several species of fungi. Dectin-1 activation results in the production of IL-1β, IL-6, and IL-23 (Chapter 9), which promotes the development of T-helper type 17 (Th17) cells—important antifungal effector cells (Chapter 16). Patients lacking dectin-1 exhibit Th17-cell deficiency, which leads to recurrent vulvovaginal *Candida albicans* infections and onychomycosis[5] (Chapter 29). Mutations in Card9, a mediator of dectin-1 signal transduction, also lead to a Th17 deficiency associated with chronic mucocutaneous candidiasis.[8]

Cutaneous Production of Cytokines and Chemokines

Types of Cytokines and Chemokines

Skin is a rich source of cytokines (Chapter 9) and chemokines (Chapter 10) that influence the magnitude and type of cutaneous immune response that is generated.[9] They are also important in driving nonimmunological processes, such as activating systemic responses to injury, wound healing, carcinogenesis, and pigmentation. Production of these mediators in skin involve several different cell types and environmental stimuli.

Keratinocytes produce a variety of immune mediators during different phases of immune reactions: proinflammatory cytokines IL-1, IL-6, IL-33, and tumor necrosis factor-α (TNF-α); immunosuppressive and antiinflammatory cytokines IL-10, TGF-αβ, and IL-1 receptor antagonist (IL-1RA); colony-stimulating factors GM-CSF, G-CSF, and M-CSF; immunomodulatory cytokines IL-7, IL-12, IL-15, IL-18, IL-19, IL-20, and IL-23; and chemokines CXCL8/IL-8, CCL2/MCP-1, CCL20/MIP3α, and CCL5/RANTES. LCs produce IL-1αβ, IL-12, and IL-23, CCL3/MIP1α, CCL4/MIP1αβ, and CCL5/RANTES. Melanocytes produce cytokines IL-1β, IL-6, IL-8, TNF-αα, type I interferons (IFN-αα and IFN-β), and chemokines CCL2, CCL3, and CCL5.

Although cytokines and chemokines produced in skin are functionally identical to those secreted by nonskin cells, in skin they can have unique effects. For example, IL-1 can remodel the dermis by inducing MMP production and augmenting collagen production. Mutant forms of cytokines can also cause cutaneous diseases. For example, a deficiency of IL-1–receptor antagonist (DIRA) is associated with a mutation in the gene that encodes the IL-1–receptor antagonist. This mutation can cause severe pustulosis, ichthyosiform lesions, psoriasis-like changes, and nail abnormalities.[10]

Phases of Active Cutaneous Cytokine Secretion

In response to immunological and inflammatory stimuli, keratinocytes produce different sets of immune mediators in three sequential phases: initiation, amplification, and resolution.

Initiation. The initial set of cytokines that are generated include proinflammatory IL-1, IL-6, TNF-α, and many others. Keratinocytes constitutively produce a pool of inactive proinflammatory precursors: pro–IL-1α, pro–IL-1β, and pro–IL-18. These molecules can be immediately processed into active mediators by inflammasomes, which are rapidly formed in the presence of DAMPs.[7] IL-1β, a potent stimulator of endothelial cell activation, permits rapid infiltration of inflammatory leukocytes and T cells from blood into the dermis. TNF-α contributes to dermal microvascular endothelial cell activation and augments expression of ICAM-1 (CD54) on keratinocytes and endothelial cells. Infiltrating T cells express the β_2 integrin molecule leukocyte function–associated antigen-1 (LFA-1), which binds to ICAM-1 expressed on affected dermal and epidermal layers. IL-6 helps drive the development of IL-17–producing Th-17 cells. IL-6 can also perpetuate chronic inflammation.

Amplification. Activated keratinocytes respond to the first wave of proinflammatory mediators by generating of a second set of cytokines and chemokines (including GM-CSF, TNF-α, and CXCL8/IL-8) with autocrine and paracrine effects. They can activate unaffected neighboring keratinocytes to amplify proinflammatory signals. These cytokines are key to triggering adaptive immune responses by activating cutaneous LCs and DCs to become mature potent antigen-presenting cells (APCs). APC interaction with antigen-specific, skin-resident memory T cells stimulates the proliferation and expansion of differentiated effector T cells at the affected skin site. Keratinocyte-derived IL-8 is a potent chemoattractant for neutrophils. The influx of leukocytes further amplify local immune responses, inflammation, and associated tissue damage. Stimulation of TLRs 7 and 9 within small numbers of resident dermal pDCs leads to abundant production of type I interferons (IFN-α), inducible nitric oxide synthase (iNOS), and arginase, which aid in the eradication of viral infection and other invading organisms.

Resolution. After the pathogen is eliminated, keratinocytes switch their cytokine production program in response to IL-6 and generate anti-inflammatory factors, such as IL-10 and IL-1 receptor antagonist (IL-1RA). During this phase, recruited leukocytes can also produce IL-10. IL-10 is a potent inhibitor of adaptive and inflammatory cellular responses, whereas IL-1RA effectively blocks the activity of IL-1β. Both activities function to return skin to its normal resting state. In addition, macrophages and certain DC subsets can inhibit ongoing adaptive immune responses through their expression of indoleamine 2,3 dioxygenase (IDO), an enzyme that alters tryptophan metabolism in T cells.

Antimicrobial Peptides

Skin is a production site for antimicrobial peptides, which are important components of the innate immune response.[11] Over 20 antimicrobial peptides have been described. They function at the earliest stages of infection by microorganisms that breach the stratum corneum. These antimicrobial peptides have the ability to disrupt bacterial and fungal membranes, as well as viral envelopes. These peptides have broad effects on innate and adaptive immune responses. The two best characterized skin antimicrobial peptides are the cathelicidins and β-defensins. They are synthesized by keratinocytes, cells of the epidermal sebaceous and eccrine glands, and dermal mast cells. Cathelicidins are secreted as precursor proteins (*e.g.,* hCAP18), which are then processed to an active form (*e.g.,* IL-37). They are detected at low levels in unperturbed skin but are strongly increased following infection or disruption of the epidermal barrier.

Cathelicidins interact with a variety of cell-surface receptors, including TLR-like and endothelial growth factor (EGF) receptors. They enhance leukocyte migration, stimulate the secretion of cytokines and chemokines, and promote angiogenesis. Interestingly, vitamin D_3, the synthesis of which is initiated in skin by UVR, plays an important role in regulating cathelicidin production through epigenetic mechanisms.[12]

Abnormalities in antimicrobial peptides have been implicated in a variety of immunological and nonimmunological skin diseases.[10] Deficiencies in antimicrobial peptide levels have been found in atopic dermatitis (Chapter 44), which may explain why patients with these deficiencies are at greater risk for viral (herpes simplex, HPV-induced warts, and poxvirus-induced molluscum contagiosum) and bacterial (*Staphylococcus aureus*) skin infections. Rosacea manifests as an acneiform eruption with erythema, telangiectasias, and flushing of the face. In rosacea, excessive amounts of cathelicidins are associated with characteristic inflammation and angiogenesis.[13] In psoriasis (Chapter 64), cathelicidins are postulated to combine with self DNA from damaged keratinocytes. This stimulates type I IFN production by pDCs, which is a key to the development of pathogenic Th17 cells.[14]

Nonimmune mechanisms that cause physical disruption of the skin barrier (*e.g.,* shear forces, chemical, thermal, or UV damage) stimulate "sterile" inflammatory responses that produce similar proinflammatory mediators and antimicrobial peptides to protect vulnerable sites during wound healing.

KEY CONCEPTS

- Skin is the major interface between an individual and his or her environment. Skin-resident and recirculating immunocompetent cells and immune mediators are used to defend against intrusion by microorganisms, noxious exogenous chemicals, and mutant cells that arise within it.
- When the immune system overperforms, it can give rise to autoimmune skin diseases. Pemphigus, for example, is a blistering disease in which antibodies are directed at desmogleins-1 and -3.
- When the immune system underperforms, cutaneous immunodeficiency occurs. For example, patients taking immunosuppressive medications suffer an increased risk of developing squamous cell carcinomas of skin.
- Skin-specific immune system components provide targets for new therapies. For instance, T-helper type 17 (Th17) cells play a dominant role in psoriasis, and neutralizing antibodies to IL-17 have recently been approved as a highly effective treatment for this disease.

ADAPTIVE IMMUNITY AND SKIN

Although the innate immune response forms the first line of defense against microbial threats in an antigen-nonspecific manner, the host's adaptive immune response provides antigen-specific protection against intracellular and extracellular pathogens. In contrast to innate immunity, adaptive immunity is able to distinguish self from nonself to eliminate microbial-infected or mutated cancer cells without damaging normal cells. The adaptive immune response in skin is carried out primarily by T cells and by immunoglobulin E (IgE) antibodies that are bound to mast cells. T cell–mediated immunity is dependent on activation by antigen-presenting DCs.

Dendritic Cells

DCs are a heterogeneous group of leukocytes that typically exhibit a DC surface, providing a large surface area for interaction with T cells (Fig. 19.2) (Chapters 5 and 6). LCs are specialized epidermal DCs originally discovered by Paul Langerhans in 1868. Under resting conditions, two different types of DCs are typically found in the dermis–myeloid derived conventional DCs (cDCs) and lymphoid plasmacytoid DCs (pDCs). cDCs are the most potent APC type. pDCs are less dendritic and less potent APCs but are still potent producers of IFN-α.[15] In perturbed skin, inflammatory DCs (iDCs) accumulate to produce abundant levels of TNFα and iNOS. They are derived from circulating monocytes that differentiate into DCs upon transendothelial migration into inflamed dermis.

To perform APC function, LCs and DCs express a wide array of antigen capture receptors that internalize bound microbes and antigens into processing endosomes, where antigens are cleaved and trimmed into short peptides and then loaded into the peptide-binding groove of MHC (or human leukocyte antigen [HLA]) molecules (Chapters 5 and 6). Peptide–MHC molecular complexes are presented at the surface of DCs for recognition by antigen-specific T-cell receptors (TCRs) (Chapter 4) on naïve and/or memory T cells. Through integration of environmental signals, LCs and DCs are able to program naive T cells to differentiate into appropriate, functionally distinct subsets of skin-homing effector cells that can eradicate a particular pathogen.

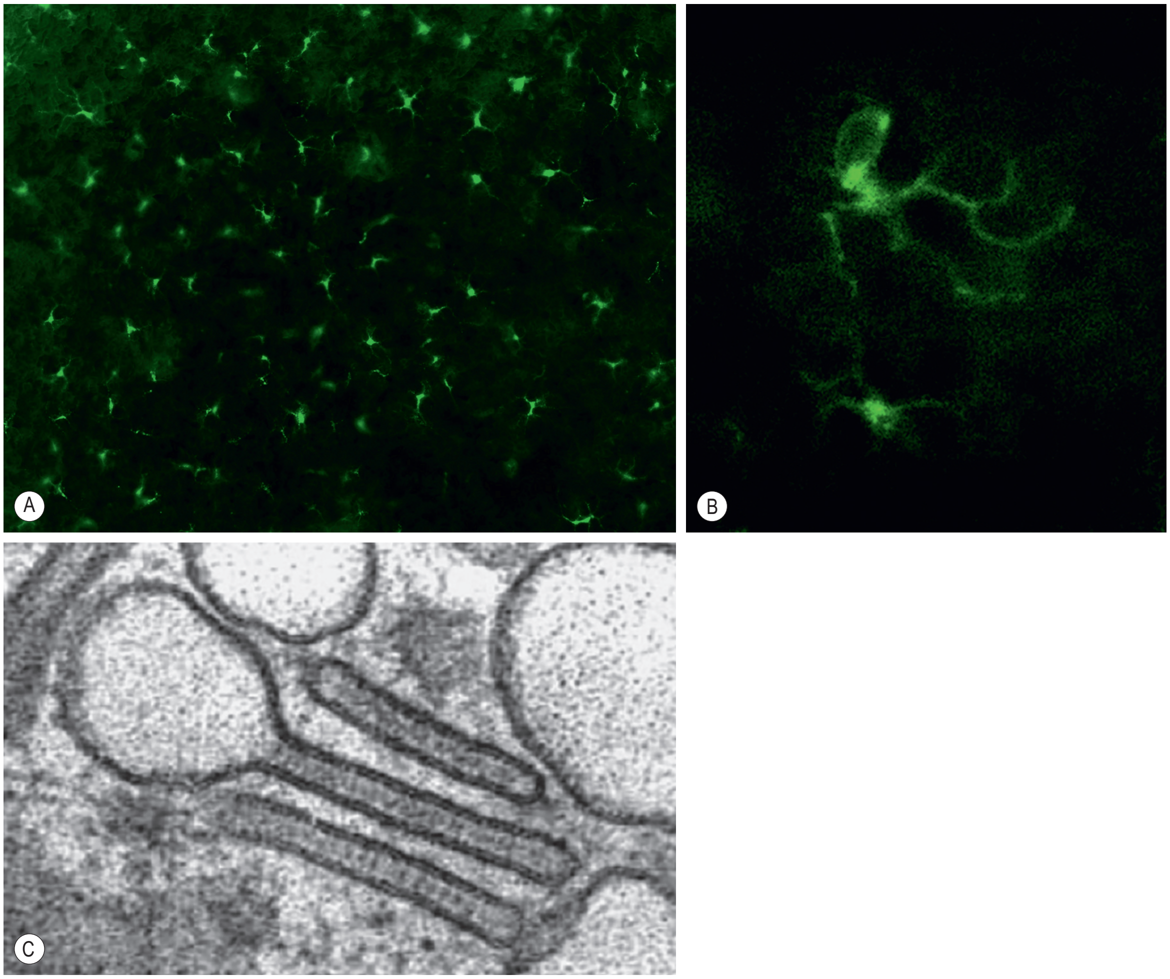


FIG 19.2 Langerhans Cells (LCs). (A) A horizontal view of epidermis (murine) stained with an antibody to major histocompatibility complex (MHC) class II molecules, of which only LCs are positive. Dendritic shaped LCs are spaced throughout the epidermis with dendrites stretching out to meet and communicate with neighboring LCs. (B) High-power view of an LC. Note contact between adjacent LCs. (C) Transmission electron microscopy of Birbeck granules in LCs from human epidermis. The Birbeck granule has a "tennis racket" morphology and is formed by internalized langerin molecules. (From: Romani N. et al. Langerhans cells and more: langerin-expressing dendritic cell subsets in the skin. Immunol Rev 2010:234(1):120–41.)

Epidermal Langerhans Cells

In normal skin, resting LCs reside in the suprabasal layer of the epidermis, where they cover the entire skin surface in a net-like structure. They represent 1–3% of all cells in the epidermis. They are poised to sense and respond to the earliest stages of a microbial or physical breach of epidermal integrity (see Fig. 19.2). LCs are identified by selective expression of antigen capture receptor langerin (CD207), a C-type lectin. Langerin internalizes to form a definitive "tennis racket–shaped" organelle called the *Birbeck granule,* which is thought to be a specialized antigen-processing compartment (see Fig. 19.2). Human LCs are also identified by their expression of CD1a, a protein structurally similar to MHC/HLA class I molecules that present lipid and glycolipid microbial antigens, rather than peptides. *In vivo* studies of mouse LCs reveal that they can extrude their dendritic processes across the skin tight junction, present in the granular layer, through a specialized pore, called *tricellulin,* to sample the microbial environment in the stratum corneum. Langerin receptors concentrate at the tips of the extruded dendrites to aid in the capture of external pathogens.

LCs have long been considered the paradigm-defining conventional DCs that transfer site-specific antigens to the lymph node, where they inform and activate CD4 and CD8 T cells (Chapter 8) to develop into differentiated skin-homing helper and cytotoxic effector effectors to eliminate the invading threat. Limited studies of primary human LCs indicate they are potent at inducing cytotoxic CD8 T cells and CD4 Th2 cells.[16] This is in contrast to relatively recent *in vivo* studies of mouse LCs that appear unable to activate CD8 T cells. Thus mouse and human LCs differ with regard to their division of labor.

In addition to their potent cross-presenting activity to CD8 T cells (Chapter 17); human LCs, like mouse LCs, also favor development of CD4 regulatory T cells (Tregs) (Chapter 18), antimicrobial Th17 cells, and follicular helper T (Tfh) cells for antibody development[17,18] (Chapter 16). LCs that emigrate from the epidermis leave a hole in the skin surface network, which is subsequently replenished, under inflammatory conditions by iDCs for the short term and by resident LC-committed stem cells that seed the epidermis during fetal development for the long term.[17,19]

Dermal Dendritic Cells

Dermal DCs efficiently present antigens that have reached the dermis to T cells. In normal skin, multiple subpopulations of resident dermal DCs can be detected.[19] In mice, dermal langerin positive and negative DCs are found, both of which are developmentally distinct from LCs. In humans, most dermal DCs are langerin negative; only a small number of langerin-dull dermal DCs can be detected. Nevertheless, human langerin negative dermal DCs share functional characteristics with mouse langerin-positive dermal DCs. They are able to cross-present captured intracellular bacterial or viral antigen from infected skin cell debris to activate CD8 T cells without being infected themselves. Thus they are able to forgo the canonical MHC class I–antigen processing pathway, which requires infection for endogenous expression of the pathogen antigens. Mouse langerin positive dermal DCs are also critical for activating Tfh cells for IgG2a/c and IgG2b isotype antibody production. In the resting state, all DCs are poor stimulators of T cells. Interaction with so-called immature DCs can generate unresponsive antigen-specific T cells, a key mechanism for maintaining immune tolerance, which is needed to prevent the immune system from attacking self tissues.

T Cells and Immune Responses in Skin

Phases of the Cell-Mediated Immune Response in Skin

The adaptive immune response in skin occurs in two phases: *(i)* sensitization (or immunization phase) by first exposure to antigen and *(ii)* elicitation (or effector phase) in response to re-exposure to antigen (Fig. 19.3).

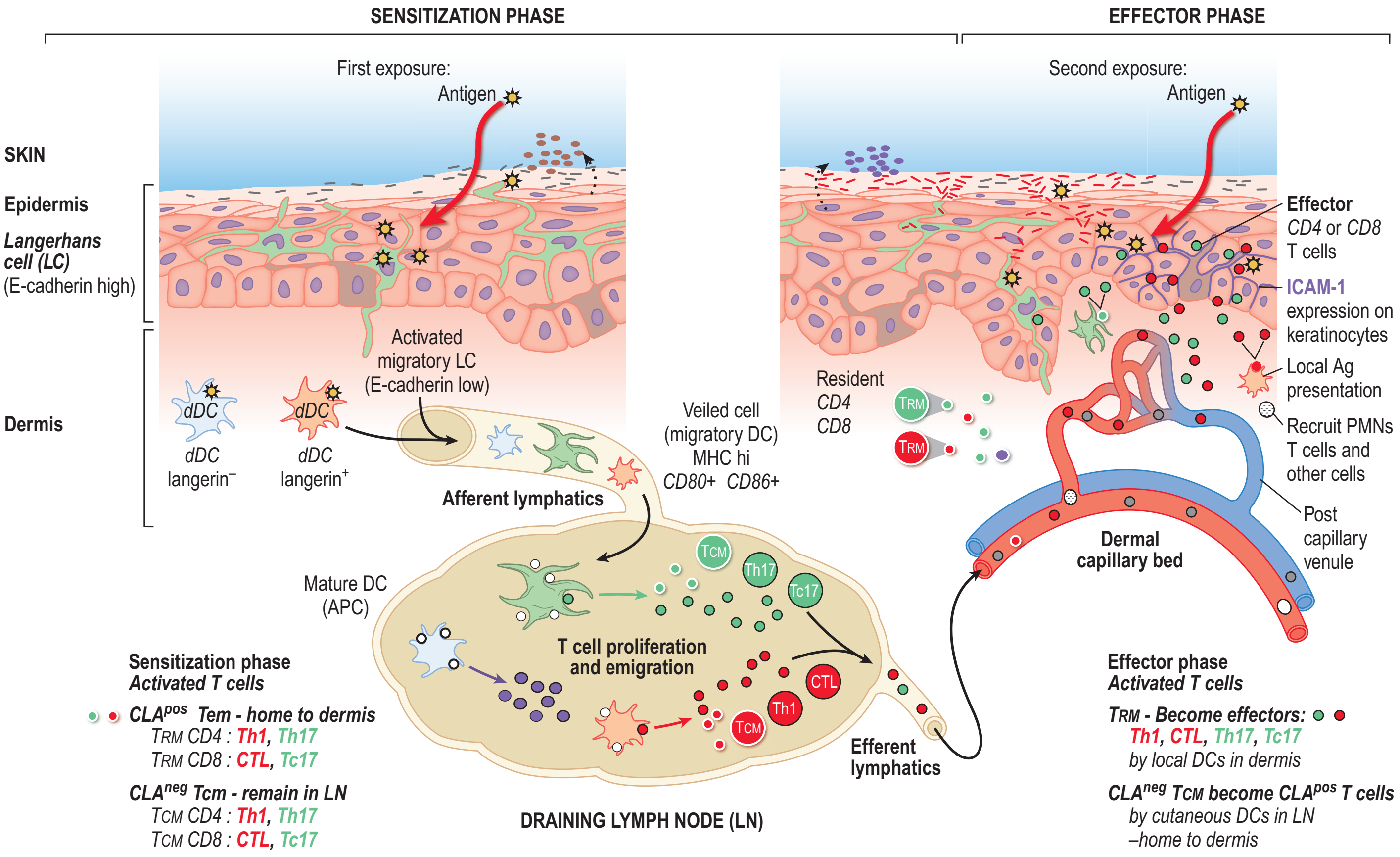

SENSITIZATION PHASE
EFFECTOR PHASE
First exposure:
Antigen
SKIN
Epidermis
Langerhans cell (LC)
(E-cadherin high)
Dermis
Activated migratory LC
(E-cadherin low)
dDC
dDC
dDC langerin⁻
dDC langerin⁺
Afferent lymphatics
Veiled cell (migratory DC)
MHC hi
CD80+ CD86+
Mature DC (APC)
TCM
Th17
Tc17
T cell proliferation and emigration
CTL
Th1
TCM
Efferent lymphatics
DRAINING LYMPH NODE (LN)
Sensitization phase
Activated T cells
CLApos Tem - home to dermis
TRM CD4 : Th1, Th17
TRM CD8 : CTL, Tc17
CLAneg Tcm - remain in LN
TCM CD4 : Th1, Th17
TCM CD8 : CTL, Tc17
Second exposure:
Antigen
Resident
CD4
CD8
TRM
TRM
Effector
CD4 or CD8
T cells
ICAM-1
expression on keratinocytes
Local Ag presentation
Recruit PMNs T cells and other cells
Post capillary venule
Dermal capillary bed
Effector phase
Activated T cells
TRM - Become effectors:
Th1, CTL, Th17, Tc17
by local DCs in dermis
CLAneg TCM become CLApos T cells
by cutaneous DCs in LN
–home to dermis

FIG 19.3 Phases for the Development of Immune Responses in Skin. During the sensitization phase, dendritic cells (DCs) capture pathogens in skin and migrate through afferent lymphatics to the skin draining lymph nodes (LNs), where they present the antigens to naïve T cells. T cell differentiation is induced by activated cutaneous DCs, and they develop into resident memory T (T_{RM}) cells and central memory T (T_{CM}) cells representing different lineages (Th1, CTL, Th17, Tc17), depending on DC programming. Activated T_{CM} that leave the LN and migrate to skin become skin resident $T_{RM,}$ whereas nonactivated T_{CM} stay in lymph nodes. The effector phase occurs in response to a second exposure to antigen, which activates DCs, which, in turn, activates local T_{RM} to become effector cells *in situ.* Activated cutaneous DCs that reach the draining LN activate T_{CM} to become skin-homing T effector cells that are recruited to the inflamed dermal site. T effector cells, in concert with recruited cells, will enter into the epidermal layer by binding intercellular adhesion molecule 1 (ICAM-1) molecules present on activated keratinocytes, to clear infection, eliminate pathogen-infected cells, remove debris, and repair the skin barrier.

Sensitization/immunization phase. Epidermal LCs and dermal DCs are responsible for carrying antigens from the affected skin site to draining lymph nodes, where they present those antigens to activate and expand antigen-specific T cells. To emigrate from the epidermis, activated LCs downregulate E-cadherin to detach from their surrounding keratinocytes. During migration to the draining lymph node, LCs and DCs undergo a maturation process that stops antigen capture activity (by downregulation of antigen capture receptors) and increases expression of T-cell interactive molecules: MHC class I and II molecules, costimulatory molecules CD80 and CD86, and ICAM-1 (CD54).

Naïve T cells in the lymph nodes that express TCRs for presented antigens form long-term APC–T cell conjugates (on the order of 30 minutes) during which T cells are programmed to differentiate into functionally distinct T-cell subsets. These new T cells leave the lymph node by downregulating their lymph node–homing receptors, CCR7, and L-selectin (CD62L) and then upregulating the skin-specific homing receptor CLA and chemokine receptors CCR4 and CCR10 to enter the affected skin site.[20] Unless antigen is present in skin for an extended period, the sensitization phase does not cause an inflammatory reaction and therefore often goes unnoticed.

Elicitation/effector phase. In the dermis of resting skin, a large number of resident memory T cells (T_{RM}), previously activated during the sensitization phase, persist in the absence of antigens and are readily capable of effector functions. They provide rapid on-site immune protection to secondary antigen exposure.[20] Central memory T cells (T_{CM}) reside in lymph nodes and differ from T_{RM}, as they do not express high levels of skin-homing receptors CLA and CCR4 or exhibit immediate robust responses to antigen restimulation. T_{CM} in lymph nodes or spleen can be activated by secondary presentation of Ag by skin-draining LCs or DCs. Activated T_{CM} enter the circulation, express skin-homing receptors, and extravasate into the affected skin site, where they can contribute to the elicitation response. CD4 and CD8 T_{RM} and recruited T_{CM} that interact with antigen-activated LCs and DCs in the dermis are further stimulated to proliferate *in situ* to amplify *(i)* effector cell numbers, *(ii)* cytokine and chemokine production, and *(iii)* further recruitment of inflammatory leukocytes, which act in concert to eradicate microorganisms, chemically modified, or neoplastic cells.

Skin Resident T-Cell Subpopulations

In response to instructional cues from cutaneous DCs, antigen-specific T cells differentiate into specialized subsets, each with functions that focus on eliminating specific types of pathogens (Table 19.2). These presensitized, functionally committed T cells that home to skin remain in the dermis as long-lived T_{RM}s. The

TABLE 19.2 T-Cell Subsets and Trafficking in Skin

T-Cell Subset	Homing-R[a]	Chemokine-R	From	Homing to	Ligands for Homing and Chemokine Receptors
Naïve T cells	L-selectin (CD62L) Lymph node (LN) homing receptor	CCR7^{+}	Blood	LN	L-selectin ligand: GlyCAM-1 integrin on LN HEV[b]
Central memory T cells (T_{CM}) CD45RA^{+}	L-selectin	CCR7^{+}	LN	LN and the circulation	CCR7 ligand: LN-derived chemokine CCL21 (also called 6Ckine, secondary lymphoid-tissue chemokine (SLC)
Resident memory T cells (T_{RM}) CD45RO^{+}	Cutaneous lymphocyte-associated antigen (CLA^{+}) (P-selectin glycoprotein ligand-1 [PSGL-1]) Skin-homing receptor	CCR4^{+}	LN	Blood/dermis	CLA ligand: E-selectin on activated postcapillary venules CCR4 ligand: skin-derived inflammatory chemokines: MCP-1, MIP-1, RANTES, TARC, CCL22
T_{RM} Subsets Resident in Dermis	**Homing-R**	**Chemokine-R**	**Function**	**Cytokines Produced**	**Notes**
T regulatory (Treg)	CLA^{+}	CCR5^{+}	Regulatory; resolution phase; inhibit auto reactivity	Interleukin (IL)-10, transforming growth factor (TGF)-β	Tregs are 5–10% of CLA^{+} dermal T cells Treg two-way trafficking–migration from dermis to LN, then recirculate back to dermis CCR5 binds RANTES, MIP-1
T-helper (Th)1, Tc1	CLA^{+}	CCR4^{+}	Type I antiviral	Interferon (IFN)-γ, tumor necrosis factor (TNF)-α, IL-2	Promotes cellular immunity Development of cytotoxic T cells (CTLs)
Th2, Tc2	CLA^{+}	CCR4^{+}	Type II Parasites cleared	IL-4, IL-5, IL-10, and IL-13	Promotes humoral immunity Recruitment of eosinophils
Th17, Tc17	CLA^{+}	CCR4^{+}	Inflammatory; anti fungal; antibacterial		Activation is dependent on IL-23 derived from keratinocytes/Langerhans cells (LCs)/dendritic cells (DCs)

Twenty billion T cells are resident in healthy dermal tissue, representing twice the number of circulating T cells present in the entire circulation (T-cell distribution density in skin ≈1 million T cells/cm^2).

[a] Homing-R – an adhesion molecule on leukocytes that recognizes and binds site-specific adhesion molecules expressed on high endothelial venules in the lymph node (LN) and activated postcapillary venules in the dermis and other organs. L-selectin is the lymph node homing receptor, whereas CLA is the skin-homing receptor.

[b] HEV – high endothelial venules; these are postcapillary venous swellings in LNs and most secondary lymphoid organs.

skin of a healthy adult contains nearly 20 billion T_{RM} cells, about twice the number present in the entire blood volume.[20] Among the best-characterized differentiated T-cell subsets are CD4 Th1, Th2, Th17, and Tregs, each of which produces a distinct profile of cytokines. The Th22 subset has recently been discovered and is implicated in human inflammatory skin diseases.[21] CD8 T cells primarily function as cytotoxic effector cells, but they also express cytokine profiles analogous to CD4 subsets, termed *Tc1, Tc2,* and *Tc17.*[22]

Th1 responses. Th1 cells preferentially produce IFN-α and IL-2 and are the principal regulators of type 1 immunity, which eradicates intracellular pathogens and tumors. The major effector cytokine produced by Th1 cells is IFN-α. Macrophages are stimulated by IFN-α to phagocytose and generate oxidative bursts that aid intracellular killing of microbes. IFN-α also upregulates expression of class I and class II MHC molecules and ICAM-1 (CD54) on keratinocytes, dermal microvascular endothelial cells and fibroblasts, and induces them to secrete proinflammatory cytokines and chemokines.[23] In animal models, deficiencies in IFN-α render them more susceptible to cutaneous microbial infections and tumors. Th1 cells develop cutaneous delayed-type hypersensitivity responses, such as the tuberculosis skin test reaction, and allergic contact dermatitis to haptens. Haptens are small compounds that penetrate skin and bind to epidermal proteins, generating "modified self" proteins as immunogenic antigens. Examples include urushiol in poison ivy and nickel.

Th2 responses. Th2 cells are involved in type 2 immune responses, which are important for eradication of extracellular parasites and bacterial infection. They produce IL-4, IL-5, IL-10, and IL-13, which are important for the induction and development of humoral immune responses. IL-4 and IL-13 activate B-cell proliferation, Ig class-switching, and antibody production. Th2 cell-mediated inflammation is characterized by the presence of eosinophils and basophils, as well as extensive mast cell degranulation—a process dependent on cross-linking surface-bound IgE.[24] IL-5 is a potent hematopoietic cytokine, which stimulates bone marrow production of eosinophils as well as activation and chemotaxis of eosinophils and basophils to affected tissue. In mice, Th2-cell deficiency profoundly increases susceptibility to Leishmania infection in skin. In humans, Th2 cells appear to play a critical role in the pathogenesis of atopic dermatitis (Chapter 44). A recent clinical trial with dupilumab, a fully human mAb that targets the IL-4 receptor-αα and blocks IL-4 and IL-13 signaling, improved atopic symptoms .

Th17 responses. Th17 cells produce IL-17A, IL-17F, IL-21, and IL-22. (IL-17A is commonly called IL-17.) Th17 cells have protective effects against extracellular bacterial and fungal infections in skin. Both IL-17A and IL-17F enhance protective immune responses by inducing the production of CXC chemokines, G-CSF, and antimicrobial peptides by keratinocytes and epithelial cells of other organ systems. Mice deficient in IL-17A are highly susceptible to *Staphylococcus aureus* infections in skin. IL-17 and IL-22 are important mediators of psoriasis and have been implicated in other autoimmune skin disorders. In clinical trials testing anti-human IL-17A antibodies (secukinumab and ixekizumab), three-quarters of patients with psoriasis exhibited reduced symptoms. Brodalumab, an IL-17 receptor antagonist, has shown efficacy in psoriasis. IL-17 may also promote tumor development and UVR-induced immune suppression in skin.[25]

Regulatory T cells. The principal aim of the immune system is to establish a balance between defense against pathogens and avoidance of autoimmune responses. This balance is achieved partly through the action of Tregs (Chapter 18). Natural Tregs (nTregs) develop in the thymus in response to self antigens, whereas induced Tregs (iTregs) develop in the periphery in response to foreign antigens. Both are commonly characterized as $CD4^+/CD25^+/Foxp3^+$, and both produce immunosuppressive cytokines (IL-10 and/or TGF-β). They can also express immunoregulatory surface molecules, such as CTLA-4, PD-1, glucocorticoid-induced TNFR family–related gene (GITR), and lymphocyte activation gene 3 (LAG-3). They can directly inhibit the activation and function of T cells and suppress the activity of antigen-presenting DCs. A defect in Treg function will result in autoimmune and inflammatory diseases. About 5–10% of the T cells resident in normal human skin are Tregs. Tregs diminish immune responses in skin, an important function for dampening immune and inflammatory responses and limit tissue damage.[26] Conversely, increased Treg suppression can impair host immunosurveillance against tumors. UVR-induced immunosuppression is considered a critical mechanism for development of UVR-induced tumors, and this is, to a great extent, mediated by UVR-induced Tregs.[27]

CD8 T-cell immunity. CD8 T cells differentiate into cytotoxic T lymphocytes (CTLs), which are essential for the elimination of viral infections and cancerous cells (Chapter 17). The CD8 T-cell subsets Tc1, Tc2, and Tc17 also express cytokines analogous to CD4 helper subsets.[22] Tc1 cells express the killer molecules Fas-L, perforin, and granzyme B, as well as large amounts of IFN-α, all of which act to eradicate virally infected and neoplastic cells. Since CD8 T cells are important effectors of immune surveillance,[28] their presence in tumor tissues is a good prognostic factor for certain types of cancer.[29] Diminished CD8 T-cell function increases the susceptibility to squamous cell and basal cell carcinomas in the skin of humans and animal models. Tc1 and Tc17 cells, producing IFN-α and IL-17 respectively, are key mediators in allergic contact dermatitis[30] (Chapter 44). Tissue damage by cytotoxic CD8 T cells also contributes to the pathology of allergic contact dermatitis.

γδ T cells. Dendritic epidermal T cells (DETCs) bridge innate first line of defense and acquired immune mechanisms. In place of conventional αβ chain TCRs, they express a more restricted repertoire of γ and δ chain TCRs and reside in epithelial tissues of skin, the gut, lungs, and the reproductive tract.[31] In humans, γδ T cells account for at least 10% of the T cells in the epithelium. Epithelium-resident γδ T cells differ from circulating γδ T cells with respect to their development, selection, TCR diversity, and effector functions. Epithelial γδ T-cell activation is not restricted by MHC class I or class II molecules. Instead they recognize PAMP and DAMP molecules presented by nonclassic MHC molecules, such as CD1. They use alternate sets of costimulatory molecules, such as the junctional adhesion molecule–like protein (JAML) and its ligand, the coxsackievirus and adenovirus receptor (CAR). They are thought to be partially activated under static conditions. This enables them to rapidly mount responses to pathogenic stimuli. DETCs are involved in epidermal wound repair, can influence DC activation, and can directly present antigens.

Cytokines and Chemokines and the Adaptive Immune Response in Skin

A number of proinflammatory mediators are intimately involved in adaptive immunity in skin. IL-1 activates LC and dermal DC migration and maturation processes. Mature DCs produce T-cell differentiation factors IL-12 and IL-23, which are critical for the development of Th1 and Th17 cells, respectively.[32]

Chemokines facilitate leukocyte migration and thus skin inflammation and antigen-specific responses (see Table 19.2).[9,33] Secondary lymphoid chemokine (SLC, CCL21) attracts CCR7$^+$ DCs and T cells to lymph nodes, and stromal cell-derived factor 1 (SDF1; CXCL12) facilitates DC migration out of skin. CCL17 and CCL27, produced by keratinocytes and dermal cells, facilitate transmigration of circulating skin-homing T cells (expressing the chemokine receptors CCR4 and CCR10). Human skin–resident T cells are retained through interaction of their chemokine receptor CCR8 and its ligand CCL1 bound to the dermal matrix, which is produced by cutaneous cells.[33] During inflammation, CXCL1 and CXCL8/IL-8 expedite infiltration of granulocytes and monocytes into skin.

Cutaneous T-cell lymphoma (Chapter 79) is a disease in which a malignant population of skin-homing T cells accumulate in skin. Clinical manifestations include erythematous, pruritic patches, plaques, and skin tumors. The malignant T cells express CCR4 chemokine receptor. Clinical trials of humanized mAbs directed against the CCR4 receptor on T-cells (mogamulizumab) have shown promise.

MAST CELLS AND SKIN

Skin is a rich source of mast cells[22] (Chapter 23). These cells are derived from CD34$^+$ progenitors in bone marrow and migrate into many different tissues in the body, especially those at environmental interfaces. In skin, they are found in the dermis at a density of up to 20 000/mm^2. They are typically concentrated around the dermal microvasculature, appendages, and nerves. Metachromatic stains, such as toluidine blue and Giemsa, stain mast-cell granules purple, a distinguishing feature. Their cytoplasmic granules contain a plethora of preformed mediators that include histamine and heparin; the proteases tryptase, chymase, carboxypeptidase, arylsulfatase A, β-hexosaminodase, and β-glucuronidase; and the cytokines TNF-α, GM-CSF, IL-3, IL-4, IL-5, IL-6, IL-8, and IL-13.

Mast cells express the high affinity surface receptor for the Fc portion of the IgE molecule (FcεRI), which bind and retain IgE for long periods. Antigen-induced cross-linking of IgE bound to FcεRI initiates a sequence of calcium- and energy-dependent events that culminate in the fusion of granules to the plasma membrane and the release of granule contents. Degranulation releases potent prostanoids, histamines, and inflammatory cytokines that rapidly recruit inflammatory cells. Nonantigenic stimuli, such as opiates, C5a, anaphylatoxin, stem cell factor, and substance P, can activate degranulation through FcR-independent opioid, adenosine, and β-adrenergic receptors. The products and contents of mast-cell granules are causative agents of acute vasodilatation, edema, pruritus, and rapid influx of leukocytes and eosinophils into skin. Eosinophils are especially important for fighting parasitic infections (Chapter 24). Mast-cell activation is implicated in the pathogenesis of bullous pemphigoid, leukocytoclastic vasculitis, atopic dermatitis, allergic contact dermatitis, and mastocytosis (urticaria pigmentosa) and plays a prominent role in the pathogenesis of urticaria and angioedema.

ANTIBODIES AND SKIN

Less is known regarding B cells in skin. Although the role of IgG, IgA, and IgE antibodies in normal skin homeostasis is less certain, Igs play key roles in the pathogenesis of several different dermatological diseases. IgE-dependent skin diseases include urticaria and angioedema and bullous pemphigoid. Hyper-IgE syndrome (HIES) is an autosomal dominant disease that develops cutaneous manifestations of severe dermatitis, recurrent infections of staphylococcal abscesses, and, to a lower extent, cutaneous candidiasis.[33] Dominant negative mutations in the signal transducer and activator of transcription 3 (STAT3) gene are associated with this disease. Deficient STAT3 signaling leads to impaired β-defensin expression, which may explain some of the clinical manifestations of HIES.[34] Pathogenic IgG causes blistering disorders, including pemphigus, bullous pemphigoid, epidermolysis bullosa acquisita, and paraneoplastic pemphigus, as well as leukocytoclastic vasculitis. IgA is key to the pathogenesis of dermatitis herpetiformis, linear IgA bullous dermatosis, and IgA-mediated cutaneous vasculitis.

ULTRAVIOLET RADIATION AND CUTANEOUS IMMUNITY

Sunlight is the major environmental agent to which skin is exposed. Although the presence of sunlight is essential for life, injudicious exposure to wavelengths in the ultraviolet (UV) spectrum can lead to sunburn, aging of skin, skin cancers, and a variety of photosensitivity diseases, many of which have an immunological pathogenesis. Although much of the investigation into the immunological effects of UVR has been conducted in animal models, many of the observations have been corroborated in humans.

In mice, as in humans, chronic exposure to UVR results in the development of highly antigenic skin cancers that are capable of stimulating a vigorous antitumor response in untreated mice. Despite their antigenic nature, these tumors grow progressively in their original host. This apparent paradox was resolved in studies that showed that in addition to producing mutant neoplastic cells, UVR also impairs cell-mediated immune surveillance that normally eliminate mutant cells before they develop into clinically apparent tumors. Thus mutant cells develop tumors only in an environment of immune suppression. In fact, organ transplant recipients who are treated with immunosuppressive medications have a greatly increased risk of developing UV-induced skin cancers, and those tumors tend to behave more aggressively.[35]

Studies have shown that UVR mediates its effects, in part, by perturbing the function of skin APCs. UV-irradiated APCs are poor stimulators of Th1 cells but are able to activate Tregs, which promote antigen-specific immunological tolerance. The immunosuppressive effect is mediated, at least in part, through UV-induced enhanced production of the suppressive cytokine IL-10 and reduction of the Th1 activation cytokine IL-12. It may seem surprising that an environmental carcinogen, such as UVR, suppresses immunological function in skin. One proposed explanation is that altered epithelial proteins are constantly generated by UVR exposure, necessitating chronic induction of immune tolerance to UV-damaged proteins to preserve the integrity of the skin barrier.

The immunosuppressive effect of UVR has been exploited for therapeutic purposes. UVR phototherapy is used to manage pathogenic immune responses that cause certain skin diseases, such as psoriasis and atopic dermatitis.[36]

ON THE HORIZON

- Identification of T-helper type 17 (Th17) cells as important mediators of the inflammatory response has provided the impetus for development of interleukin-17 (IL-17) inhibitory antibodies for the treatment of psoriasis. Because IL-23 promotes Th17-cell development, antibodies that block that molecule are being evaluated as therapeutic agents as well.
- The availability of cytokine inhibitors provides an opportunity for physicians to broaden the spectrum of skin diseases for which they can be used. Agents that inhibit tumor necrosis factor-α (TNF-α), for example, are being evaluated for the management of immunologically mediated skin diseases, including hidradenitis suppurativa.
- Cell-surface receptors have been detected that allow specific subpopulations of T cells to migrate into skin. Agents that block these receptors are being evaluated as a treatment for cutaneous T-cell lymphoma and inflammatory diseases of skin, in which skin-homing T cells are fundamental to the disease process. This provides enhanced treatment specificity leaving the functions of those T cells that circulate to other parts of the body intact.
- Several different types of dendritic cells (DCs) are present in skin that are specialized in the subpopulations of T cells that they activate. Recognition of this fact provides the basis for methods to improve vaccination procedures. It is likely that in the next several years cutaneous DCs will be employed for vaccination against tumors, such as melanoma and selected infectious agents.
- As involvement of cytokines in cutaneous diseases continues to be clarified, they will be employed to manipulate the immune response in an expanding list of immunologically mediated skin diseases, such as vitiligo, alopecia areata, cutaneous lupus erythematosus, and dermatomyositis.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Amagai M. Pemphigus. In: Bolognia J, Jorizzo JL, Schafer JV, editors. Dermatology. St. Louis, Mo: Mosby/Elsevier; 2012. p. 461–74.
2. Yohn JJ, et al. Modulation of melanocyte intercellular adhesion molecule-1 by immune cytokines. J Invest Dermatol 1990;95(2):233–7.
3. Kupper TS, Fuhlbrigge RC. Immune surveillance in the skin: mechanisms and clinical consequences. Nat Rev Immunol 2004;4(3):211–22.
4. Heath WR, Carbone FR. The skin-resident and migratory immune system in steady state and memory: innate lymphocytes, dendritic cells and T cells. Nat Immunol 2013;14(10):978–85.
5. Nestle FO, et al. Skin immune sentinels in health and disease. Nat Rev Immunol 2009;9(10):679–91.
6. Watanabe H, et al. Activation of the IL-1β-processing inflammasome is involved in contact hypersensitivity. J Invest Dermatol 2007;127(8):1956–63.
7. Ferwerda B, et al. Human dectin-1 deficiency and mucocutaneous fungal infections. N Engl J Med 2009;361(18):1760–7.
8. Saijo S, et al. Dectin-2 recognition of α-mannans and induction of Th17 cell differentiation is essential for host defense against Candida albicans. Immunity 2010;32(5):681–91.
9. Tan SY, Roediger B, Weninger W. The role of chemokines in cutaneous immunosurveillance. Immunol Cell Biol 2015;93(4):337–46.
10. Aksentijevich I, et al. An autoinflammatory disease with deficiency of the interleukin-1-receptor antagonist. N Engl J Med 2009;360(23):2426–37.
11. Schauber J, Gallo RL. Antimicrobial peptides and the skin immune defense system. J Allergy Clin Immunol 2008;122(2):261–6.
12. Schauber J, et al. Injury enhances TLR2 function and antimicrobial peptide expression through a vitamin D-dependent mechanism. J Clin Invest 2007;117(3):803–11.
13. Dorschner RA, Williams MR, Gallo RL. Rosacea, the face of innate immunity. Br J Dermatol 2014;171(6):1282–4.
14. Morizane S, Gallo RL. Antimicrobial peptides in the pathogenesis of psoriasis. J Dermatol 2012;39(3):225–30.
15. Heath WR, Carbone FR. Dendritic cell subsets in primary and secondary T cell responses at body surfaces. Nat Immunol 2009;10(12):1237–44.
16. Clausen BE, Kel JM. Langerhans cells: critical regulators of skin immunity? Immunol Cell Biol 2010;88(4):351–60.
17. Henri S, et al. Disentangling the complexity of the skin dendritic cell network. Immunol Cell Biol 2010;88(4):366–75.
18. Clark RA. Skin-resident T cells: the ups and downs of on site immunity. J Invest Dermatol 2010;130(2):362–70.
19. Eyerich S, et al. Th22 cells represent a distinct human T cell subset involved in epidermal immunity and remodeling. J Invest Dermatol 2009;119(12):3573–85.
20. He D, et al. CD8+ IL-17-producing T cells are important in effector functions for the elicitation of contact hypersensitivity responses. J Immunol 2006;177(10):6852–8.
21. Schroder K, et al. Interferon-gamma: an overview of signals, mechanisms and functions. J Leuk Biol 2004;75(2):163–89.
22. Ogg G. Role of T cells in the pathogenesis of atopic dermatitis. Clin Exp Allergy 2009;39(3):310–16.
23. Simpson EL, et al. Dupilumab therapy provides clinically meaningful improvement in patient-reported outcomes (PROs): A phase IIb, randomized, placebo-controlled, clinical trial in adult patients with moderate to severe atopic dermatitis (AD). J Am Acad Dermatol 2016.
24. Li H, et al. Interleukin-17 mediated inflammatory responses are required for ultraviolet radiation-induced immune suppression. Photochem Photobiol 2015;91(1):235–41.
25. Beissert S, Schwarz A, Schwarz T. Regulatory T cells. J Invest Dermatol 2006;126(1):15–24.
26. Elmets CA, et al. Analysis of the mechanism of unresponsiveness produced by haptens painted on skin exposed to low dose ultraviolet radiation. J Exp Med 1983;158(3):781–94.
27. Yusuf N, et al. Antagonistic roles of CD4+ and CD8+ T-cells in 7,12-dimethylbenz(a)anthracene cutaneous carcinogenesis. Cancer Res 2008;68(10):3924–30.
28. Finn OJ. Cancer immunology. New Eng J Med 2008;358(25):2704–15.
29. He D, et al. IL-17 and IFN-gamma mediate the elicitation of contact hypersensitivity responses by different mechanisms and both are required for optimal responses. J Immunol 2009;183(2):1463–70.
30. Witherden DA, Havran WL. Molecular aspects of epithelial gammadelta T cell regulation. Trends Immunol 2011;32(6):265–71.
31. Pasparakis M, Haase I, Nestle FO. Mechanisms regulating skin immunity and inflammation. Nat Rev Immunol 2014;14(5):289–301.
32. Longley J, Duffy TP, Kohn S. The mast cell and mast cell disease. J Am Acad Dermatola 1995;32(4):545–61, quiz 562–4.
33. Minegishi Y, et al. Molecular explanation for the contradiction between systemic Th17 defect and localized bacterial infection in hyper-IgE syndrome. J Exp Med 2009;206(6):1291–301.
34. Kripke ML. Antigenicity of murine skin tumors induced by ultraviolet light. J Natl Cancer Inst 1974;53(5):1333–6.
35. Hartevelt MM, et al. Incidence of skin cancer after renal transplantation in The Netherlands. Transplantation 1990;49(3):506–9.
36. Krutmann J, Honigsmann H, Elmets CA. Dermatological phototherapy and photodiagnostic methods. 2nd ed. Berlin: Springer; 2009. p. xvi. 447 p.

MULTIPLE-CHOICE QUESTIONS

1. Pembrolizumab and nivolumab are neutralizing antibodies directed at the programmed death ligand type 1 (PD-L1) and ipilimumab is an antibody that neutralizes cytotoxic T lymphocyte antigen-4 (CTLA-4). They are used to treat which of the following cutaneous malignancies?
   A. Basal cell carcinoma
   B. Fibrosarcoma
   C. Kaposi sarcoma
   D. Melanoma
   E. Sebaceous gland carcinoma

2. Skin is a rich source of cytokines and chemokines. Which cell type does not produce them in skin?
   A. Fibroblasts
   B. Keratinocytes
   C. Mast cells
   D. Melanocytes
   E. All of the above produce cytokines in the skin

3. Cathelicidins are antimicrobial peptides that are produced by skin. They have been implicated in which of the following diseases?
   A. Atopic dermatitis
   B. Cutaneous T-cell lymphoma
   C. Dermatitis herpetiformis
   D. Hyper-IgE syndrome (HIES)
   E. Pemphigus

4. In a normal healthy individual, which of the following has the largest number of lymphocytes?
   A. Blood
   B. Brain
   C. Kidney
   D. Liver
   E. Skin

5. Which of the following is *true* about epidermal Langerhans cells?
   A. They possess a specialized organelle called the *Birbeck granule.*
   B. They are a type of myeloid dendritic cell (DC).
   C. Following cutaneous exposure to danger signals, they extrude their processes to capture external antigens.
   D. After cutaneous antigen exposure, they migrate to draining lymph nodes.
   E. All of the above are true.

20

Host Defenses at Mucosal Surfaces

Prosper N. Boyaka, Kohtaro Fujihashi

Mammals have evolved a sophisticated network of cells and molecules that serves to maintain homeostasis on exposed mucosal surfaces. This system is anatomically and functionally distinct from its bloodborne counterpart and is strategically located at the portals through which most pathogenic microorganisms enter the body. This specific branch of the immune system may have developed in response to the size of the mucosal surfaces, which cover an area of ≈70 m^2 in the airways and ≈400 m^2 in the gut of an adult human, and the large numbers of commensal bacteria and exogenous antigens to which mucosa are exposed.[1,2]

THE INNATE MUCOSAL DEFENSE SYSTEM

Cells and molecules that contribute to innate defense of the mucosa include the physical barrier provided by epithelial cells, the movement of the epithelial cilia, the production of mucus by goblet cells, the secretion of molecules with innate antimicrobial activity, and the cytolytic activity of natural killer (NK) cells (Fig. 20.1A). Innate lymphoid cells (ILCs; Chapter 3) were identified recently as key players in innate mucosal immunity (see Fig. 20.1B). In concert with the commensal microbiota (Chapter 14), these innate mechanisms provide a first line of defense against exogenous antigens and invading pathogens.

KEY CONCEPTS

Innate Defenses of the Mucosal Immune System

The innate defenses of the mucosal immune system provide a first line of defense against exogenous antigens and invading pathogens. These defenses include:

- *Physical barriers:* the epithelium, the epithelial cilia, goblet cell mucus production
- *Mucosal antimicrobial molecules:* Paneth cell production of α defensins in the small intestine; epithelial cell production of β defensins in the oral mucosa, trachea, bronchi, mammary glands, and salivary glands; lactoferrin, lysozyme, lactoperoxidase, and secretory leukocyte protease inhibitor (SLPI)
- *Cellular innate immunity:* mucosal natural killer (NK) cells, innate lymphoid cells (ILCs), dendritic cells (DCs). and polymorphonuclear neutrophils (PMNs).

Epithelial Cells and Other Effectors of the Mucosal Physical Barrier

All mucosal surfaces are covered by epithelial cells, which contribute to their selective barrier function. In the gastrointestinal (GI) tract, tightly joined enterocytes constitute the cellular component of the physical barrier and are covered by a blanket of mucus. Mucus consists of glycoproteins secreted into the lumen by goblet cells. This layer of glycoproteins interferes with the attachment of microorganisms to the mucosal surface. The replacement of damaged or infected enterocytes by crypt epithelial cells, which differentiate into enterocytes as they migrate toward the desquamation zone at the villus tip, ensures the integrity of this barrier. Multilayered squamous epithelial cells cover other mucosal surfaces, including the oral cavity, pharynx, tonsils, urethra, and vagina. These epithelia lack tight junctions. Instead, mucus coats the intercellular space between the lower stratified epithelial cell layers. Polymeric immunoglobulin A (pIgA) and commensal microbes support the physical barrier function of mucosal tissues. Perturbation of the commensal microbiome, which is present in the mucus ecosystem, facilitates opportunist infections by pathogens, such as *Clostridium difficile.*

Defensins and Other Mucosal Antimicrobial Peptides

Selected epithelial cell subsets contribute to innate responses through the production of antimicrobial peptides, iron transporters, and enzymes. Defensins are 30–40 amino acid β-sheet peptides with antiviral activity and antimicrobial effects similar to those of antibiotics. Defensins are structurally segregated into α and β categories. α-Defensins are secreted by tracheal epithelial cells and by Paneth cells in intestinal crypts. α-Defensins are homologous to peptide mediators of nonoxidative microbial cell killing in neutrophils (termed *human neutrophil peptides* [HNPs]).[3] Human β-defensin 1 (HBD-1) is expressed in the epithelial cells of the oral mucosa, trachea, bronchi, mammary glands, and salivary glands, whereas HBD-5 is expressed in the gut. Inflammatory cytokines (Chapter 9), including interleukin-1 (IL-1), IL-17, tumor necrosis factor-α (TNF-α) and bacterial lipopolysaccharide (LPS), regulate defensin production.

Other antimicrobial products of the epithelium include lactoferrin, lysozyme, peroxidases, secretory phospholipase A2 (S-PLA2), and cathelin-associated peptides. Lactoferrin, a member of the transferrin family, is found in exocrine secretions. High concentrations of lysozyme (1209–1325 μg/mL) are found in tears, saliva, colostrum, serum, and urine. Human milk contains lysozyme in concentrations ranging from 20 to 245 μg/mL, depending on the lactation period. Milk leukocytes produce myeloperoxidase (MPO), and mammary gland cells produce human lactoperoxidase (hLPO). Both peroxidases display properties similar to those of human salivary peroxidases (hSPO). S-PLA2 is released by Paneth cells. Secretory leukocyte protease inhibitor (SLPI) is found in human saliva, nasal secretions, tears, cervical mucus, and seminal fluid. It is believed to be responsible for the anti–human immunodeficiency virus (HIV) properties of external secretions.

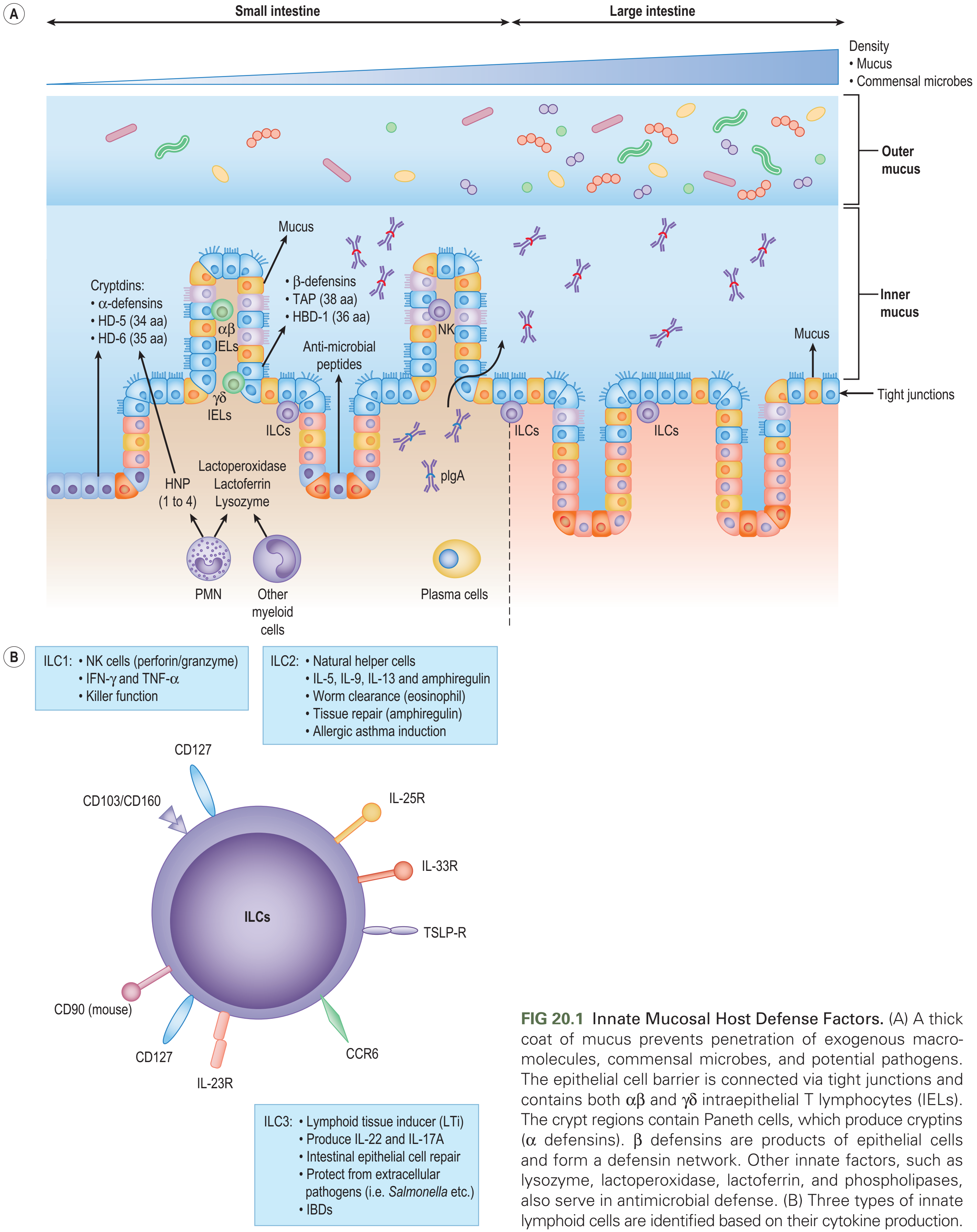

A
Small intestine
Large intestine
Density
• Mucus
• Commensal microbes
Outer mucus
Inner mucus
Mucus
Cryptdins:
• α-defensins
• HD-5 (34 aa)
• HD-6 (35 aa)
αβ IELs
γδ IELs
• β-defensins
• TAP (38 aa)
• HBD-1 (36 aa)
NK
Anti-microbial peptides
Mucus
Tight junctions
ILCs
ILCs
ILCs
HNP (1 to 4)
Lactoperoxidase
Lactoferrin
Lysozyme
pIgA
PMN
Other myeloid cells
Plasma cells
B
ILC1: • NK cells (perforin/granzyme)
• IFN-γ and TNF-α
• Killer function
ILC2: • Natural helper cells
• IL-5, IL-9, IL-13 and amphiregulin
• Worm clearance (eosinophil)
• Tissue repair (amphiregulin)
• Allergic asthma induction
CD127
CD103/CD160
IL-25R
IL-33R
ILCs
TSLP-R
CD90 (mouse)
CD127
IL-23R
CCR6
ILC3: • Lymphoid tissue inducer (LTi)
• Produce IL-22 and IL-17A
• Intestinal epithelial cell repair
• Protect from extracellular pathogens (i.e. *Salmonella* etc.)
• IBDs

FIG 20.1 Innate Mucosal Host Defense Factors. (A) A thick coat of mucus prevents penetration of exogenous macromolecules, commensal microbes, and potential pathogens. The epithelial cell barrier is connected via tight junctions and contains both αβ and γδ intraepithelial T lymphocytes (IELs). The crypt regions contain Paneth cells, which produce cryptins (α defensins). β defensins are products of epithelial cells and form a defensin network. Other innate factors, such as lysozyme, lactoperoxidase, lactoferrin, and phospholipases, also serve in antimicrobial defense. (B) Three types of innate lymphoid cells are identified based on their cytokine production.

Mucosal Innate Lymphoid Cells

ILCs constitute recently described groups of lymphoid cells that are predominantly found in mucosal tissues. ILCs lack T- or B-cell receptors. They contribute to innate regulation of homeostasis through their ability to rapidly produce cytokines. Similar to CD4 T helper (Th) cells (Chapter 16), ILCs can be divided into three major groups according to cytokines they produce.[4,5] Group 1 ILCs (ILC1) produce interferon-γ (IFN-γ) and are considered similar to Th1 cells. Cells with ILC1 phenotypes express CD103/CD160 and CD127 and include NK cells, which are large granular lymphocytes found in the lamina propria and the intraepithelial compartment. Group 2 ILCs are similar to Th2 cells and produce IL-5, IL-9, IL-13, and amphiregulin. ILC2 are responsible for innate responses in allergies and asthma. As with Th17 cells, group 3 ILCs (ILC3) secrete IL-17 and/or IL-22. Although lymphoid tissue inducer (LTi) cells are ILC3, functional properties of these cells are distinct from other mucosal ILC3.[4,5]

Two types of ILC1 occur in tonsils and the mucosa of the GI tract. $CD103^+$ $CD160^+$ ILCs are found in the intraepithelium of the intestines and produce perforin and granzyme like NK cells. This ILC1 subset requires Nfil3 and T-bet transcription factors for its development. An ILC1 population expressing low levels of RORγt and aryl hydrocarbon receptor (Ahr) similar to ILC3 was identified in humans. These ILCs express IL-7Rα (CD127), but not conventional lymphocyte lineage or NK cell markers. Both subsets of ILC1 may be involved in the induction of inflammatory bowel disease (IBD).[4] ILC2 are seen in the lungs, upper respiratory mucosa, gut, and skin. By producing IL-5, IL-13, and amphiregulin, they play key roles in the clearance of parasites, including *Nippostrongylus brasiliensis* and *Trichuris muris.* ILC2 are also important in the development of asthma and allergies, including atopic dermatitis. Although ILC2 were reported to promote airway hypersensitivity during acute influenza virus infection, the amphiregulin they produce could also help maintain lung epithelial cell homeostasis.[4]

ILC3 in both humans and mice showed heterogenic cytokine profiles. Thus human ILC3 bearing NKp44 and CCR6 produce IL-22 alone. However, human ILC3 isolated from patients with Crohn disease produce IL-17 and IFN-γ. Mouse LTi and LTi-like ILC3 produce both IL-17 and IL-22. $NKp46^+$ ILC3 secrete IL-22, as well as IFN-γ under certain conditions, but not IL-17. Another subset of mouse ILC3 that produces IFN-γ, IL-17, and IL-22 can be found in the large intestine. ILC3-derived IL-22 induces antimicrobial peptide responses by intestinal epithelial cells (IECs). Furthermore, IL-22 from ILC3 was reported to enhance innate immunity against *Salmonella* and bacterial infection secondary to influenza virus infection.[4,5] Since ILC3 are also involved in the induction of colonic inflammation and colorectal cancer, they may act as a double-edged sword in the mucosal immune system.[4,5]

A COMMON MUCOSAL ADAPTIVE IMMUNE SYSTEM

Higher-order mammals have developed an organized secondary lymphoid tissue system in the GI and upper respiratory tracts. Gut-associated lymphoreticular tissues (GALTs) include Peyer patches (PPs), the appendix, and solitary lymphoid nodules in the GI tract (Chapter 2). Tonsils and adenoids comprise nasal-associated lymphoid tissues (NALTs). Experimental animals, such as rabbits, rats, and guinea pigs, exhibit organized bronchus-associated lymphoid tissues (BALTs) that rarely occur in human airway branches.[1] Together, GALTs and NALTs in humans and GALTs, BALTs, and NALTs in experimental species are termed *mucosa-associated lymphoreticular tissue* (MALT).

The vast areas of the mucosal immune system characterized by diffuse collections of lymphoid cells are termed *effector tissues.* These include the interstitial tissues of the mammary, lacrimal, salivary, sweat, and all other exocrine glands, as well as the lamina propria and the epithelium of the GI tract. The lamina propria areas of the upper respiratory and genitourinary tracts are also lymphoid effector sites. MALT is connected with effector sites through the migratory patterns of effector cells.

KEY CONCEPTS

The Common Mucosal Immune System

The term *mucosa-associated lymphoreticular tissue* (MALT) comprises discrete and diffuse collections of lymphoid tissues that share distinctive features, including a unique type of epithelium, a distinct architecture, a unique set of antigen-presenting cells (APCs), and B cells, where switching to immunoglobulin A (IgA) predominates. The involved tissues include:

- *Gut-associated lymphoid tissues (GALTs):* Peyer patches (PPs), the appendix, and solitary lymphoid nodules in the gastrointestinal (GI) tract
- *Nasal-associated lymphoid tissues (NALTs):* tonsils and adenoids
- *Effector tissues:* the interstitial tissues of the mammary, lacrimal, salivary, sweat, and all other exocrine glands; the lamina propria and the epithelium of the GI tract; and the lamina propria areas of the upper respiratory and genitourinary tracts.

MALT as an Inductive Site

MALT has a unique type of epithelium for antigen uptake. Its features include a characteristic architecture, antigen-presenting–cell (APC) and B cell areas with germinal centers where switches to IgA predominate. The columnar epithelium that covers MALT is infiltrated with lymphocytes and APCs, leading to the term follicle-associated epithelium (FAE). Lacking goblet cells, the FAE is covered with far less mucus than normal enterocytes. Soluble and particulate luminal antigens are taken up by microfold (M) cells and are delivered to adjacent APCs. M cells have been described in PPs, the appendix, and tonsils and represent 10–15% of cells within the FAE.[6] M cells are also found in isolated lymphoid follicles (ILFs) and at the tips of the villus, where they are termed *villous M cells.*[7] The microvilli of these cells, which are less dense than those of adjacent enterocytes, offer a portal of entry into MALT (Fig. 20.2). The M cell is often identified by an invagination of the basolateral membrane into a "pocket" normally occupied by lymphocytes and APCs (Fig. 20.3).

M cells appear ideal for antigen uptake owing to a well-developed microvesicle system that contains endosomes. However, it remains unclear whether M cells act as classic APCs. M cells also provide a portal of entry for some pathogens, such as invasive strains of *Salmonella typhimurium*, but not for noninvasive strains of *S. typhimurium* and reoviruses.

Gut-Associated Lymphoreticular Tissues

Each PP contains a dome region that is positioned under the FAE. This dome region is populated by T cells, B cells, macrophages (MØs), and dendritic cells (DCs). It includes follicles that contain germinal centers. The presence of all three major APC types in the dome (*i.e.*, memory B cells, MØs, and DCs) makes it likely that antigen uptake occurs immediately following

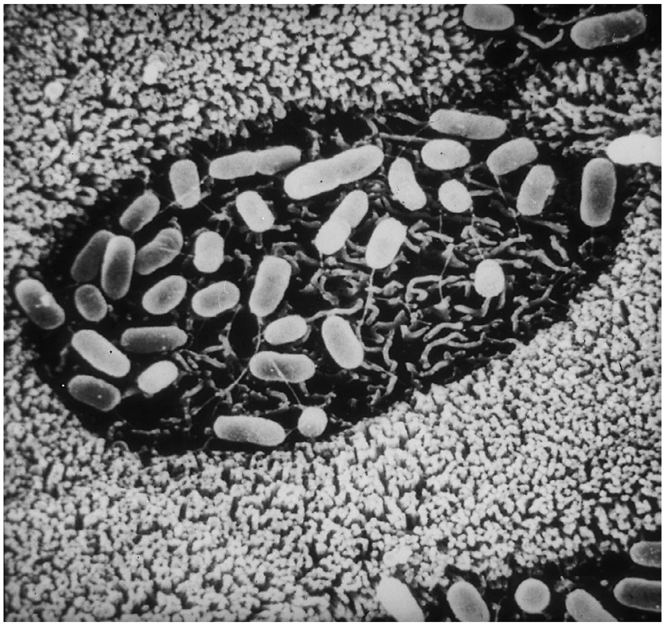


FIG 20.2 **The Microfold (M) Cell.** A scanning electron micrograph of an M cell with adjacent enterocytes. The M cell has selectively bound *Escherichia coli* 0157. Note that a thick brush border is lacking, facilitating the binding and uptake of microparticles. (Courtesy of Dr. Tatsuo Yamamoto, Niigata University.)

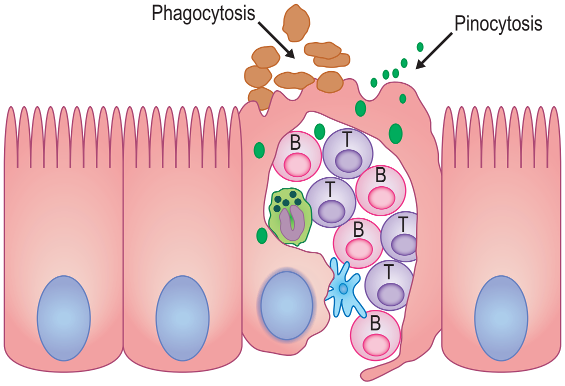


FIG 20.3 **Microanatomical Features of Microfold (M) Cells.** The M cell forms a "pocket" containing memory lymphocytes. It actively pinocytoses soluble antigens and phagocytoses particulates such as viruses, bacteria, and microspheres. (Courtesy of Dr. Svein Steinsvoll, University of Oslo.)

release from M cells (Fig. 20.4A). M cell pockets in PPs contain approximately equal numbers of T and B cells, but fewer MØs. Approximately 75% of the T cells are Th cells.

GALT B-cell follicles are enriched in IgA-bearing B cells,[8] suggesting that they are major sites for B-cell μ to α switching (Chapters 4 and 7). The interfollicular regions of PPs contain high endothelial venules (HEVs) (Chapter 2, 11). Both CD4 and CD8 TCR$\alpha\beta$ T cells are found in these interfollicular regions, with CD4 T cells representing the predominant phenotype. Both naïve and memory T cells are present in PPs, with one-third in cell cycle (see Fig. 20.4; Table 20.1). Lymphotoxin-α (LT-α), lymphotoxin-β (LT-β), and TNF-α (Chapter 9) are critical for lymphoid tissue organogenesis (Chapter 2). LT-$\alpha^{-/-}$ mice are deficient in secondary lymph nodes,[9,10] whereas LT-$\beta^{-/-}$ mice have mesenteric and cervical lymph nodes but lack peripheral lymph nodes and PPs. Tumor necrosis factor–receptor I (TNF-RI)$^{-/-}$ mice lack or display abnormal PP structures, whereas TNF-$\alpha^{-/-}$ mice exhibit normal patches.

TABLE 20.1 **Major T-Cell Subpopulations Associated With Murine Peyer Patches**

T-Cell Phenotype	Percentage of Total T Cells
CD3$^+\alpha\beta$ T-cell receptor (TCR)$^+$	95–97
CD3$^+\gamma\delta$ TCR$^+$	3–5
CD3$^+$, CD4$^+$ (precursors of T-helper [Th] cells)	65–70
CD3$^+$, CD8$^+$ (precursors of cytotoxic T lymphocytes [CTLs])	30–35
Naïve (CD45RB^{Hi})	50–60
Memory (CD45RB^{Lo}, CD45RO^{Hi})	40–50
Blasts (in cell cycle)	30–35

Nasal-Associated Lymphoid Tissues

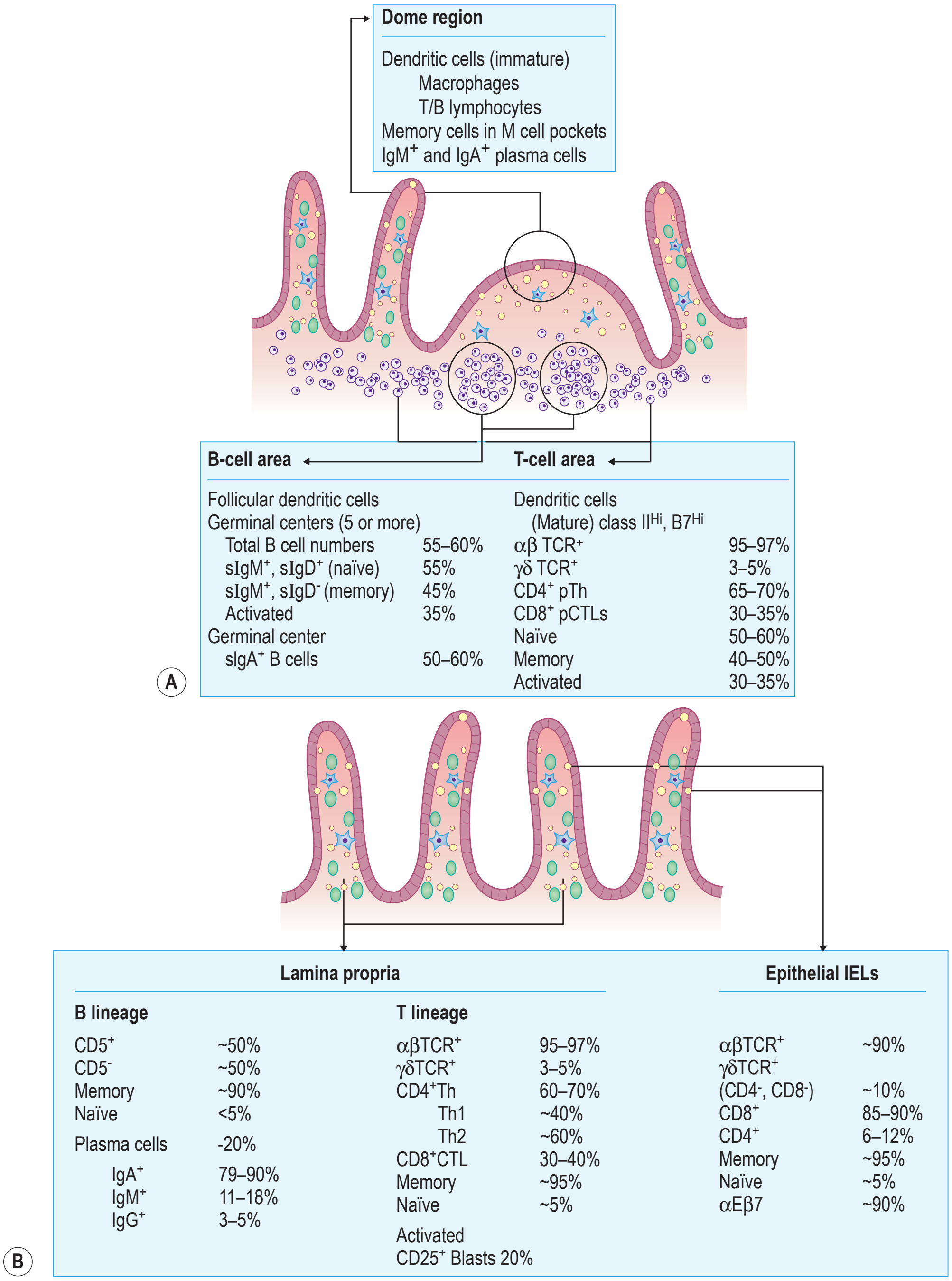
Strategically positioned at the entry of the respiratory and the digestive tracts are the accumulations of lymphoid tissues that comprise the palatine, lingual, and nasopharyngeal tonsils, which collectively form the Waldeyer ring. These tissues resemble both lymph nodes and PPs, including an FAE with M cells in the tonsillar crypts that is essential for selective antigen uptake (see Fig. 20.3). Germinal centers containing B and T cells, plasma cells, and APCs are also present. Tonsillar tissues can serve as a source of precursors of IgA plasma cells found in the upper aerodigestive tracts, as well as inductive sites for systemic and mucosal immune responses.[11] The LT$\alpha_1\beta_2$ signaling pathway is essential for the maintenance, but not the initiation, of NALT organogenesis.[7] Signaling via the IL-7/IL-7R and the L-selectin/peripheral lymph node addressin (PNAd) adhesion molecules both play important roles in the organization of NALTs.[7]

Other Sites for Mucosal Induction of an Immune Response

The follicular structures analogous to PPs in the large intestine are known as rectal-associated lymphoid tissues (RALTs). Unlike most other mucosal tissues, the large intestine lamina propria is home to more IgA2- than IgA1-producing cells.[8] Eye drop administration of antigen elicits secretory immunoglobulin A (SIgA) antibody responses in ocular and nasal mucosae. Thus both tear-duct associated lymphoid tissue (TALT) and conjunctiva-associated lymphoid tissue (CALT) take up antigens for the initiation of mucosal immune responses as a component of MALT.[12,13] Immunization via the epicutaneous and sublingual routes is emerging as a potential method for induction of mucosal immunity, and structures facilitating these responses are being investigated.

LYMPHOCYTE HOMING INTO MUCOSAL COMPARTMENTS

Mesenteric lymph node cells of orally immunized animals can repopulate the lamina propria of the gut, mammary glands, lacrimal glands, and salivary glands with antigen-specific IgA plasma cells (see Fig. 20.4B),[1] pointing to the existence of a "common" mucosal immune system. This concept has undergone



Dome region

Dendritic cells (immature)
Macrophages
T/B lymphocytes
Memory cells in M cell pockets
IgM$^+$ and IgA$^+$ plasma cells

B-cell area

Follicular dendritic cells	
Germinal centers (5 or more)	
Total B cell numbers	55–60%
sIgM$^+$, sIgD$^+$ (naïve)	55%
sIgM$^+$, sIgD$^-$ (memory)	45%
Activated	35%
Germinal center	
sIgA$^+$ B cells	50–60%

T-cell area

Dendritic cells	
(Mature) class II^{Hi}, B7^{Hi}	
αβ TCR$^+$	95–97%
γδ TCR$^+$	3–5%
CD4$^+$ pTh	65–70%
CD8$^+$ pCTLs	30–35%
Naïve	50–60%
Memory	40–50%
Activated	30–35%

Lamina propria

B lineage

CD5$^+$	~50%
CD5$^-$	~50%
Memory	~90%
Naïve	<5%
Plasma cells	-20%
IgA$^+$	79–90%
IgM$^+$	11–18%
IgG$^+$	3–5%

T lineage

αβTCR$^+$	95–97%
γδTCR$^+$	3–5%
CD4$^+$Th	60–70%
Th1	~40%
Th2	~60%
CD8$^+$CTL	30–40%
Memory	~95%
Naïve	~5%
Activated	
CD25$^+$ Blasts 20%	

Epithelial IELs

αβTCR$^+$	~90%
γδTCR$^+$	
(CD4$^-$, CD8$^-$)	~10%
CD8$^+$	85–90%
CD4$^+$	6–12%
Memory	~95%
Naïve	~5%
αEβ7	~90%

FIG 20.4 Structural Features and Cellular Components of Gut-Associated Lymphoreticular Tissues (GALTs). (A) The dome region is covered by the follicle-associated epithelium (FAE) with its characteristic microfold or microfold (M) cells. Major features of the dome include M cells with lymphocyte pockets, scattered plasma cells, and immature dendritic cells (DCs). The B-cell area contains five or more germinal centers with high frequencies of surface immunoglobulin (Ig) A$^+$ B cells. The adjacent T-cell area contains mature interdigitating DCs and precursors of CD4$^+$ Th and CD8$^+$ CTL. (B) Structural features and cellular characteristics of mucosal effector sites. The lamina propria is equally populated by B1 and B2 cells, both of which differentiate into IgA$^+$ plasma cells. Note that memory B and T lymphocytes are also both present in this compartment. Although intraepithelial lymphocytes (IELs) in human are mainly T-cell receptor (TCR)αβ$^+$, significant numbers of TCRγδ$^+$ T cells are also found in this compartment.

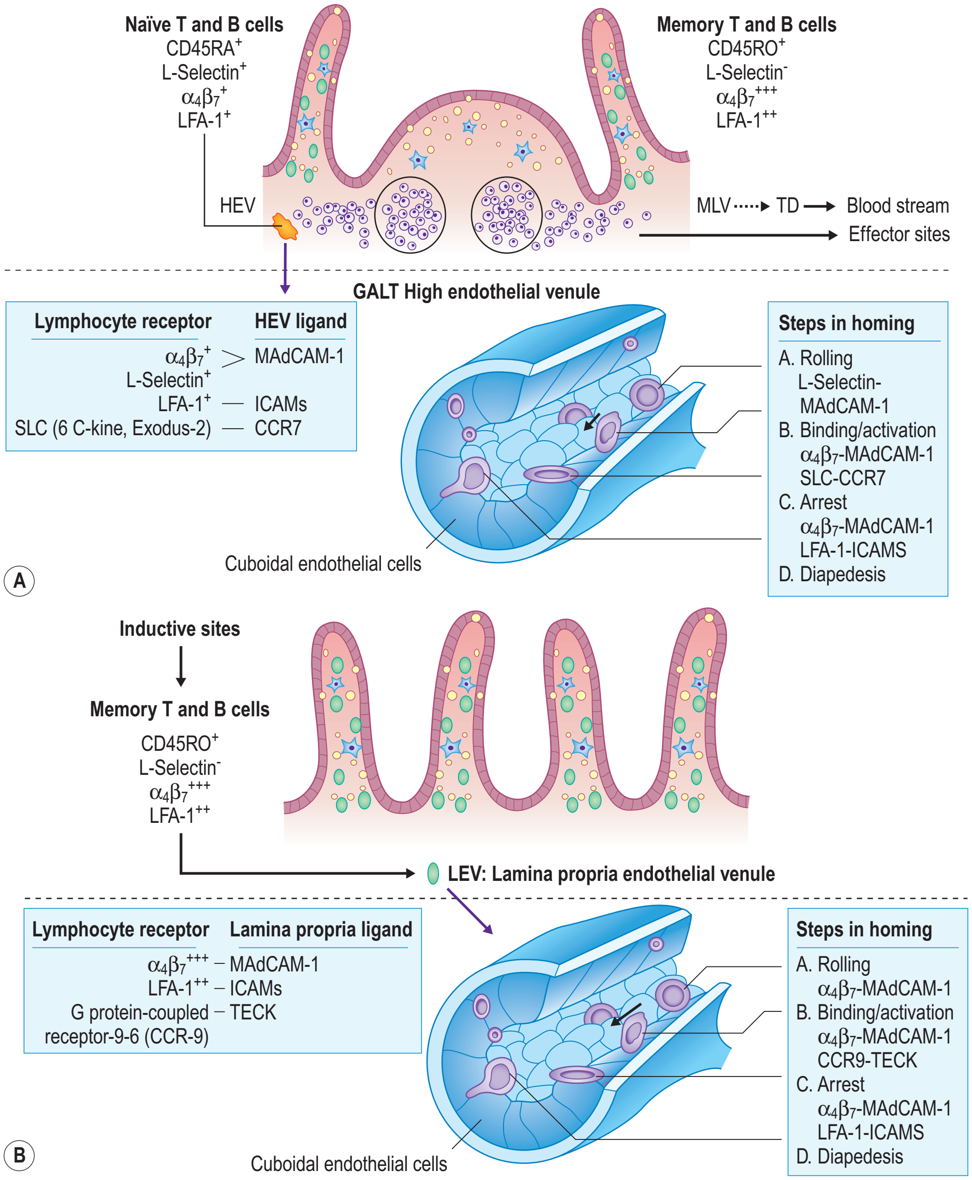

Naïve T and B cells
CD45RA$^+$
L-Selectin$^+$
$\alpha_4\beta_7^+$
LFA-1$^+$
HEV
Memory T and B cells
CD45RO$^+$
L-Selectin$^-$
$\alpha_4\beta_7^{+++}$
LFA-1^{++}
MLV → TD → Blood stream
Effector sites
GALT High endothelial venule
Lymphocyte receptor
HEV ligand
$\alpha_4\beta_7^+$ > MAdCAM-1
L-Selectin$^+$
LFA-1$^+$ — ICAMs
SLC (6 C-kine, Exodus-2) — CCR7
Steps in homing
A. Rolling
L-Selectin-
MAdCAM-1
B. Binding/activation
$\alpha_4\beta_7$-MAdCAM-1
SLC-CCR7
C. Arrest
$\alpha_4\beta_7$-MAdCAM-1
LFA-1-ICAMS
D. Diapedesis
Cuboidal endothelial cells
A
Inductive sites
Memory T and B cells
CD45RO$^+$
L-Selectin$^-$
$\alpha_4\beta_7^{+++}$
LFA-1^{++}
LEV: Lamina propria endothelial venule
Lymphocyte receptor
Lamina propria ligand
$\alpha_4\beta_7^{+++}$ – MAdCAM-1
LFA-1^{++} – ICAMs
G protein-coupled – TECK
receptor-9-6 (CCR-9)
Steps in homing
A. Rolling
$\alpha_4\beta_7$-MAdCAM-1
B. Binding/activation
$\alpha_4\beta_7$-MAdCAM-1
CCR9-TECK
C. Arrest
$\alpha_4\beta_7$-MAdCAM-1
LFA-1-ICAMS
D. Diapedesis
Cuboidal endothelial cells
B

FIG 20.5 Structural Features and Lymphocyte Homing to Gut-Associated Lymphoreticular Tissues (GALTs). (A) High endothelial venules (HEVs) occur in T-cell areas and express the ligands mucosal addressin cell adhesion molecule-1 (MAdCAM-1), intercellular adhesion molecule 1 (ICAM-1), and CCR7. Naïve T and B cells, which are L-selectin$^+$, $\alpha_4\beta_7^+$, and leukocyte function–associated antigen-1 (LFA-1)$^+$, all participate in rolling, binding activation, arrest, and diapedesis in the HEV. Memory B and T cells express $\alpha_4\beta_7$ and LFA-1 at a higher level. (B) Lymphocyte receptors and addressin ligands involved in homing to mucosal effector sites of the gastrointestinal (GI) tract. The majority of B and T cells exhibit a memory phenotype with coexpression of high levels of $\alpha_4\beta_7$ and LFA-1. The expression of the G-protein–coupled receptor CCR9 allows the homing steps that occur on lamina propria venules.

further refinement, with evidence that migration of cells into and from NALTs follows rules different from those for GALTs and the GI tract.

Lymphocyte Homing in the GI Tract

Naïve lymphocytes enter mucosal or systemic lymphoid tissues from blood through specialized HEVs (Fig. 20.5A) (Chapter 11). In GALTs, HEVs are present in the interfollicular T-cell zones.[14] In effector sites, such as the lamina propria of the gut, the HEVs tend to occur near villi crypts (see Fig. 20.5B). Mucosal addressin cell adhesion molecule-1 (MAdCAM-1) is the most important addressin expressed by PP HEVs or lamina propria venules (LPVs). PNAd and vascular cell adhesion molecule 1 (VCAM-1) are the principal addressins expressed by peripheral lymph node and skin HEVs, respectively (Chapter 11).

Integrins are a large class of homing receptors characterized by heterodimeric α and β chains (Chapter 11). In general, the β_1 integrin characterizes the homing receptor for skin, whereas the β_7 integrin characterizes the receptor for the gut. The pairing of α_4 with β_7 is thus responsible for lymphocyte binding to MAdCAM-1, which is expressed on HEVs in PPs and gut LPVs (see Fig. 20.5).[14] The C-type lectins—L-, E-, and P-selectins (Chapter 11)—also serve as homing receptors. For example, L-selectin can bind to carbohydrate-decorated MAdCAM-1 and is an important initial receptor for homing into GALT HEVs.

Chemokines (Chapter 10) are also involved in immune-cell homing in mucosal tissues. For example, loss of secondary lymphoid tissue chemokine (SLC) results in lack of naïve T-cell or DC migration into the spleen or PPs. Conversely, memory $\alpha_4\beta_7^{hi}$ T cells that express the receptor for thymus-expressed chemokine (TECK), CCR9, migrate into the lamina propria of the GI tract. Both human $\alpha_E\beta_7^+$ and $\alpha_4\beta_7^{hi}$ CD8 T cells express CCR9, suggesting that TECK-CCR9 is also involved in lymphocyte homing and the arrest of IELs in the GI tract epithelium (see Fig. 20.4) (Chapters 10, 11).

PPs and GALTs contain both naïve and memory T- and B-cell subsets, whereas the lamina propria consists of memory T and B cells and terminally differentiated plasma cells (see Table 20.1 and Fig. 20.4). Naïve lymphocytes destined for GALTs express L-selectin, moderate levels of $\alpha_4\beta_7$ ($\alpha_4\beta_7^+$) and lymphocyte function–associated antigen-1 (LFA-1). Memory lymphocytes destined for lamina propria express higher levels of $\alpha_4\beta_7$ ($\alpha_4\beta_7^{hi}$) and lack L-selectin. Initial rolling is dependent on $\alpha_4\beta_7$ interactions with LPV MAdCAM-1. Activation-dependent binding and extravasation require LFA-1–ICAM binding. $\alpha_4\beta_7$ also mediates binding to E-cadherin, and CCR9 expression can result in activation-dependent entry into the epithelial cell compartment.

Cryosections of human tissues have revealed naïve lymphocytes in HEVs that express both L-selectin and $\alpha_4\beta_7$, whereas memory lymphocytes in efferent lymphatics express $\alpha_4\beta_7$, but not L-selectin. The majority of cells in mesenteric lymph nodes, including B-cell blasts, tend to be of the memory phenotype and are $\alpha_4\beta_7^{hi}$, L-selectin^{lo}. Ig-containing B-cell blasts also express high levels of $\alpha_4\beta_7$. This separation of naïve and memory T and B cells for entry into GALT HEVs or LPVs has important implications in vaccine development (Chapter 90).

An oral cholera vaccine was reported to elicit transient IgA antibody-forming cells (AFCs) in blood and subsequent IgA anticholera toxin AFCs in duodenal tissues.[15,16] In a separate study, peripheral blood AFCs induced after parenteral immunization were L-selectin$^+$, whereas those induced after oral and rectal immunization were predominantly $\alpha_4\beta_7^+$ AFCs.[15,16] In the latter study, most of the AFCs produced IgA, but some also expressed IgG. After nasal immunization, AFCs expressed both L-selectin and $\alpha_4\beta_7$ homing receptors. The APCs in GALTs were shown to produce high levels of retinoic acid, which promotes expression of $\alpha_4\beta_7$. Thus enteric immunization of GALTs more effectively triggers $\alpha_4\beta_7$ memory IgA and IgG B cells, which can then migrate into the bloodstream.

Lymphocyte Homing in NALTs and Lung-Associated Tissues

Unlike PP HEVs, which are found in T-cell zones, murine NALT HEVs are found in B-cell zones and express PNAd, either alone or associated with MAdCAM-1. Moreover, anti–L-selectin antibodies—but not anti–MAdCAM-1 antibodies—block the binding of naïve lymphocytes to NALT HEVs, suggesting a role for L-selectin and PNAd in the binding of naïve lymphocytes to these HEVs.[17]

During pulmonary immune responses, induction of VCAM-1, E-selectin, and P-selectin in the pulmonary vasculature is matched by increased expression of P-selectin ligand on peripheral blood CD4$^+$ and CD8$^+$ T cells.[18] As the cells accumulate in the bronchoalveolar lavage (BAL) fluid, the number of cells that express P-selectin ligand in blood declines. Very late antigen-4 (VLA-4) appears to be involved, as migration of VLA-4$^+$ T cells into BAL fluid is impaired following treatment with anti-α_4 antibody. Following systemic immunization, most NALT effector B cells express L-selectin, with only a few cells expressing $\alpha_4\beta_7$.[19] Effector B cells induced by nasal immunization display a more promiscuous pattern of adhesion molecules, with a large majority expressing both L-selectin and $\alpha_4\beta_7$.

The Common Mucosal Immune System Revisited

The homing pattern that has been elucidated in the GI tract after immunization of GALTs has been the model for all mucosal immune sites. As summarized above, the specific set of homing receptors and ligand addressins expressed in the GI tract are absent in NALTs and associated lymph nodes. It remains possible, and even likely, that memory lymphocytes from the gut may enter NALTs for additional priming and reprogramming of homing receptors. Likewise, memory lymphocytes induced in NALTs may traffic to lung and genitourinary tract tissues as well as to the GI tract. Thus the rules for the homing of naïve lymphocytes precursors to NALTs need to be more clearly defined.

INDUCTION OF MUCOSAL IMMUNITY

Mucosal Immune responses are typified by SIgA antibodies, the predominant Ig isotype in external secretions. The resistance of SIgA to endogenous proteases makes them uniquely suited to protect mucosal surfaces. The development of mucosal adaptive immunity requires cytokine signals from CD4, as well as from CD8, T cells; DCs, MØs, and B cells; and nonclassic APCs (*e.g.*, epithelial cells). B cell commitment (Cμ to Cα switching) and B-/T-cell interactions are of central importance for induction of pIgA-producing cells.

Mucosal Antigen-Presenting Cells

Large macromolecules are taken up by M cells in the GI tract. N418$^+$, 2A1$^+$, NLDC-145$^-$, M342$^-$ DCs form a dense layer of cells in the subepithelial dome (SED) just beneath the follicle epithelium, where CD4 T cells can be found,[20] whereas N418$^+$, 2A1$^+$, NLDC-145$^+$, and M342$^+$ DCs populate the interfollicular T-cell regions, where both CD4 and CD8 T cells reside. DCs in the dome region are immature, highly endocytic, and express low levels of major histocompatibility complex (MHC) (Chapter 5) and B7 molecules. DCs in the T-cell area are mature, with low endocytic activity and high levels of MHC class I and II molecules and B7 molecule expression (see Fig. 20.4A). DCs are also found in NALTs, where they play essentially the same role as in GALTs.

IECs express MHC class II and class I molecules and present peptides to primed CD4 and CD8 T cells. Human and murine IECs also express CD1d, a nonclassic MHC class I molecule involved in the presentation of lipid and glycolipid antigens (Chapter 5).

CD4 T-Helper Cell Subsets in Mucosal Immunity

Th cells are classified as Th1, Th2, Th17, regulatory T cells (Tregs), or follicular Th (Tfh) cells according to the cytokines they produce (Chapter 16). Th1 cells produce IFN-γ, LT-α,

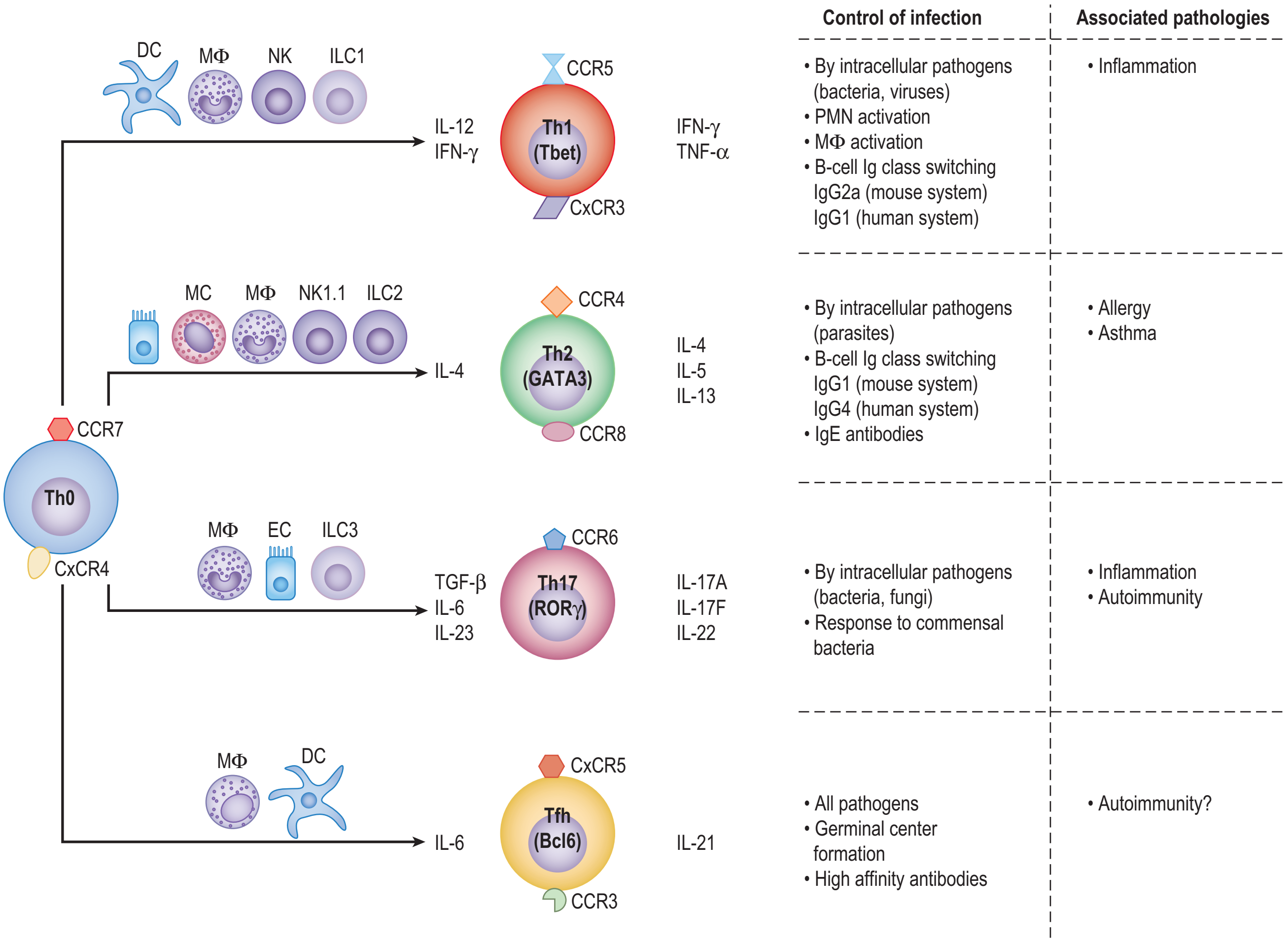

Control of infection
Associated pathologies
DC
MΦ
NK
ILC1
CCR5
IL-12
IFN-γ
Th1
(Tbet)
CxCR3
IFN-γ
TNF-α
• By intracellular pathogens (bacteria, viruses)
• PMN activation
• MΦ activation
• B-cell Ig class switching IgG2a (mouse system) IgG1 (human system)
• Inflammation
MC
MΦ
NK1.1
ILC2
CCR4
IL-4
Th2
(GATA3)
CCR8
IL-4
IL-5
IL-13
• By intracellular pathogens (parasites)
• B-cell Ig class switching IgG1 (mouse system) IgG4 (human system)
• IgE antibodies
• Allergy
• Asthma
CCR7
Th0
CxCR4
MΦ
EC
ILC3
CCR6
TGF-β
IL-6
IL-23
Th17
(RORγ)
IL-17A
IL-17F
IL-22
• By intracellular pathogens (bacteria, fungi)
• Response to commensal bacteria
• Inflammation
• Autoimmunity
MΦ
DC
CxCR5
IL-6
Tfh
(Bcl6)
CCR3
IL-21
• All pathogens
• Germinal center formation
• High affinity antibodies
• Autoimmunity?

FIG 20.6 **T-Helper (Th) Cell Subset Development in Mucosal Tissues.** The cellular and cytokine environment induces Th0 cells to develop into Th1, Th2, or Th17 subsets that can be discriminated based upon their cytokine production. Antigen-presenting cells (APCs) produce IL-12 in response to microbial assault and, together with interferon (IFN)-γ produced by natural killer (NK) cells and group 1 innate lymphoid cells (ILC1), induce mature Th1 cells. Th1 cells express select chemokine receptors and, through IFN-γ synthesis, activate macrophages (MØs) and induce B cells to produce opsonizing antibodies. Other cells, such as NK1.1, mast cells, and ILC2, respond to parasite/antigen/allergen with IL-4 production. IL-4 induces Th0 to Th2 differentiation. Epithelial cells (ECs) also produce cytokines that facilitate Th2 cell differentiation. Th2 cells produce IL-4, -5, -6, -9, -10, and -13, which help regulate mucosal secretory immunoglobulin A (SIgA) antibody responses. Transforming growth factor (TGF)-β, IL-6, and IL-23 produced by epithelial cells, MØs, and other cells help promote the differentiation of Th17 cells. The cytokines produced by Th17 cells contribute to several functions for the host response to commensal bacteria and protection against fungal infections. Follicular T helper cells (Tfh) are a subset of Th cells that help germinal center formation and the development of high-affinity antibodies.

LT-β, and TNF-α, whereas Th2 cells produce IL-4, IL-5, IL-6, IL-9, IL-10, and IL-13 (Fig. 20.6). In mice, mucosal Th1-type responses are associated with cell-mediated immunity and B-cell responses with characteristic IgG2a antibodies. Th2 cells support the production of IgA, as well as IgG1, IgG2b, and IgE. In humans and mice, Th1 and Th2 cells regulate the development of the opposite subset reciprocally through IFN-γ and IL-4 secretion, respectively (see Fig. 20.6). Human Th1 cells and IFN-γ responses are associated with IgG1 and IgG3 C-fixing antibodies with low IgG2 and undetectable IgG4 antibody levels,[21] suggesting that in humans IL-4 promotes IgG4 and IFN-γ promotes IgG1 (see Fig. 20.6).

Tregs and Th17 cells play a role in mucosal homeostasis and inflammatory responses (Chapter 18). Human tonsil CD4 T cells expressing the B-cell follicle homing receptor CXCR5 are identified as Tfh cells that help B-cell differentiation.[22] Foxp3$^+$ Tregs in PPs can apparently differentiate into Tfh cells, which express the chemokine CXCR5, the transcription factor Bcl-6, and the cytokine IL-21, to promote germinal center formation and IgA synthesis in the gut.[23,24]

B-Cell Isotype Switching and IgA Plasma Cell Differentiation

Isotype switching is preceded by transcriptional activation of the isotype in question (Chapter 4). IL-4 and TGF-β induce surface IgM-positive (sIgM$^+$) B cells to switch to IgE and IgA. TGF-β_1 can induce sIgM$^+$ to sIgA$^+$ B-cell switches, and addition of TGF-β_1 to LPS-triggered mouse B-cell cultures increased IgA synthesis. In humans, anti-CD40 stimulation of tonsillar B cells, together with TGF-β_1 in the presence of IL-10, stimulates IgA synthesis.[1] Cα_1 transcripts can also be induced by B-cell mitogen plus TGF-β, and Cα_2 transcripts can be induced by TGF-β together with IL-10.

DCs can also induce surface IgA$^+$ B cells via direct stimulation of B cells with B-cell activation factor of the TNF family (BAFF) and a proliferation-inducing ligand (APRIL).[25] APRIL–transmembrane activator and CAML interactor (TACI) signaling plays a key role in CD40-independent IgA class switching in mice.[25] In humans, functional mutations in TACI can result in IgA deficiency (IgAD; Chapter 34). Differentiation of sIgA$^+$ B cells into IgA-producing plasma cells is dependent on IL-5 and IL-6.[26]

VACCINE DEVELOPMENT AND MUCOSAL IMMUNE RESPONSES

Mucosal sIgA antibodies, as well as Th cell and cytotoxic T lymphocyte (CTL) responses, can be induced by pathogens triggering the organized mucosal inductive sites. Effective protection against virulent mucosal pathogens requires prophylactic immune responses that can be achieved through mucosal vaccines. In contrast to conventional injected vaccines, those administered via mucosal routes can trigger both mucosal immune responses as a first line of defense at the portal of pathogen entry and systemic immune responses that neutralize pathogens that have penetrated that barrier. Thus safe adjuvants boosting SIgA antibodies and mucosal immunity are being developed for mucosal vaccines. These efforts are in large part as a result of knowledge gained from studies of bacterial enterotoxins and nontoxic derivatives (Fig. 20.7).

Lessons From Studies of Bacterial Enterotoxins

Earlier studies of cholera toxin (CT) and heat-labile toxin I (LT-I) from *Escherichia coli* helped establish that mucosal (*i.e.,* oral or nasal) administration of vaccines was an effective approach for the induction of both mucosal and systemic immunity to coadministered vaccine antigens (Chapter 90). These closely related molecules are AB-type toxins consisting of two structurally and functionally separate enzymatic A subunits and binding B subunits (see Fig. 20.7). The B subunit of cholera toxin (CT-B) binds to GM1 gangliosides, whereas the B subunit of heat-labile toxin I (LT-B) binds to GM1 as well as GM2 asialo-GM1 gangliosides. The A subunits of these toxins are adenosine diphosphate (ADP)–ribosyl transferases. Binding of the B subunits to ganglioside receptors on target cells allows the A subunits to reach the cytosol where they elevate cyclic adenosine monophosphate (cAMP) levels.

Cellular Targets of Vaccine Adjuvants Can Shape the Immune Response

Studies with CT and LT-I revealed the importance of the cellular targets for shaping the profile of immune responses induced by mucosal adjuvants. In fact, CT promotes CD4 Th2 and Th17 responses, whereas LT-I also induces a CD4 Th1 (*i.e.,* IFN-γ) response.[27] As discussed below, studies with mutants of these enterotoxins and other toxins (*e.g., Bacillus anthracis* edema toxin[28]) have shown that their enzymatic activity are dispensable for vaccination.

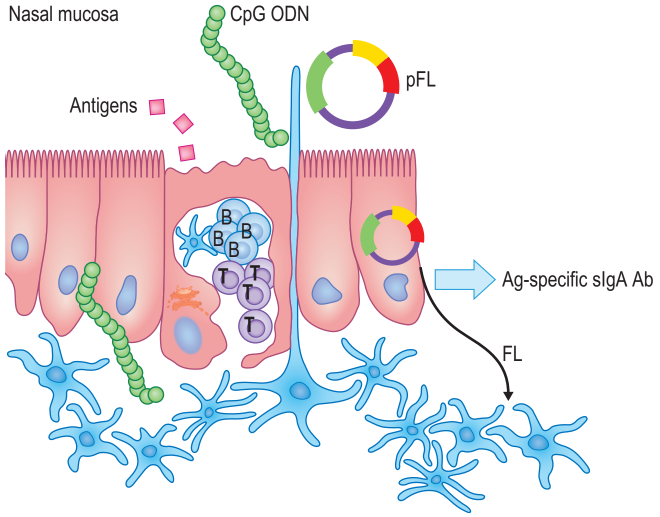


FIG 20.7 **Dendritic Cell (DC) Targeting With a Nasal Adjuvant.** Nasal administration of CpG oligodeoxynucleotide (ODN) and plasmid expressing FLT3 ligand cDNA (pFL) specifically target DCs in nasal-associated lymphoid tissues (NALTs). These nasal DC-targeting vaccines successfully elicit protective antigen-specific secretory immunoglobulin A (SIgA) antibody responses in older adults.

Central Nervous System Targeting Is a Safety Concern With Nasal Vaccines

Cholera induces diarrhea as a result of its ability to elevate cAMP in epithelial cells, thereby promoting secretion of water and chloride ions into the intestinal lumen. Diarrhea is thus the primary limiting factor for the use of oral enterotoxin as an adjuvant in humans. The olfactory neuroepithelium in the nasopharynx constitutes approximately 50% of the nasal surface and has direct neuronal connection to the olfactory bulbs (OBs) in the central nervous system (CNS). Nasally delivered enterotoxins can enter and/or target olfactory neurons and therefore gain access to OBs and deeper structures in the brain parenchyma. These adverse effects are, in large part, mediated by the ADP–ribosyl transferase activity and the nature of the cellular receptors targeted. Both CT and LT-I bind to GM1 on epithelial cells and require endocytosis followed by transport across the epithelial cell to reach the basolateral membrane. GM1 gangliosides are also abundantly expressed by the neuronal and microglial cells of the CNS.[29] CT or CT-B, when administered nasally to mice, enters the olfactory nerves and epithelium (ON/E) and OBs by mechanisms that are selectively dependent on GM1.[30] The targeting of CNS tissues by nasally administered bacterial enterotoxins is clearly related to a higher incidence of Bell palsy (facial paresis) among volunteers of a nasal vaccination trial given LT-I as mucosal adjuvant. Bell palsy among study subjects that in 2000 received nonliving nasal influenza vaccine (Nasalflu) led to its withdrawal from the market (www.niaid.nih.gov/dmid/enteric/intranasal.htm).

New Mucosal Adjuvants and Delivery Systems

CLINICAL RELEVANCE

Examples of Mucosal Adjuvants and Delivery Systems for the Induction of Targeted Immunity

Genetically Engineered Bacterial Toxins

- Examples are derivatives of the enterotoxin cholera toxin (CT) and heat-labile toxin (LT-I) from *Escherichia coli.*

Nucleic Acid Toll-Like Receptor (TLR) Ligands

- These sequences typically contain a transcription unit designed to express the antigen in question that is coupled to an adjuvant/mitogen unit, such as CpG motifs.

Mucosal Cytokines and Innate Factors as Adjuvants

- Mucosal delivery of specific cytokines or innate factors can reduce the risk of adverse systemic effects while targeting the immune response to the mucosa.

Transgenic Plants

- Plants, such as potatoes, bananas, and rice, can be engineered to express both B- and T-cell antigen epitopes, providing a simple delivery system for oral vaccination or oral tolerance induction.

Nontoxic Derivatives of Bacterial Enterotoxins

To circumvent the toxicity of enterotoxins, mutants of CT (mCT) and LT (mLT) molecules were generated by site-directed mutagenesis in the active site of the A subunit of CT or LT, or in the protease sensitive loop of LT. These mutants induced comparable levels of antigen-specific serum IgG and sIgA antibodies as wild-type CT and significantly higher levels than those induced by recombinant CT-B.[31] One of the mutants also induces Th2-type responses through a preferential inhibition of Th1-type CD4 T cells. mLT molecules, whether possessing a residual ADP–ribosyltransferase activity (*e.g.,* LT-72R) or totally devoid of it (*e.g.,* LT-7 K and LT-6 K3), can also function as mucosal adjuvants for nasal vaccine antigen in mice.[32] As LT induces a mixed CD4 Th1- and Th2-type response,[27] one might envisage the use of mLTs when both Th1- and Th2-type responses are desired.

The use of GM1-receptor binding holotoxins as nasal mucosal adjuvants is currently not recommended because of the risk for their accumulation in the CNS. However, nontoxic mCT could overcome these potential problems. To this end, a model adjuvant has been developed by combining the ADP-ribosylating ability of native CT (nCT) with a dimer of an Ig-binding fragment, D, of *Staphylococcus aureus* protein A.[33] This CTA1-DD molecule directly binds to B cells of all isotypes, but not to MØs or DCs. Despite the lack of a mucosal binding element, the B cell–targeted CTA1-DD molecule is as strong an adjuvant as nCT. Notably, CTA1-DD promoted a balanced Th1/Th2 response with little effect on IgE antibody production. CTA1-DD did not induce inflammatory changes in the nasal mucosa and, most importantly, did not bind to or accumulate in the OBs or the CNS.[33] CTA1-DD is an example of the use of nonganglioside targeting adjuvants and delivery systems as new tools for the development of safe and effective nasal vaccines.

Nucleic Acid Toll-Like Receptor Ligands

Toll-like receptor 3 (TLR3) and TLR9 recognize the pathogen-associated microbial pattern double-stranded RNA (dsRNA) and unmethylated DNA, respectively (Chapter 3). The latter contains immunostimulatory sequences consisting of short palindromic nucleotides located around a CpG dinucleotide core (*e.g.,* CpG motifs). CpG motifs bind to intracellular TLR9 and induce cytokine secretion (*i.e.,* IL-6, IFN-α, IFN-β, IFN-γ, IL-12, and IL-18) by a variety of immune cells. CpG motifs can enhance both systemic and mucosal immune responses when given nasally to mice,[34] and injection of bacterial DNA or CpG motifs with a DNA vaccine or with a protein antigen promotes Th1-type responses even in mice with preexisting Th2-type immunity.

Stimulation of TLR3 by dsRNA results in the production of type I IFNs (*i.e.,* IFN-α/β), which stimulate antibody responses to injected vaccines.[35] The synthetic TLR3 ligand polyinosinic–polycytidylic acid (poly I:C) has been shown to enhance CD8 responses to an experimental nasal influenza vaccine in mice and promote heterosubtypic protection via stimulation of TLR3 signaling by nonhematopoietic radioresistant cells.[36]

Mucosal Cytokines and Innate Factors as Adjuvants

Mucosal delivery of cytokines offers a means to prevent the adverse effects associated with the large and repeated parenteral doses often required for the effective targeting of tissues and organs. For example, nasal delivery permits acquisition of significant serum levels of IL-12 at one-tenth the dose required for inhibition of serum IFN-γ by parenteral administration.[1] Earlier studies have shown that nasal administration of tetanus toxoid with IL-12 as adjuvant induced high titers of sIgA antibody responses in the GI tract, vaginal washes, and saliva.[37] Similar results were reported when mice were nasally immunized with soluble influenza H1 and N1 proteins and IL-12. Related studies showed that mucosally administered IL-12 can redirect antigen-specific Th2-type responses toward the Th1 type or promote mixed Th1- and Th2-type responses, depending on the mucosal route and timing of delivery.[1]

FMS-like tyrosine kinase 3 ligand (FL) binds to the FMS-like tyrosine kinase receptor Flt3/Flk2. FL mobilizes and stimulates myeloid and lymphoid progenitor cells, DCs, and NK cells. Although FL dramatically augments numbers of DCs *in vivo,* it fails to induce their activation. Treatment of mice by systemic FL injection can induce marked increases in the numbers of DCs in both systemic *(i.e.,* spleen) and mucosal lymphoid tissues (*i.e.,* iLP, PPs, and mesenteric lymph nodes). Although this increase in mucosal DCs can, in some cases, initially enhance induction of oral tolerance,[38] it favors the induction of immune responses by mucosal or systemic vaccines. Nasal administration of plasmid or adenovirus encoding FL cDNA (pFL or Ad-FL) with protein antigens was shown to induce antigen-specific sIgA and protective immunity.[39-41] Thus FL cDNA may be an alternative to costly treatments with FL protein.

Transgenic Plants

Edible plants have been engineered to synthesize and assemble one or more antigens that retain both T- and B-cell epitopes, thereby inducing systemic and mucosal immune responses in both mice and humans.[42,43] To circumvent potential denaturation of the plant antigen during cooking, recombinant bananas that can accumulate up to 1 mg of vaccine antigen per 10 g of banana were developed. Most recently, the CT-B subunit has been expressed under the control of the rice seed storage protein glutelin promoter (MucoRice-CT-B). Oral feeding of powdered MucoRice-CT-B to mice and nonhuman primates resulted in the induction of both systemic and mucosal antibody responses for protection against CT.[44-46]

ON THE HORIZON

Development of Transgenic Plants as Vehicle for Vaccine Administration

- The MucoRice system is a novel strategy for vaccine development.
- The MucoRice system may also be used as a passive neutralizing antibody delivery system.

SYNTHESIS AND FUNCTIONS OF SECRETORY ANTIBODIES

Mucosal sIgA differs from serum IgA in both molecular composition and specific antibody activity. Humans possess two $C\alpha$ gene segments, $C\alpha 1$ and $C\alpha 2$ (Chapter 4), the use of which defines the two IgA subclasses, IgA1 and IgA2.[8] These IgA subtypes differ primarily in their hinge regions (Chapter 15). IgA1 antibodies contain an additional 13 amino acids in the hinge region, and this renders them more flexible and susceptible to IgA1-specific proteases produced by certain bacteria. IgA1-secreting cells are prevalent in most human mucosal tissues, especially the small intestine and the respiratory tract, whereas the human colon and genital tract are enriched by IgA2-secreting cells. SIgA are mostly viewed as a barrier at mucosal surfaces to prevent adhesion and colonization of pathogens, as well as an effective means to neutralize viruses and toxins, although these antibodies confer the additional advantage of providing antiinflammatory properties.[8]

In external secretions, adult levels of sIgA are reached considerably earlier (1 month to 2 years) than in the serum (adolescence). Approximately 98% of SIgA antibodies are produced locally in mucosal tissues, with only a minor fraction deriving from the circulation.

KEY CONCEPTS

Secretory Immunoglobulin A (SIgA)

- Unlike serum IgA, mucosal secretion of IgA reaches adult levels early in life (1 month to 2 years after birth).
- The polymeric Ig receptor (pIgR) is expressed on the basolateral surface of epithelial cells and facilitates the active transport of secretory IgA, as well as pentameric IgM, into mucosal secretions.
- SIgA protects the host by inhibiting microbial adherence; neutralizing viruses, enzymes, and toxins; and engaging in antiinflammatory activities by means of inhibiting IgM and IgG complement activation.
- Clinically, selective IgA deficiency, which is the most common primary immune deficiency, is characterized by recurrent mucosal infections, including sinusitis, otitis media, bronchitis, and pneumonias of viral or bacterial origin, as well as acute diarrhea caused by viruses, bacteria, or parasites, such as *Giardia lamblia.*

Polymeric Immunoglobulin Receptor and pIgA Transport

The polymeric Ig receptor (pIgR) is synthesized as a transmembrane protein by epithelial cells and is found on the basolateral surface of epithelial cells. It acts as a receptor for the endocytosis of pIgA and pentameric IgM, both of which contain a J-chain. The pIgR is produced by bronchial epithelial cells, renal tubules, glands, and the epithelia of the small and large intestines.[1] The pIgR is not expressed by the FAE (including M cells) of PPs, but only by the adjacent columnar epithelial cells. Further, pIgR is expressed in the upper respiratory tract, which includes the nasal cavity, tonsils, trachea, bronchi, and tracheobronchial glands. Expression in the lower lungs is restricted to the pulmonary alveolar cells.

In female reproductive tissues, the expression of pIgR is influenced by the sex hormones. It is low in the vagina, absent in the ovary and myometrium, and very high in the fallopian tubes and uterus. Normal kidneys do not express pIgR, whereas epithelial cells in the lower urinary tract may normally express pIgR and transport pIgA into urine. The expression of pIgR can be upregulated by cytokines, such as IFN-γ, TNF-α, IL-1α, IL-1β, and TGF-β.

IgA-Mediated Inhibition of Microbial Adherence

The inhibition of microbial adherence plays a critical initial role in the protection of the host. This inhibition is mediated by both specific and nonspecific mechanisms. The surface of microorganisms interacting with sIgA becomes less hydrophobic and thus more likely to be entrapped in mucus. SIgA and pIgA are more effective at agglutinating microorganisms than is membrane-bound IgA, and the agglutinating ability of sIgA specific for capsular polysaccharides of *Haemophilus influenzae* appears to be crucial to preventing colonization by *H. influenzae.*[47]

Neutralization by sIgA of Viruses, Enzymes, and Toxins

SIgA antibodies have been shown to be effective at neutralizing viruses in several experimental systems (*e.g.,* influenza virus, Epstein-Barr virus [EBV], HIV, etc.) and at different steps in the infectious process. SIgA specific for influenza hemagglutinin can interfere with the initial binding of influenza virus to target cells or with the internalization and the intracellular replication of the virus. *In vitro* experiments employing polarized murine epithelial cells have demonstrated that antibodies specific to rotavirus and hepatitis virus can neutralize the viruses inside epithelial cells. Finally, sIgA can neutralize the catalytic activity of many enzymes of microbial origin.

Antiinflammatory Actions Mediated by SIgA Antibodies

IgA antibodies are unable to activate complement by either the classical or the alternative pathway (Chapter 21). Nevertheless, they can interfere with IgM- and IgG-mediated complement activation.[8] SIgA can inhibit phagocytosis, bactericidal activity, and chemotaxis by polymorphonuclear neutrophils (PMNs), monocytes, and MØs. IgA can downregulate the synthesis of TNF-α and IL-6, as well as enhancing the production of IL-1R antagonists by LPS-activated human monocytes. Thus the antiinflammatory properties of IgA are of significant importance for the integrity of the mucosa in that IgA can limit bystander tissue damage that may result from the continuous interactions of the mucosa with myriad dietary and environmental antigens. Systemically, circulating IgA also appears to help limit inflammatory reactions that result from complement fixation and phagocyte activation, and it contributes to the inhibition of IgE-dependent anaphylactic responses.

IgA Deficiency

Selective IgAD is the most common primary immune deficiency (PID) in individuals of European descent (Chapter 34). The clinical diagnosis of IgAD depends on the relative absence of IgA in the serum. However, the most important manifestations of the disorder primarily reflect the absence of both sIgA1 and sIgA2 in the external secretions. Thus IgAD affects both the mucosal and systemic immune compartments, with only rare individuals exhibiting a superselective loss of either IgA1 or IgA2 alone.[48]

MUCOSAL CTLS

M cells have specific receptors for mucosal virus that allow certain viruses, such as reoviruses, to enter the cells in both NALTs and GALTs. It is likely that enteric viruses, such as rotavirus, and respiratory pathogens, such as influenza virus and respiratory syncytial virus (RSV), also enter the mucosal inductive pathway via M cells.[49] After enteric infection or immunization, antigen-stimulated CTLs are disseminated from PPs into mesenteric lymph nodes via the lymphatic drainage. Oral immunization with live virus can thus induce antigen-specific CTLs in both mucosal and systemic lymphoid tissues.

KEY CONCEPTS

Mucosal Cytotoxic T Cells (CTLs)

- After enteric infection or immunization, antigen-stimulated CTLs are disseminated from Peyer patches into mesenteric lymph nodes via the lymphatic drainage.
- Oral immunization with live virus can induce antigen-specific CTLs in both mucosal inductive and effector tissues for mucosal immune responses, as well as in systemic lymphoid tissues for serum immune responses.

Enteric Viruses and Mucosal CTLs

CD8 CTLs (Chapter 17) play a central role in rotavirus and reovirus immunity.[50,51] Reovirus-induced CTL precursors (pCTLs) in GALTs migrate to the systemic compartment. Reovirus-specific CD8 CTLs associated with the $\alpha\beta$ T cell population are also observed in intraepithelial T lymphocytes. Oral delivery of rotavirus increases pCTLs in GALTs and results in their dissemination throughout the murine lymphoid system within 3 weeks. Moreover, adoptively transferred CD8 T cells mediate the clearance of rotavirus infection in severe combined immunodeficiency mice.

Respiratory Viruses and Mucosal CTLs

Studies of immune responses after intranasal infection with influenza virus in CD4-coreceptor knock-outs or other mice in which this subset had been depleted have shown that CD4 T cells do not affect the induction of pCTLs or significantly alter clearance of infection.[52] Clearance of influenza is unaltered by the use of β_2 microglobulin knock-out mice, which lack CD8 T cells, or of mice that have been treated with monoclonal anti-CD8. $\gamma\delta$ T cells with several Vδ chain specificities increase in the infected site as clearance occurs, which suggests a regulatory role for $\gamma\delta$ T cells in antiviral immunity.[53]

Mucosal AIDS Models for CTL Responses

Approximately 80% of new HIV-1 infections result from sexual transmission (Chapter 39). Studies using the rhesus macaque and the simian immunodeficiency virus (SIV) vaginal infection model have provided evidence that pCTLs occur in female macaque reproductive tissues and that infection with SIV induces CTL responses. This important finding was extended to vaginal infection with an SIV/HIV-1 chimeric virus (SHIV) containing the HIV-1 89.6 *env* gene.[54] Other work has shown that intranasal immunization with SIV/HIV components induces antibody responses in vaginal secretions.

Other Mucosal CTL Systems

Salmonella can elicit CD8 T-cell responses, including CTLs, to expressed proteins, and CD8 T cells induced to the parasite *Toxoplasma gondii* have been shown to be protective. Thus mucosal CD8 CTLs can also be induced in nonviral situations. Significant questions remain as to the mechanism by which naïve CD8 T cells can be triggered to expand into pCTLs and to the rules for expression of effector CTLs and memory in the actual mucosal compartment that manifests the infection. pCTLs accumulate in immunologically privileged sites, but they do not develop a cytotoxic function until they encounter infected class I MHC-presenting target cells. It is possible that this mechanism protects the common mucosal immune system network from inadvertent cytotoxic inflammatory events.

MUCOSAL IMMUNE RESPONSES IN EARLY LIFE AND AGING

Although most of its structures are present at birth, the mucosal immune system requires further postnatal development and maturation before becoming fully functional (Chapter 38). GALTs, NALTs, and tonsils are present in humans at birth. The bacterial colonization after birth increases the number of immune cells and germinal centers in these sites and the number of secondary lymph nodes (*e.g.,* mesenteric lymph nodes and cervical lymph nodes; Chapter 2) and generate innate lymphoid follicles (ILFs). Significant changes that occur after bacterial colonization include an increase in SIgA levels and in the numbers of IgA secreting cells, Tregs, and Th17 cells. BALT only develops after birth. It is clear that the immature mucosal immune system in early life cannot protect against infectious pathogens entering mucosal surfaces. This gap is filled by maternal antibodies, which are acquired either before birth through the placenta or after birth via ingestion of milk.

KEY CONCEPTS

Mucosal Immunosenescence

- Early mucosal aging is evident in the gastrointestinal (GI) tract immune system.
- Nasal immunization is an effective route for the induction of mucosal and systemic immune responses in aging mice.
- Dendritic cell (DC)–targeting mucosal adjuvants are able to elicit protective pathogen-specific secretory immunoglobulin A (SIgA) antibody responses in aged mice.

Immune functions are known to deteriorate as a result of aging in several species (Chapter 38). The risk and severity of infections are higher, and the susceptibility to certain types of autoimmune diseases and cancer are greater in older adults,[55] and responses to vaccination are diminished. Aging-associated alterations of the systemic immune compartments have been studied extensively. Dysfunctions occur in both B and T cells, although the latter are considered more susceptible to immunosenescence.

In humans, older subjects were reported to have significantly higher concentrations of salivary sIgA antibodies compared with younger subjects, whereas whole gut lavages of aged and young subjects contain similar amounts of antibodies.[56] Analogous results have also been obtained for total IgA antibody responses in the serum of aged animals and humans. These results indicate an absence of aging-associated impairment in total IgA antibody levels in external secretions.

The GI tract in older adults is particularly susceptible to infectious diseases. Antigen-specific mucosal IgA antibody responses are diminished in aged animals, especially those in GALTs.[56] In older humans, pathogens that invade through mucosal surfaces, such as influenza virus and the bacterial pathogen *Streptococcus pneumoniae,* cause more severe and more frequently lethal infections. The development of effective vaccines for older adults remains a largely unmet goal. To provide effective protection against influenza and *S. pneumoniae* for this population, one should strongly consider developing a new generation of vaccines that could induce pathogen-specific immunity in the respiratory tract. Although it has been shown that effective protection can be provided by pathogen-specific systemic IgG without mucosal IgA responses, pathogen-specific sIgA responses are a necessary component for providing a first line of effective immunity against these respiratory pathogens at their entry site.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Kiyono H, Kunisawa J, McGhee JR, et al. The Mucosal Immune System. In: Paul WE, editor. Fundamental immunology. 5th ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2008. p. 983–1030.
2. Baker K, Blumberg RS, Kaetzel CS. Immunoglobulin transport and immunoglobulin receptors. In: Mestecky J, Strober W, Russell MW, et al, editors. Mucosal immunology. 4th ed. Cambridge, MA: Elsevier; 2015. p. 349–407.
3. Clevers HC, Bevins CL. Paneth cells: maestros of the small intestinal crypts. Annu Rev Physiol 2013;75:289–311.
4. Collonna M, Fuchs A, Cella M. Innate lymphoid cells in mucosal homeostasis, infections, autoimmune disorders, and tumors. In: Mestecky J, Strober W, Russell MW, et al, editors. Mucosal immunology. 4th ed. Cambridge, MA: Elsevier; 2015. p. 1003–12.
5. Serafini N, Di Santo JP. Effector cells of mucosal immune system: innate lymphoid cells. In: Mestecky J, Strober W, Russell MW, et al, editors. Mucosal immunology. 4th ed. Cambridge, MA: Elsevier; 2015. p. 787–804.
6. Williams IR, Owen RL. M cells: Specialized antigen sampling cells un the follicle-associated epithelium. In: Mestecky J, Strober W, Russell MW, et al, editors. Mucosal immunology. 4th ed. Cambridge, MA: Elsevier; 2015. p. 211–30.
7. Kiyono H, Fukuyama S. NALT- versus Peyer's-patch-mediated mucosal immunity. Nat Rev Immunol 2004;4:699–710.
8. Woof JM, Mestecky J. Mucosal immunoglobulins. In: Mestecky J, Strober W, Russell MW, et al, editors. Mucosal immunology. 4th ed. Cambridge, MA: Elsevier; 2015. p. 287–324.
9. Randall TD, Carragher DM, Rangel-Moreno J. Development of secondary lymphoid organs. Annu Rev Immunol 2008;26:627–50.
10. van de Pavert SA, Mebius RE. New insights into the development of lymphoid tissues. Nat Rev Immunol 2010;10:664–74.
11. Brandtzaeg P. Potential of nasopharynx-associated lymphoid tissue for vaccine responses in the airways. Am J Respir Crit Care Med 2011;183:1595–604.
12. Knop N, Knop E. Conjunctiva-associated lymphoid tissue in the human eye. Invest Ophthalmol Vis Sci 2000;41:1270–9.
13. Nagatake T, Fukuyama S, Kim DY, et al. Id2-, RORγt-, and LTβR-independent initiation of lymphoid organogenesis in ocular immunity. J Exp Med 2009;206:2351–64.
14. Habtezion A, Nguyen LP, Hadeiba H, et al. Leukocyte Trafficking to the Small Intestine and Colon. Gastroenterology 2016;150:340–54.
15. Czerkinsky C, Holmgren J. Mucosal delivery routes for optimal immunization: targeting immunity to the right tissues. Curr Top Microbiol Immunol 2012;354:1–18.
16. Brandtzaeg P. The mucosal B cell system. In: Mestecky J, Strober W, Russell MW, et al, editors. Mucosal immunology. Cambridge, MA: Elsevier; 2015. p. 623–81.
17. Csencsits KL, Jutila MA, Pascual DW. Nasal-associated lymphoid tissue: phenotypic and functional evidence for the primary role of peripheral node addressin in naïve lymphocyte adhesion to high endothelial venules in a mucosal site. J Immunol 1999;163:1382–9.
18. Wolber FM, Curtis JL, Milik AM, et al. Lymphocyte recruitment and the kinetics of adhesion receptor expression during the pulmonary immune response to particulate antigen. Am J Pathol 1997;151:1715–27.
19. Quiding-Jarbrink M, Nordstrom I, Granstrom G, et al. Differential expression of tissue-specific adhesion molecules on human circulating antibody-forming cells after systemic, enteric, and nasal immunizations. A molecular basis for the compartmentalization of effector B cell responses. J Clin Invest 1997;99:1281–6.
20. Lambrecht BN, Iwasaki A, Kelsall BL. Mucosal dendritic cells: Origins, Subsets, and Biology. In: Mestecky J, Strober W, Russell MW, et al, editors. Mucosal immunology. 4th ed. Cambridge, MA: Elsevier; 2015. p. 489–541.
21. Widhe M, Ekerfelt C, Forsberg P, et al. IgG subclasses in Lyme borreliosis: a study of specific IgG subclass distribution in an interferon-gamma-predominated disease. Scand J Immunol 1998;47:575–81.
22. Moser B. CXCR5, the defining marker for follicular B helper T (TFH) cells. Front Immunol 2015;6:296.
23. Tsuji M, Komatsu N, Kawamoto S, et al. Preferential generation of follicular B helper T cells from Foxp3+ T cells in gut Peyer's patches. Science 2009;323:1488–92.
24. Kato LM, Kawamoto S, Maruya M, et al. Gut TFH and IgA: key players for regulation of bacterial communities and immune homeostasis. Immunol Cell Biol 2014;92:49–56.
25. Cerutti A, Chen K, Chorny A. Immunoglobulin responses at the mucosal interface. Annu Rev Immunol 2011;29:273–93.
26. Fujihashi K, McGhee JR, Lue C, et al. Human appendix B cells naturally express receptors for and respond to interleukin 6 with selective IgA1 and IgA2 synthesis. J Clin Invest 1991;88:248–52.
27. Fujihashi K, McGhee JR. Th1/Th2/Th3 cells for regulation of mucosal immunity, tolerance and inflammation. In: Mestecky J, Lamm ME, Strober W, et al, editors. Mucosal immunology. 3rd ed. Part I, Chapter 28. Cambridge, MA: Elsevier Inc.; 2005. p. 539–58.
28. Duverger A, Carre JM, Jee J, et al. Contributions of edema factor and protective antigen to the induction of protective immunity by *Bacillus anthracis* edema toxin as an intranasal adjuvant. J Immunol 2010;185:5943–52.
29. Mancini P, Santi PA. Localization of the GM1 ganglioside in the vestibular system using cholera toxin. Hear Res 1993;64:151–65.
30. van Ginkel FW, Jackson RJ, Yuki Y, et al. Cutting edge: the mucosal adjuvant cholera toxin redirects vaccine proteins into olfactory tissues. J Immunol 2000;165:4778–82.
31. Yamamoto S, Kiyono H, Yamamoto M, et al. A nontoxic mutant of cholera toxin elicits Th2-type responses for enhanced mucosal immunity. Proc Natl Acad Sci USA 1997;94:5267–72.
32. Rappuoli R, Pizza M, Douce G, et al. Structure and mucosal adjuvanticity of cholera and *Escherichia coli* heat-labile enterotoxins. Immunol Today 1999;20:493–500.
33. Lycke N, Bemark M. Mucosal adjuvants and long-term memory development with special focus on CTA1-DD and other ADP-ribosylating toxins. Mucosal Immunol 2010;3:556–66.
34. Fukuyama Y, Ikeda Y, Ohori J, et al. A molecular mucosal adjuvant to enhance immunity against pneumococcal infection in the elderly. Immune Netw 2015;15:9–15.
35. Le Bon A, Schiavoni G, D'Agostino G, et al. Type I interferons potently enhance humoral immunity and can promote isotype switching by stimulating dendritic cells *in vivo*. Immunity 2001;14:461–70.
36. Perez-Giron JV, Belicha-Villanueva A, Hassan E, et al. Mucosal polyinosinic-polycytidylic acid improves protection elicited by replicating influenza vaccines via enhanced dendritic cell function and T cell immunity. J Immunol 2014;193:1324–32.
37. Boyaka PN, Marinaro M, Jackson RJ, et al. IL-12 is an effective adjuvant for induction of mucosal immunity. J Immunol 1999;162:122–8.
38. Williamson E, Westrich GM, Viney JL. Modulating dendritic cells to optimize mucosal immunization protocols. J Immunol 1999;163:3668–75.

39. Fukuyama Y, King JD, Kataoka K, et al. Secretory-IgA antibodies play an important role in the immunity to *Streptococcus pneumoniae*. J Immunol 2010;185:1755–62.
40. Kataoka K, McGhee JR, Kobayashi R, et al. Nasal Flt3 ligand cDNA elicits CD11c$^+$ CD8$^+$ dendritic cells for enhanced mucosal immunity. J Immunol 2004;172:3612–19.
41. Sekine S, Kataoka K, Fukuyama Y, et al. A novel adenovirus expressing Flt3 ligand enhances mucosal immunity by inducing mature nasopharyngeal-associated lymphoreticular tissue dendritic cell migration. J Immunol 2008;180:8126–34.
42. Haq TA, Mason HS, Clements JD, et al. Oral immunization with a recombinant bacterial antigen produced in transgenic plants. Science 1995;268:714–16.
43. Tacket CO, Mason HS, Losonsky G, et al. Immunogenicity in humans of a recombinant bacterial antigen delivered in a transgenic potato. Nat Med 1998;4:607–9.
44. Nochi T, Takagi H, Yuki Y, et al. Rice-based mucosal vaccine as a global strategy for cold-chain- and needle-free vaccination. Proc Natl Acad Sci USA 2007;104:10986–91.
45. Nochi T, Yuki Y, Katakai Y, et al. A rice-based oral cholera vaccine induces macaque-specific systemic neutralizing antibodies but does not influence pre-existing intestinal immunity. J Immunol 2009;183:6538–44.
46. Tokuhara D, Yuki Y, Nochi T, et al. Secretory IgA-mediated protection against *V. cholerae* and heat-labile enterotoxin-producing enterotoxigenic *Escherichia coli* by rice-based vaccine. Proc Natl Acad Sci USA 2010;107:8794–9.
47. Kauppi-Korkeila M, van Alphen L, Madore D, et al. Mechanism of antibody-mediated reduction of nasopharyngeal colonization by *Haemophilus influenzae* type b studied in an infant rat model. J Infect Dis 1996;174:1337–40.
48. Singh K, Chang C, Gershwin ME. IgA deficiency and autoimmunity. Autoimmun Rev 2014;13:163–77.
49. Williams IR, Owen RL. M cells: Specialized antigen sampling cells in the follicle-associated epithelium. In: Mestecky J, Strober W, Russell MW, et al, editors. Mucosal Immunology. 4th ed. Cambridge, MA: Elsevier; 2015. p. 211–30.
50. Burns JW, Siadat-Pajouh M, Krishnaney AA, et al. Protective effect of rotavirus VP6-specific IgA monoclonal antibodies that lack neutralizing activity. Science 1996;272:104–7.
51. Cuff CF, Cebra CK, Rubin DH, et al. Developmental relationship between cytotoxic α/β T cell receptor-positive intraepithelial lymphocytes and Peyer's patch lymphocytes. Eur J Immunol 1993;23:1333–9.
52. Allan W, Tabi Z, Cleary A, et al. Cellular events in the lymph node and lung of mice with influenza. Consequences of depleting CD4+ T cells. J Immunol 1990;144:3980–6.
53. Carding SR, Allan W, Kyes S, et al. Late dominance of the inflammatory process in murine influenza by γ/δ^+ T cells. J Exp Med 1990;172:1225–31.
54. Miller CJ, McChesney MB, Lu X, et al. Rhesus macaques previously infected with simian/human immunodeficiency virus are protected from vaginal challenge with pathogenic SIVmac239. J Virol 1997;71:1911–21.
55. Goronzy JJ, Weyand CM. Understanding immunosenescence to improve responses to vaccines. Nat Immunol 2013;14:428–36.
56. Fujihashi K, Kiyono H. Mucosal immunosenescence: new developments and vaccines to control infectious diseases. Trends Immunol 2009;30:334–43.

MULTIPLE-CHOICE QUESTIONS

1. The following are major characteristics of mucosal effector sites:
   - **A.** Essentially all of the T and B cells present are of a memory phenotype.
   - **B.** Most of the plasma cells are producing polymeric immunoglobulin A (pIgA).
   - **C.** The majority of T cells are $CD4^+$ and have characteristics either T-helper 1 (Th1)- or Th2-type helper cells.
   - **D.** Significant numbers of functional $CD8^+$ cytotoxic T lymphocytes (CTLs) are present.
   - **E.** All of the above.

2. Two IgA subclasses occur in humans, IgA1 and IgA2. However, their distribution in mucosal effector sites is remarkably different. For example:
   _____ predominates in tissues from the waist up, whereas _____ is prevalent in the lower gastrointestinal (GI) tract. In this regard, ______ predominates in the upper respiratory tract and nasal mucosa. On the other hand, ______ predominates in the lower ileum and colon.

   Correct Order:
   - **A.** IgA1→IgA2→IgA2→IgA1
   - **B.** IgA2 →IgA1→IgA1→IgA2
   - **C.** IgA1→IgA2→IgA1→IgA2
   - **D.** IgA2→IgA1→IgA2→IgA1

3. The existence of a "common mucosal immune" system suggests that mucosal tissues are connected. This means that:
   - **A.** Nerves connect distant mucosal sites.
   - **B.** Immune cells freely move between mucosal sites.
   - **C.** Selected homing receptors and addressins control homing to selected tissues.
   - **D.** Homing receptors and addressins play no role in homing to effectors sites.

21

The Human Complement System: Basic Concepts and Clinical Relevance

John P. Atkinson, Terry W. Du Clos, Carolyn Mold, Hrishikesh Kulkarni,
Dennis Hourcade, Xiaobo Wu

OVERVIEW: AN EVOLUTIONARY AND HISTORICAL PERSPECTIVE

The complement system arose early in evolution, possibly even in single-cell organisms. In vertebrates, it features a proteolytic cascade to coat microorganisms with cleavage fragments that are recognized by receptors on phagocytic cells.[1-4] This phenomenon, known as *opsonization,* leads to immune adherence followed by internalization. A related strategy employed by primates is to express a complement receptor on erythrocytes. In blood, the opsonized pathogen becomes adherent to the abundant erythrocytes and then, like a taxi, is taken by them to the liver and/or the spleen for transfer to monocytes and macrophages. In this manner, bacteria are both immobilized and prevented from traveling free in the circulation to such sites as the brain.

To mediate opsonization, the complement system needed to develop a system to transfer a plasma component onto the pathogen. Lectins and, subsequently in evolution, antibodies (Abs) faced this same challenge. Of note, both of the latter eventually utilized the complement system (lectin pathway and classical pathway [CP]) to "complement" their attachment strategies. In contrast to lectins and antibodies, the alternative pathway (AP) earlier solved this problem creatively through development of a thioester bond in C3. C3's homologous cousin, α_2-macroglobulin, also utilizes cleavage of a thioester bond to covalently attach to and thereby inactivate proteases. The parallels are apparent—upon its generation following C3 activation, C3b can transiently (within microseconds) attach to nearby hydroxyl or amino groups to form an ester or amide linkage, respectively. This mechanism creates essentially an almost unbreakable bond and places this complement fragment irreversibly on the pathogen's surface (Table 21.1).

To summarize, the major function of the complement system is to modulate the membrane of a microbial target, leading to immune adherence and internalization of the opsonized antigen. The target is selected by lectins in the lectin pathway and predominantly by antibodies in the CP. However, the ancient AP does not feature selective or specific recognition. Instead, 1–2% of C3 "ticks over" each hour, serving as a surveillance system. The activated C3b has a few microseconds to bind to a target, such as bacteria, or it will be inactivated by water. If C3b lands on a pathogen, C3b can rapidly amplify to put several million copies of C3b on a single *Escherichia coli* in 2–3 minutes. If C3b remains in the fluid phase, it is promptly inactivated by plasma regulators. If it binds to healthy self, it is inactivated by ubiquitously expressed membrane complement inhibitors. To summarize, the AP initially formed the original complement system and likely consisted of three elemental proteins (C3 and two proteases known as *factor B* and *factor D*). These three were sufficient to generate a C3 convertase (splitting) enzyme and also to form a feedback loop. The positive regulator of this remarkable enzymatic feedback/amplification loop, known as *properdin,* probably came later in evolution (as did lectins and antibodies).

Later to arise in evolution is the membrane attack complex (MAC, often called the *terminal pathway*) that is common to all three cascades. The goal of its five sequentially interacting, nonenzymatic proteins (C5b, C6, C7, C8, and C9) is also to attach to and then alter the surface of a pathogen by membrane perturbation, often ending in lysis. In humans, a deficiency of any one of these five proteins leads to meningococcal infections. Interestingly, properdin deficiency also predisposes to recurrent meningococcemia. Thus such infections likely drove specialization of this terminal wing of the complement system.

In contrast to the rapidly acting AP, a major limitation with the lectin and Ab systems is that the triggers of these two pathways are slow! It takes several days to ramp up synthesis of a particular lectin (acute phase type response) and at least a week to develop an appropriate immunoglobulin M (IgM) and even longer for an IgG immune response. For a host with an opportunistic pathogen invading the bloodstream, this time delay for an adaptive humoral immune response is far from adequate. Consequently, the complement system is often called the "guardian of the intravascular space." Once a "pumped" circulation developed in evolution, a rapidly acting, abundant, and nearly fail-safe system to prevent and stop "bugs" from entering, traveling, and dividing in plasma was mandatory.

The second function of the complement system is to promote the inflammatory response. This is primarily accomplished by the anaphylatoxins C3a and C5a. Upon cleavage of C3 to C3b (the major complement opsonin) and C5 to C5b (the trigger of the MAC), the ~10 kilodalton (kDa) C3a and C5a fragments are released, and these can engage their respective receptors to initiate vascular and cellular changes, rapidly leading to a proinflammatory state. These receptors are expressed on many cell types, including endothelial, epithelial, and immune cells. Upon receptor engagement, defensive strategies are initiated and result in increased blood flow and stimulated phagocytes that are now more efficient at binding and ingesting C3b-coated antigens.

Through these same interactions, the complement system instructs the adaptive immune response. Antigens decorated by complement proteins are taken up by monocytes, follicular-dendritic cells (FDCs), B lymphocytes, and other antigen-presenting cells (APCs), resulting in an adaptive immune response

TABLE 21.1 Proteins of the Complement System

Component	Function
Classical Pathway (CP)	
C1q	Part of C1. Binds to immunoglobulin M (IgM), IgG, pentraxins, and ligands on apoptotic cells to initiate CP activation.
C1r	Part of C1. After auto-activation, cleaves C1s.
C1s	Part of C1. After activation by C1r, cleaves C4 and C2.
C4	Cleaved by C1s to form C4b, part of the CP and LP C3 and C5 convertases. An opsonin.[*]
C2	Binds to C4b and then cleaved by C1s to form C2a; enzymatic component of the CP and LP C3 and C5 convertases. C2b is released.[†]Δ
Lectin Pathway (LP)	
MBL	Recognition component for LP activation. Binds to mannose-rich glycans through C-type lectin domains.
MASP-1 and MASP-3	Associated with MBL and ficolins. Cleaves C2, but not C4. Cleaves profactor D.
MASP-2	Associated with MBL and ficolins. Cleaves C2 and C4.
Ficolins 1–3	Recognition components for LP activation. Bind to glycans through fibrinogen-like recognition domains.
Alternative Pathway (AP)	
C3	Cleaved by C3 convertases to form C3b and C3a. C3b is opsonic. A small fraction becomes part of the AP C3 convertase and part of all C5 convertases. C3b is further cleaved to opsonic iC3b and the CR2 ligands C3dg and C3d. C3a is an anaphylatoxin.
Factor B	Binds to C3b and then cleaved by factor D to form Bb, the enzymatic component of the AP C3 and C5 convertases. Ba is released.[‡]○
Factor D	Cleaves factor B bound to C3b to form AP convertases.
Properdin	Stabilizes AP convertases. Binds to microbial ligands to initiate AP activation.
Membrane Attack Complex (MAC)	
C5	Cleaved by C5 convertases to form C5b and C5a. C5b initiates MAC formation. C5a is an anaphylatoxin.
C6	Part of the MAC. Binds membranes.
C7	Part of the MAC. Binds membranes.
C8	Part of the MAC. Initiates pore formation.
C9	Part of the MAC. Polymerizes to form lytic pores.
Soluble Regulatory Proteins	
C1-INH	Serine protease inhibitor of C1r, C1s, MASP-1, MASP-2, kallikrein, factor XII.
C4BP	Binds C4b. Decay accelerating and cofactor activities for C4b-containing convertases.
FH	Binds C3b and polyanions. Has decay-accelerating and cofactor activities for C3b-containing convertases.
FI	Cleaves C3b and C4b bound to a cofactor protein.
Vitronectin	Binds C5b-7, prevents membrane insertion and lysis.
Clusterin	Binds C8 and C9, prevents MAC assembly and lysis.
Membrane Regulatory Proteins	
CD55 (DAF)	Accelerates decay of C3 and C5 convertases.
CD46 (MCP)	Cofactor for FI cleavage of C3b and C4b.
CD59	Binds to C8 and C9, prevents MAC assembly and lysis.
Receptors	
CD35 (CR1)	Opsonic receptor for C3b and C4b. Has decay accelerating and cofactor activity for C4b and C3b and convertases containing these fragments.
CD21 (CR2)	Receptor for C3dg and C3d. Enhances B cell activation.
CD11b/CD18 (CR3)	Opsonic receptor for iC3b. Leukocyte adhesion integrin.
CD11c/CD18 (CR4)	Opsonic receptor for iC3b. Leukocyte adhesion integrin.
CRIg	Opsonic receptor for iC3b and C3c. Inhibits C5 convertases.
C5aR (CD88)	Proinflammatory and chemotactic receptor for C5a.
C5L2	Receptor for C5a. Function not fully defined.
C3aR	Proinflammatory and chemotactic receptor for C3a.

[*] C4a, an anaphylatoxin is released. Function and receptor are not defined.
[†] No defined function upon release.
[‡] Several putative functions reported but requires further study.

(the complement system is often called "nature's adjuvant.") Thus complement activation is required for optimal Ab responses to most foreign antigens. Individuals lacking C3 are predisposed at an early age to bacterial infections, predominantly by encapsulated bacteria.

Complement deficiencies are instructive anomalies of nature. Surprisingly, a complete deficiency in an early component of the classical complement pathway predisposes to autoimmune diseases, particularly systemic lupus erythematosus (SLE) (Chapter 51). Greater than 80% of patients with C1q or C4 deficiency present with SLE. This association indicates that the complement system is required not only for host defense against foreign agents but also to identify and safely clear self-materials (debris or garbage removal), particularly RNA and DNA species.

Much complement-mediated pathology revolves around disturbances of the AP and thus its potent amplification loop. It must be rigorously regulated to prevent activation on normal self and excessive activation on injured self. Even haploinsufficiency of its two major inhibitors predisposes to endothelial damage in atypical hemolytic–uremic syndrome and retinal damage in age-related macular degeneration (ARMD).

In clinical medicine, the complement system participates in three pathological processes: (1) bacterial infections and/or autoimmunity are caused by complete component deficiencies; (2) tissue damage results from complement activation by autoantibodies forming immune complexes (*e.g.*, SLE); and (3) excessive activation at sites of cellular and tissue injury occurs in individuals carrying dysfunctional genetic variants in complement regulators.

Knowledge of how complement is activated and how it can be controlled points to opportunities for the development of therapeutic agents. One such example is anti-C5 monoclonal antibody (mAb) therapy, which has been recently approved to treat several complement-dependent hemolytic disorders. Other new complement therapeutics and diagnostics are on the horizon as biotechnology companies pursue novel complement inhibitors and genetic evaluations are increasingly utilized for helping to define diseases in which the complement system is involved.

COMPLEMENT PATHWAYS

The three pathways of complement activation are the CP, AP, and lectin cascades (Fig. 21.1).[1-4] The CP is initiated by IgM or IgG Ab binding to antigen. The CP can also be activated by innate pattern recognition molecules, such as the pentraxins, cross-reactive protein (CRP) and serum amyloid P (SAP) component,[5] and the membrane-bound lectin, SIGN-R1.[6] It participates with natural IgM in early host defense and provides a mechanism for immune complex and apoptotic cell clearance. The lectin pathway (LP) uses most of the CP components but is activated by mannan-binding lectin (MBL) and the ficolins, which are lectins that recognize repeating carbohydrate patterns on microorganisms.[7] The AP is the most ancient pathway and also has the broadest recognition ability. The AP is engaged by surface components of all types of microorganisms, including bacteria, fungi, parasites, viruses, and virus-infected cells.[8] It is continuously turning over and autoactivates if its inhibitors are lacking. Activation of the AP can also be initiated by properdin (P), a molecule that binds to pathogens and apoptotic cells. This mechanism further promotes its function as an innate and rapid responder to infection.[9] The AP is an important amplification mechanism for CP or LP activation, resulting in greater opsonization and generation of the terminal lytic pathway. For example, the initial trigger on a pathogen surface may be IgM or a lectin, but the majority of the C3b deposited is via AP's amplification or feedback loop that is engaged by just having C3b on a target.

The cleavage of C3 to C3a and C3b is central to all three pathways of complement activation. This enzymatic step exposes a highly reactive thioester bond through which C3b covalently attaches to nearby molecules (Fig. 21.2). Activation of C3 to C3b also exposes sites for interactions with other complement proteins, inhibitors, and receptors. Recent results have shed new light on the structural basis for C3 activation. In 2006, several studies reported the first x-ray structure of C3b, the activated product C3.[10,11] The results revealed a major conformational change in C3 upon cleavage to C3b that exposes the reactive thioester group as well as cryptic binding sites for complement receptors and regulatory proteins. Moreover, the crystal structures of five binding proteins in complex with C3b have now been solved.[12]

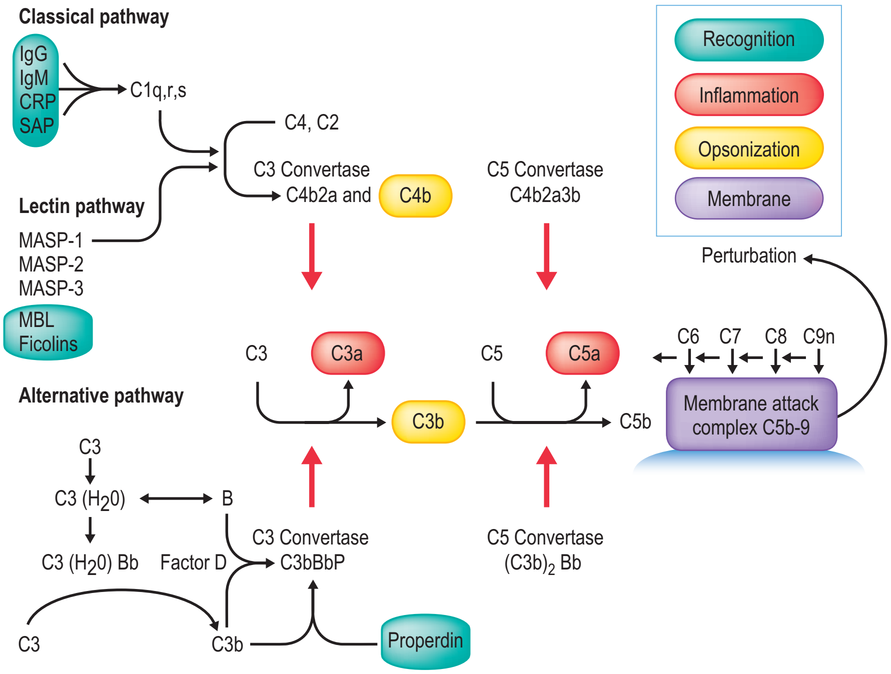


FIG 21.1 Overview of the complement pathways indicating components required for recognition, the enzymatically active fragments and complexes and the major opsonic, inflammatory, and membranolytic products.

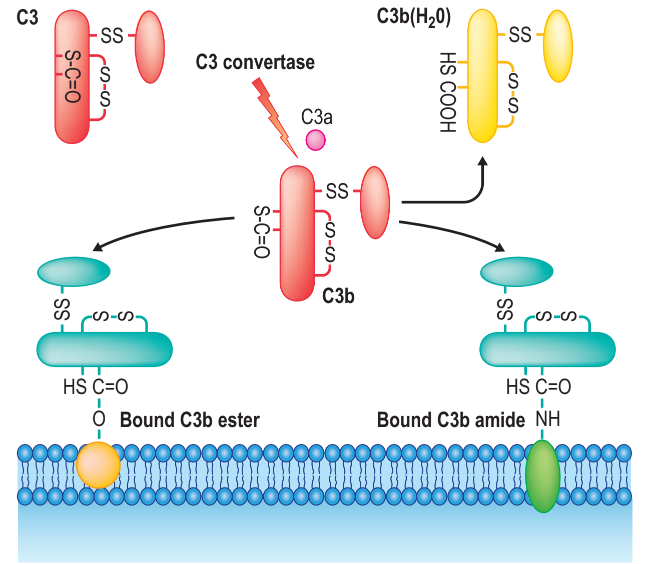


FIG 21.2 Exposure and Reactivity of the C3 Thioester Bond. C3 cleavage by a C3 convertase generates metastable C3b with a reactive thioester. Metastable C3b may be hydrolyzed to form $C3b(H_2O)$ or may react with hydroxyl (ester linkage) or amino (amide linkage) groups to become covalently bound to a surface.

Classical Pathway

KEY CONCEPTS

Structural and Functional Homologies in Complement Pathways

Recognition: C1q, MBL, Ficolins, CRP
Initiating enzymes: C1r, C1s, MASP-1, MASP-2, FD
C3 convertases: C4b2a, C3bBb
C5 convertases: C4b2a3b, $(C3b)_2Bb$
Enzyme subunits of convertases: C2a, Bb
Assembly subunits: C3b, C4b (covalently bound to target)
Anaphylatoxins: C3a, C5a
MAC subunits: C5b, C6, C7, C8, C9
Regulatory proteins: C4BP, FH, CR1, MCP, DAF
Receptor proteins: CR1 and CR2; CR3 and CR4
Major opsonins: C3b and C4b

The CP is focused by Ab binding to a target antigen. In general, the ability of Ab to activate complement is IgM > IgG3 > IgG1 > IgG2 > IgG4. Binding of these Ab exposes sites in the Fc region for attachment of the first subcomponent of complement, C1q.[13,14] C1 is a large calcium-dependent complex composed of C1q and two molecules each of the proenzymes, C1r and C1s. C1q is a 410-kDa protein with six globular heads connected by a collagen-like tail. For IgM, which is pentameric, binding to antigen creates a conformational change that exposes the C1q binding site in the Cμ3 domain. For IgG, at least two closely bound molecules are required to provide multiple attachment points for C1q binding to the Cγ2 domain. Similarly, CCRP or SAP molecules bound to ligand provide multiple C1q binding sites, resulting in CP activation. IgM, IgG, and CRP bind C1q through its globular head groups. Membrane lipids exposed on apoptotic cells or mitochondrial membranes, polyanions, nucleic acids, retroviruses, and endotoxins can also activate the CP.

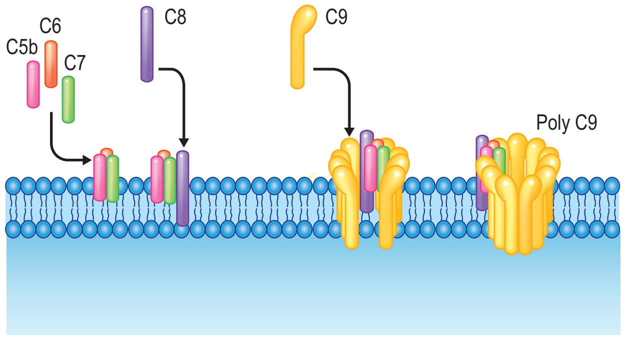


FIG 21.3 Sequence of Protein Interactions in the Assembly of the Membrane Attack Complex (MAC). C5b, generated by a convertase cleaving C5, combines with C6 and C7 to form a hydrophobic complex capable of a membrane interaction. Binding of C8 allows the complex to insert further into the membrane and forms a site for C9 polymerization. C9 polymers (10–15) form a transmembrane pore to mediate cell lysis.

Once C1q binds to an activator through several globular heads, C1r is cleaved by an autocatalytic process. Activated C1r then cleaves and activates C1s, which, in turn, cleaves circulating C4. C4 and C3 are highly homologous proteins that share an unusual posttranslational modification known as an *internal thioester bond* (see Fig. 21.2).[10,15] Cleavage of C4 releases the C4a fragment and exposes the reactive thioester bond in the larger C4b fragment. This allows C4b to attach covalently to nearby target structures, through either amide or ester bonds, to form amino or carboxyl groups on proteins and polysaccharides on cell surfaces, including Abs and a wide array of antigens. The exposed thioester bond is highly but transiently reactive as it is susceptible to rapid hydrolysis. For example, only about 5% of the C4b typically becomes attached to the target. Bound C4b provides an anchor site for C2 attachment, which is then also cleaved by C1s, releasing the smaller fragment C2b.

The complex of C4b2a is termed the *CP C3 convertase* because it has the capacity to cleave C3, releasing C3a. The C2a component of the complex contains the active enzymatic site. C3 cleavage is similar to C4 cleavage in that the larger fragment, C3b, contains a thioester site (see Fig. 21.2) that mediates covalent attachment to nearby surface structures, including the antigen, the Ab, and the attached C4b. C3 is found at a three- to fourfold higher concentration in serum compared with C4, and its cleavage is amplified by the AP. Thus efficient complement activation will result in clusters of multiple bound C3b molecules that can be recognized by cellular receptors. C3b that attaches to C4b within the C3 convertase produces the trimolecular complex C4b2a3b, which is a C5 convertase. Cleavage of C5 produces C5a, which has potent inflammatory activity, and C5b, which initiates the formation of the MAC or, as it is also known, the *terminal complement complex* (TCC) (Fig. 21.3).

Lectin Pathway

The LP is similar to the CP, except that it uses pattern recognition molecules, MBL, and ficolins-1, 2, and 3, instead of Ab to target activation.[7,16] MBL is structurally similar to C1q, with a collagen-like region and globular heads. The globular heads of MBL are C-type lectin domains specific for repeating carbohydrate structures found on microorganisms. Like C1q, MBL and ficolins are in complex with serine proteases, MBL-associated serum proteases (MASPs), which are structurally and functionally similar to C1r and C1s. MASP-1 and -2 are active proteases, but only MASP-2 cleaves both C4 and C2 to generate C4b2a, the same C3 convertase as the CP. MASP-1 can supplement activation by cleaving C2 but not C4. Two nonproteolytic splice products of the *MASP2* and *MASP1/3* genes, sMAP and MAP-1, compete with MASP-1 and MASP-2 for binding to MBL to regulate the LP. Subsequent steps in the LP are identical to those in the CP. Although still controversial, MASP-1 and MASP-3 likely contribute to the AP by cleaving profactor D to active factor D.[7,16,17]

Alternative Pathway

The AP uses proteins that are structurally and functionally homologous to those of the CP, but this pathway has unique features that play three important roles in the complement cascade. The surveillance role of the AP is mediated by a continuous low level of spontaneous activation that results from the hydrolysis of the C3 thioester bond.[8] Hydrolyzed C3, C3(H_2O), assumes a conformation similar to that of C3b and can bind factor B (homologous to C2), which is cleaved by factor D (homologous to C1s) to form a fluid-phase C3 convertase. This convertase cleaves C3 to generate C3b, which can covalently bind to nearby structures and provide the basis for a bound C3 convertase (C3bBb). Because C3b is both a part of this enzyme and a product of the reaction, a positive feedback loop that rapidly deposits more C3b is formed. This low-grade activation process is tightly regulated on host cells and tissues by plasma and membrane-bound complement regulatory proteins. It is the lack of such regulation that usually restricts AP activation to microbial targets. The plasma protein factor H (FH) is particularly important in controlling AP activation, both in the fluid phase and on "nonactivating" surfaces. The latter recruits FH through its binding sites for polyanions, including sialic acid and glycosaminoglycans. "Activating" surfaces, such as microbial polysaccharides, lipopolysaccharides, and foreign glycoproteins, provide C3b attachment sites that are protected from regulatory proteins. Similar to the CP, the AP C5 convertase (C3bBb3b) is formed when a second C3b attaches to the C3 convertase. The AP C3 and C5 convertases are stabilized by P (factor P or P), for which this pathway was originally named.

More recently, an additional role for properdin (P) in initiating AP activation was rediscovered.[8,18,19] P is a pattern-recognition molecule with specificity for microbes and damaged cells. Once bound, P can recruit fluid-phase C3b or C3b (H_2O), independent of covalent binding, and thereby provide a platform for the assembly of the AP convertase. Thus P binding can direct AP activation, similar to MBL in the LP. P binding to certain *Neisseria* species potently activates the AP, and this may account for the susceptibility of P-deficient individuals to infection with *N. meningitidis.*

The third important role of the AP is the amplification of C3b deposition and C5 convertase generation that is initiated by the CP or the LP.[8] This function of the AP is critical in complement-mediated pathology, as it increases the generation of C5a and the MAC, the most inflammatory components of the system. It is this amplification role of the AP that makes it an attractive therapeutic target.

MEMBRANE ATTACK COMPLEX

All three complement pathways merge with the cleavage of C5 into C5a and C5b. Although C5 is structurally homologous to C3 and C4, it lacks an internal thioester bond that allows covalent attachment to surfaces. C3a and C5a are also structurally homologous and, as described below, are the most potent proinflammatory mediators of the complement system. C5b initiates the formation of the MAC (see Fig. 21.3), a complex of C5b, C6, C7, C8, and multiple (5–10) C9 molecules.[20] This complex, as indicated by its name, penetrates membrane bilayers to form pores that disrupt the osmotic barrier, leading to swelling and lysis of susceptible cells. Lysis of Ab-sensitized erythrocytes by the MAC is the basis of the total hemolytic complement (THC) assay or CH_{50}. C5b initiates the formation of the MAC without further proteolytic steps. C5b binds to C6, and this complex binds to C7. The C5b67 complex is lipophilic and associates with cell membranes, if available, or with serum lipoproteins. Once bound to a membrane, C5b67 recruits C8, and the complex penetrates more deeply into the membrane. However, efficient lysis requires C9, a pore-forming molecule with homology to perforin, a protein used by cytotoxic T cells and natural killer (NK) cells for killing virus-infected targets. The complex of C5b678 forms a nidus for C9 binding and polymerization. Although complement-dependent lysis of bacteria can be observed *in vitro*, many pathogens have evolved mechanisms to circumvent this activity of complement.[21] Opsonization by C3b is the most potent mechanism for destruction (adherence followed by ingestion) of bacteria by the complement system. The sublytic MAC is proinflammatory because of its membrane perturbing capabilities for host cells and contributes to the deleterious effects of complement activation in inflammatory diseases.[22]

REGULATION OF COMPLEMENT ACTIVATION

The complement cascade is rapidly activated and highly amplified by the generation of C3 and C5 convertases. There are three main levels of control that limit the potential harm that uncontrolled complement activation might cause: (1) the initiation step in the CP and the LP; (2) the C3 and C5 convertases of all three pathways; and (3) the assembly of the MAC. Both soluble and membrane-bound regulatory proteins serve these functions, which help terminate complement activation and direct it to appropriate targets.[23]

C1 Esterase Inhibitor

C1 esterase inhibitor (C1-INH) is a plasma serine proteinase inhibitor (serpin). C1-INH covalently binds to activated C1r and C1s, irreversibly inhibiting their activity and thereby limiting CP activation. C1-INH inactivation of C1r and C1s also removes them from the C1 complex, exposing sites on the collagen-like region of C1q. Likewise, C1-INH inhibits MASP-1 and MASP-2, kallikrein, factor XIa, factor XIIa, and plasmin of the LP and the contact, coagulation, and fibrinolytic systems. Inherited deficiency of C1-INH is the basis of hereditary angioedema, a disease characterized by recurrent attacks of subcutaneous or submucosal edema (Chapter 42).[24]

Regulators of the C3 and C5 Convertases

The C3 and C5 convertases are central to the generation of the inflammatory and opsonic products of complement activation and are highly regulated by fluid-phase and membrane-bound regulatory proteins. The membrane deposited C4b and C3b may be bound by the regulator to prevent an association with C2 or FB. The convertases themselves are complexes of two or three components, and one mechanism of regulation is the dissociation of these complexes. This type of regulation is termed *decay*

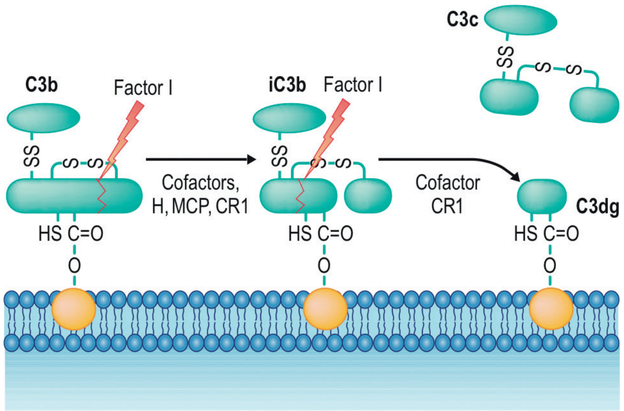


FIG 21.4 FI dependent cleavage of C3 showing the structures of the products and the required cofactors. The cofactor protein binds first and then the serine protease Factor I cleaves the C3b.

acceleration. A second mechanism of regulation is the enzymatic inactivation of the C4b and C3b components of the convertases (Fig. 21.4). This is accomplished by the plasma enzyme factor I (FI), which, however, only acts on C4b or C3b in complex with one of several regulatory proteins. The binding of regulatory proteins to C4b or C3b to enable FI cleavage is termed *cofactor activity.*

Factor I

FI (C3b inactivator, C3bINA) cleaves C4b and C3b into products that are recognized by specific cellular receptors (as discussed below). The sequential cleavages of C3b by FI to iC3b and C3dg are depicted in Fig. 21.4. C4b is cleaved in an analogous manner to C4d. (The iC4b intermediate is found only transiently.) The regulatory proteins that facilitate this cleavage by cofactor activity and those that inactivate C3 and C5 convertases by decay-accelerating activity are members of a family of structurally related proteins encoded within the regulators of complement activation (RCA) genetic locus.[25] This family is characterized by a repeating structure that consists of subunits, termed *complement control protein repeats* (CCP), of about 60 amino acids with a conserved pattern of two disulfide bonds per repeat and are usually encoded by a single exon.

Soluble Regulatory Proteins, C4b-Binding Protein, and FH

C4b-binding protein (C4bp) and FH are fluid-phase regulatory proteins with both decay-accelerating and cofactor activities. C4bp is multimeric, being composed of seven identical subunits, each containing eight CCPs. FH is a single-chain protein composed entirely of 20 CCPs. C4bp is specific for C4b and the C4b-containing convertases of the CP (C4b2b, C4b2b3b), whereas FH regulates C3b and C3b-containing convertases (C3bBb, C3bBb3b, C4b2b3b). FH is essential for regulation of C3 "tickover," and FH deficiency results in an acquired deficiency of C3. Additional binding sites on FH that recognize polyanions, such as sialic acid and glycosaminoglycans, provide targeted regulation of AP activation on surfaces.[23,26-28]

Membrane Regulatory Proteins

The RCA family includes the membrane regulatory proteins decay-accelerating factor (CD55, DAF), membrane cofactor protein (CD46, MCP), and complement receptors CR1 (CD35) and CR2 (CD21).[25,29] CD55 (DAF) and CD46 (MCP), as their names imply, have decay-accelerating and/or cofactor activity, respectively, that inhibits complement activation on cell membranes.[25] Each has an extracellular domain composed exclusively of four CCPs. CD55, a glycophosphatidylinositol (GPI)–anchored protein, and CD46, a transmembrane protein, are widely distributed on cells in contact with blood, with the notable exception of erythrocytes that lack CD46. Soluble CD55 is also found in most biological fluids. Both protect cells from complement-mediated lysis. CD35 (CR1) has decay-accelerating and cofactor activity and is a receptor for bound C3b. The function of CD35 as a complement receptor is discussed later in the chapter.

Complement C2 receptor inhibitor trispanning (CRIT) is a non-RCA membrane regulator of the CP. CRIT was originally identified on *Schistosoma* and *Trypanosoma* parasites and later found to be widely expressed on human tissues and blood cells, except for neutrophils and erythrocytes.[30] CRIT competes with C2 for binding to C4b, blocking the formation of the CP C3 convertase.

Properdin

In contrast to the regulatory proteins discussed above, the plasma protein P (factor P) stabilizes C3 and C5 convertases of the AP, increasing their activity.[8,18,19] This enhancer of AP activation is found as noncovalently linked dimers, trimers, tetramers, and larger species composed of identical 56-kDa chains. The majority of this plasma protein consists of a series of six thrombospondin type 1 modules. P binds to C3b and to Bb, preventing the spontaneous or induced decay of the AP C3 and C5 convertases. Its multimeric structure promotes interaction with clustered C3b. As discussed above, bound P can also recruit C3b to provide a site of assembly for the AP C3 convertase.

Regulators of the Membrane Attack Complex

The MAC is also regulated by both fluid-phase and membrane regulatory proteins.[1,20,23,29]

Soluble MAC Inhibitors: Vitronectin and Clusterin

Soluble hydrophobic proteins block the incorporation of the MAC into membranes. Two well-characterized proteins with this activity are vitronectin (S protein) and clusterin (SP-40,40, apolipoprotein J).[23,29] Vitronectin is in plasma and the extracellular matrix and binds to C5b-7. C8 and C9 can still bind to the complex, but membrane insertion and C9 polymerization are prevented. Soluble complexes of vitronectin and C5b-9 are in plasma during complement activation, and an enzyme-linked immunosorbent assay (ELISA) specific for this complex has been used to monitor activation of the MAC. Clusterin forms a complex with C5b-9, preventing membrane insertion. It is found in plasma, in the male reproductive tract, and on endothelial cells of normal arteries. It is also associated with amyloid deposits, including β amyloid in Alzheimer disease.

Membrane MAC Inhibitor CD59

The primary membrane-bound inhibitor of the MAC is CD59.[23,29] CD59 is a GPI-anchored protein expressed by most cells. CD59 binds to C8 and C9, preventing the incorporation and polymerization of C9.

COMPLEMENT RECEPTORS

Many of the biological effects of complement activation are mediated by cellular receptors for fragments of complement

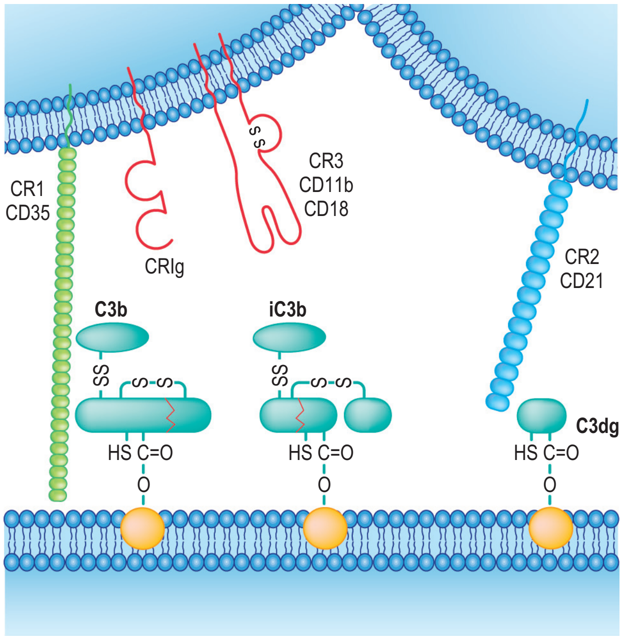


FIG 21.5 **Receptors for Bound C3b and Its Cleavage Products.** Receptors shown are CD35 and CD21 composed of CCP (SCR) subunits; CD11b/CD18 (CR3), a β_2 integrin; and CRIg with one or two immunoglobulin domains. The specificities of the receptors are CD35 for C4b and C3b, (C4b > C3b), CRIg for iC3b > C3b; CD11b/CD18 for iC3b; CD21 for C3dg and C3d. CD11c/CD18 (CR4) is similar to CD11b/CD18 and is not shown. Receptors are not drawn to scale. Their molecular weights are listed in Table 21.1.

proteins. These include receptors for the small soluble complement fragments, C5a and C3a, and receptors for bound complement fragments, C1q and C4b and C3b and their cleaved fragments. Receptors are specific for C3b and for its further breakdown products generated by the enzymatic processing by FI in conjunction with the cofactor proteins mentioned above. The breakdown of C3b and intermediate products are shown in Fig. 21.4 and the receptors for these components in Fig. 21.5.

C1q Receptors

C1q is one of a family of proteins termed *soluble defense collagens*, which includes the "collectins" (MBL, surfactant proteins A and D, conglutinin), and the ficolins. Each of these proteins is composed of a collagen-like linear stem region terminated by multiple globular recognition domains or head groups. The collectins recognize carbohydrates with their C-type lectin head groups, and the ficolins recognize acetyl groups on carbohydrates and other molecules with fibrinogen-like recognition domains. In contrast, the globular head groups of C1q do not recognize carbohydrates but, rather, bind to amino acid motifs on IgG, IgM, and pentraxins. In general, the soluble defense collagens broadly recognize pathogen-associated carbohydrate patterns and damaged or apoptotic cells. Reported direct effects of this group on leukocytes include the enhancement of phagocytosis, triggering of the respiratory burst, and regulation of cytokine responses. Several cell surface proteins have been proposed to facilitate these activities, including CD93 (C1qRp), CD35 (CR1), $\alpha_2\beta_1$ integrin, calreticulin in complex with CD91, gC1q binding protein, and SIPRα. However, none has been definitively established as a receptor in the classic sense.[1,5,16,31]

Complement Receptor 1 (CR1, CD35)

There are five identified receptors for bound fragments of C3 and/or C4. CD35 is a large protein composed of a linear string of CCPs, a transmembrane region, and a short intracytoplasmic domain.[32] Different allelic forms of CD35 are found, the most common being composed of 30 CCPs with a molecular weight of 190 kDa. These CCPs are organized into groups of seven, creating structures termed *long homologous repeats* (LHRs), each of which contains a single binding site. The predominant allele of CD35 contains two binding sites for C3b, three for C4b, and one for C1q. CR1 is expressed on human erythrocytes, monocytes and macrophages, neutrophils, B lymphocytes, a small percentage of T lymphocytes, eosinophils, FDCs, and glomerular podocytes.

CD35 on primate erythrocytes provides a mechanism for clearing soluble immune complexes from the circulation. Although the number of receptors on each erythrocyte is low, the large number of erythrocytes provides the major pool of CR1 in the circulation. Soluble immune complexes that fix complement attach quickly to erythrocytes in the circulation, bypassing monocytes and neutrophils. These erythrocyte-bound complexes are taken to the liver, where they are transferred to Kupffer cells expressing Fc and complement receptors and destroyed. The erythrocytes exit into the circulation to pick up more immune complexes. This clearance pathway is impaired in patients with SLE because of decreased complement in the circulation, decreased CD35 on erythrocytes, and saturated Fc receptors in the liver and spleen.

CD35 on monocytes and neutrophils promotes binding of microbes carrying C3b and C4b on their surface (immune adherence reaction), facilitating their phagocytosis through Fc receptors. CD35 can directly mediate the uptake of microbes when phagocytic cells have been activated by chemokines or integrin interactions with matrix proteins. CD35 is a member of the RCA family and has decay-accelerating and cofactor activity in addition to its function as a receptor. It differs from the membrane regulatory proteins DAF (CD55) and MCP (CD46) in its ability to also bind to C3b and C4b extrinsically (on targets other than the cell expressing it) and in its cofactor activity for iC3b processing. CD35 is the most effective cofactor for FI cleavage of C3b and iC3b to the smallest covalently bound fragment C3dg. C3dg is the major ligand for CR2 on B lymphocytes (described below). The cofactor activity of CD35 on B lymphocytes can process bound C3b to C3dg, facilitating binding to CR2 and lowering the threshold for B-cell activation.[29,33,34]

Complement Receptor 2 (CR2, CD21)

CD21 is also an RCA family protein composed of 15–16 CCPs. CD21 has a limited range of expression that includes B lymphocytes, FDCs, and some epithelial cells. CD21 is specific for the smallest covalently bound C3 fragments, C3dg and C3d, and has weaker binding to iC3b. CD21 is also the Epstein-Barr virus (EBV) receptor on B cells and nasopharyngeal epithelial cells and binds to CD23, a low-affinity IgE receptor.[33,34]

CD21 on B lymphocytes serves a costimulatory role. It is expressed on mature B cells as a complex with CD19 and CD81 (TAPA-1). Coligation of CD21 and the B-cell antigen receptor induces the phosphorylation of CD19, activating several signaling pathways and strongly amplifying B-cell responses to antigen. This role of CD21 is believed to contribute to the strong adjuvant effect produced by attaching C3d to antigen.[33,34]

Complement Receptors 3 and 4

CR3 and CR4 are the β_2 integrins commonly known as *CD11b/CD18 (Mac-1)* and *CD11c/CD18.*[34,35] β_2 integrins are large heterodimers found on neutrophils and monocytes with multiple roles in adhesion to endothelium and matrix molecules as well as direct recognition of microbial pathogens. The binding activities of β_2 integrins are regulated by cellular activation often through chemokine receptors. Both CD11b/CD18 and CD11c/CD18 are expressed primarily on neutrophils, monocytes, and NK cells and bind to iC3b and, to a lesser extent, C3b. CD11b/CD18 has been studied more extensively than CD11c/CD18. CD11b/CD18 expression, clustering, and conformation are all rapidly upregulated by chemokine activation of neutrophils, leading to increased responses to ligand. CR3 plays an essential role in neutrophil attachment to and migration through activated endothelium to sites in inflammation and in the regulation of neutrophil apoptosis. Deficiency of the β_2 chain (CD18) results in leukocyte adhesion deficiency, characterized by recurrent pyogenic infections, and defects in inflammatory and phagocytic responses. Complement receptors CD11b/CD18 and CD11c/CD18 provide an essential function for the removal of microbial pathogens following complement activation, since C3b processing to iC3b often occurs rapidly after deposition.

Complement Receptor of the Immunoglobulin Superfamily (CRIg)

CRIg is a receptor for iC3b and C3b present on Kupffer cells in the liver as well as other tissue macrophages but is absent from splenic macrophages, peripheral blood cells, bone marrow–derived macrophages, and monocyte/macrophage cell lines.[35,36] Two alternative-spliced forms of human CRIg were identified with one and two Ig domains. The mouse receptor has a single Ig domain. CRIg removes C3b or iC3b-opsonized particles from the circulation by the liver.

C5a and C3a Receptors

During complement activation, the homologous proteins C3 and C5 are each cleaved near the amino-terminus of the α chains to release a soluble peptide fragment of approximately 8 kDa. These fragments are designated C3a and C5a. C5a may also be generated locally by direct cleavage of C5 by thrombin or leukocyte proteases.[37] C3a and C5a are termed *anaphylatoxins* because of their ability to increase vascular permeability, contract smooth muscle, and trigger the release of vasoactive amines from mast cells and basophils.[38-41] C5a is 10- to 100-fold more active than C3a. These peptides are also chemotactic: C5a is specific for neutrophils, monocytes, and macrophages, whereas C3a is specific for mast cells and eosinophils. Other biological activities of complement anaphylatoxins are summarized in Table 21.2.

Structurally, anaphylatoxins are compact structures consisting of multiple helices cross-linked by disulfide bonds with more flexible carboxy-terminal regions. The C-terminal peptide of C3a interacts with the C3aR and can reproduce C3a agonist activity. In contrast, C5a interacts with the C5aR at multiple sites. Plasma carboxypeptidases cleave the C-terminal arginine from C3a and C5a producing the des-Arg forms. This inactivates C3a; however, $C5a_{desarg}$ retains much of its biological activity. The C5aR (CD88 and C5L2) and the C3aR are rhodopsin-type receptors with seven transmembrane-spanning domains coupled to G-protein signaling pathways. C5aR is expressed at high levels on neutrophils and is also found by macrophages, mast cells, basophils, smooth muscle cells, and endothelial cells. If C5a is generated locally, for example, in an extravascular site of infection, it helps induce an acute local inflammatory response, including vasodilation, edema, neutrophil chemotaxis, and activation of neutrophils and macrophages for enhanced phagocytosis and killing. The inflammatory activities of C5a can also contribute to complement-mediated pathology in some conditions, such as sepsis, acute respiratory distress syndrome, and ischemia/reperfusion (I/R) injury, making the C5a–C5aR interaction an attractive therapeutic target.

TABLE 21.2 Cellular Targets and Effects of Complement Anaphylatoxins

	Targets Bearing Receptors	Effects
C3a, C5a	Mast cells, basophils	Degranulation, release of vasoactive amines: contraction of smooth muscle, increased vascular permeability
C3a	Eosinophils	Chemotaxis, degranulation
C5a	Endothelium	Increased adhesion of leukocytes; augmented chemokinesis and cytokine synthesis
C5a	Neutrophils, monocytes/macrophages, eosinophils, basophils, astrocytes	Chemotaxis
C5a	Neutrophils, monocytes/macrophages	Priming: activation of receptors, assembly of nicotinamide adenine dinucleotide phosphate (NADPH) oxidase; activation: degranulation, respiratory burst
C5a	Resident macrophages	Regulation of FcγR expression (↑activating, ↓inhibitory)
C5a	Hepatocytes	Acute phase protein synthesis
C3a, C5a	Lymphocytes (antigen-presenting cells)	Regulation of T-cell responses to antigen

The C5L2 receptor binds to both C5a and $C5a_{desarg}$. C5L2 was initially believed to be a default or decoy receptor for C5a because it is uncoupled from G proteins. Genetic deletion of C5L2 ($Gpr77^{-/-}$) in mice resulted in enhanced neutrophil infiltration and cytokine production in the pulmonary Arthus reaction, supporting an antiinflammatory role for C5L2 in immune complex disease, where genetic deletion of C5aR is fully protective.[40] However, studies in a cecal ligation and puncture (CLP) model of sepsis found increased survival in mice lacking either C5aR or C5L2.[41] The results suggest an active proinflammatory role for C5L2 that requires C5a and results in the release of the inflammatory signal, high-mobility group box-1 protein (HMGB1) from phagocytic cells. Thus both C5aR and C5L2 may contribute synergistically to harmful inflammatory events during sepsis.

COMPLEMENT IN HOST DEFENSE AND IMMUNITY

Complement in Host Defense

Complement activation provides a coordinated response to infection that results in the opsonization of microbial pathogens and the attraction and activation of phagocytic cells to kill them.

Complement-dependent opsonization is of greatest importance in infections with encapsulated extracellular bacteria, and individuals with deficiencies in Ab production, neutrophil function, or C3 share increased susceptibility to these organisms, including *Streptococcus pneumoniae* and *Haemophilus influenzae.* MBL deficiency is also associated with recurrent pyogenic infections in young children. In general, activation of complement by natural Ab or MBL results in C3b and iC3b deposition on these pathogens, overcoming the antiphagocytic effects of the capsule. Phagocytic cells ingest and kill the organisms using CD35, CD11b/CD18, and CD11c/CD18 receptors in conjunction with other innate and Fc receptors. C5aR signaling activates these receptors, leading to increased phagocytosis. Gram-negative bacteria are susceptible to complement-dependent lysis. This is evident in the increased incidence of disseminated neisserial infection in individuals deficient in C3, any of the MAC components, or P, as discussed below.

Complement in Inflammation

An essential function of complement in host defense is the coordination of the local inflammatory response. C5a is the most potent complement product in this activity.[39] Sublytic deposition of the MAC on endothelial cells and platelets and C3a interaction with the C3aR also contribute to the proinflammatory effects of complement activation. As discussed below, these potent inflammatory fragments of complement, when generated in high amounts or targeted inappropriately, result in many of the disease-related deleterious effects of complement. Local production of C5a at a site of infection occurs either through local complement activation or through direct cleavage of C5 by tissue macrophages or thrombin.[37,38] This C5a is released and sets up a chemotactic gradient for neutrophils and macrophages. In addition, C5a activates endothelial cells to express P-selectin and synthesize chemokines, including interleukin-8 (IL-8). Interaction of C5a with mast cells releases vasoactive amines, increasing endothelial permeability. Neutrophils and macrophages are "primed" by interaction of C5a with its receptor. Priming includes enhancement of chemotaxis, activation of complement receptors for phagocytosis increased expression of activating FcγR, and assembly of the nicotinamide adenine dinucleotide phosphate (NADPH)–oxidase that is required for effective killing of microbes after phagocytosis. C5a also prevents neutrophil apoptosis, prolonging survival and contributing to local accumulation. Together, these actions result in the attraction and activation of potent antimicrobial cells and resolution of infection (Fig. 21.6).

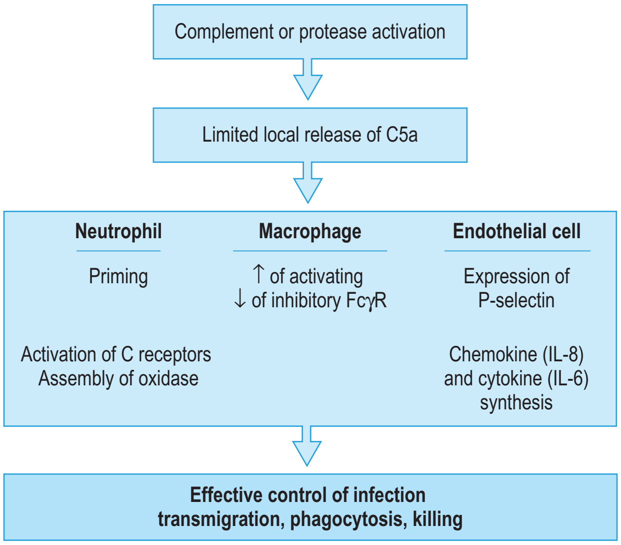


FIG 21.6 C5a in Local Host Defense. Note that C5a can also be released by a direct protease event, such as by thrombin. IL-8, Interleukin 8.

Pathogen Evasion of Complement

Further evidence of the host defense function of complement is the association of complement evasion strategies with virulence. Pathogenic gram-negative bacteria, such as *Salmonella,* have lipopolysaccharides with long O-polysaccharide side chains that promote rapid shedding of the MAC and prevent its insertion into the cell membrane. *Neisseria* species have several FH-binding components that help restrict AP activation and protect against lysis. Group A and B streptococci and *S. pneumoniae* have cell surface components (M protein, Bac or beta, PspC, Hic) that bind to FH and/or C4bp, restricting complement activation. Other organisms, including type 3 group B streptococci, elaborate sialic acid–containing capsules or cell walls to limit AP activation.

Although complement deficiencies are not generally associated with viral infections, the importance of complement in host defense against viruses is suggested by the multiple strategies used by viruses to evade complement.[21,42] Several viruses produce complement regulatory proteins, including vaccinia virus complement control protein and herpes virus glycoprotein C, which facilitate breakdown of C3b and C4b. Some viruses, such as human immunodeficiency virus (HIV), incorporate complement regulatory proteins into the viral envelope, a strategy that is also used by other pathogens, such as *Schistosoma.*[21,42,43]

There are also many examples of complement receptors and membrane regulatory proteins being exploited as receptors for pathogens to invade cells. Examples of these include strategies of direct pathogen binding to receptors as well as deposition of C3 fragments followed by invasion through host C3 receptors.[21]

Role of Complement in Adaptive Immunity

Over the past 10 years there has been renewed interest in the role of the innate immune system in adaptive immune responses.[1,8] The importance of complement in humoral immunity has been recognized since the observation that complement depletion of mice before immunization decreased Ab responses to thymus-dependent antigens. Further studies have shown that complement receptors CR1 (CD35) and CR2 (CD21) are also required.[33,34] In humans, these receptors are found together on B cells and FDCs. CD35 is also expressed on a number of other cell types (described above), including erythrocytes and phagocytic cells.

Effects of Complement on the Humoral Immune Response

Results obtained by experimental manipulation of C3, C4, and their receptors in mouse models indicate roles for these complement components at multiple levels in the humoral immune response.[33,34] One caveat regarding these studies is that in the mouse CD35 and CD21 are alternative splice products of the same gene, and genetically deficient animals lack both receptors.[34,44] In humans, CD35 and CD21 are encoded by separate genes. The first role of CD35/CD21 is in B-cell development, indicated by a pronounced defect in B-1-cell development in CD35/CD21–deficient mice. B-1 cells are generally found outside lymphoid follicles, have a restricted repertoire, and are essential in the production of natural Ab to pathogens, such as *S.*

KEY CONCEPTS

Abuse by and Evasion Strategies of Complement by Pathogens: Some Examples

1. Bacteria

Block C1, C3b deposition
    Streptococcus pneumoniae
Block MAC access to plasma membrane
    Salmonella
Limit access of C3b, iC3b to C receptors by capsule
    Streptococcus pneumoniae
    Haemophilus influenzae
Block AP activation by sialylation
    Streptococcus agalactiae (GBS) type III,
    Neisseria
Bind FH, C4BP to inhibit complement activation
    Streptococcus pneumoniae (Hic)
    Streptococcus pyogenes (GAS)(M protein)
    Neisseria
Use CD55 (DAF), CD46 (MCP) for attachment to cells
    Streptococcus pyogenes (GAS)(M protein)
    Neisseria
    Escherichia coli
Use complement receptors for entry
    Mycobacterium tuberculosis (CR3)
    Bacillus anthracis spores (CR3)

2. Viruses

Express complement regulatory proteins homologous to those synthesized by the host
    Herpes simplex virus (glycoprotein C)
    Poxviruses (SPICE/VICE)
Express unique complement regulatory proteins
    Flaviviruses (Dengue, West Nile)
Use CD55 (DAF), CD46 (MCP) for attachment to cells
    Measles virus, adenovirus, herpes virus 6 (CD46)
    Picornaviruses, hantavirus (CD55)
Use complement receptors for entry
    Epstein-Barr virus (CD21)
    Human immunodeficiency virus (CD35, CR3)

3. Parasites

Express complement regulatory proteins
    Schistosoma (CRIT)
Acquire complement regulatory proteins from host
    Schistosoma (CD55)
Use complement receptors for entry
    Leishmania (CR1, CR3)

pneumoniae, and to self antigens exposed on damaged cells, such as phosphatidylcholine and DNA. Although the mechanism of this defect in CD35/CD21–deficient mice is not fully understood, these mice have an altered repertoire of natural Ab and B-1 cells.[33,34,44] Decreased natural Ab may contribute to susceptibility to infection and autoimmune disease in hereditary complement deficiency (discussed below).

A second role for complement in the Ab response is the well-described function of CD21 as a coreceptor for the mature B-cell response to antigen.[33,34,44,45] As described above, CD21 is associated with the signaling complex of CD19 and CD81 (TAPA-1) in the B-cell membrane. Coligation of CD21 with the B-cell antigen receptor occurs naturally when the antigen activates complement and covalently binds C3dg. This coligation of the B-cell receptor with CD21 greatly decreases the threshold for B-cell activation and blocks Fas-initiated apoptosis of B cells. B cells activated by complement-opsonized antigen have increased ability to present antigen as well as survival and proliferation during encounters with T-dependent antigens.

The expression of CD35 and CD21 on FDCs is also important in the Ab response. FDCs trap antigen in the germinal centers and provide selection of somatically mutated high-affinity B-cell clones. Antigen trapped on FDCs also provides a source of long-term stimulation for maintenance of memory B cells. FDCs use complement receptors (CD35 and CD21) and FcγR to trap and retain antigen for these functions. Expression of CD21 on both FDC and B cells is required for effective affinity maturation of the Ab response and for the development and maintenance of memory B cells.

Complement and T-Cell Activation

Studies in primary pulmonary infection with influenza indicate that C3-deficient mice have a defect in influenza-specific CD4 and CD8 T-cell priming.[45] CR1/2 deficiency had no effect. The mechanism is unknown but may be more efficient uptake and presentation of C3-opsonized virus by APC through CR3 and CR4 or stimulation of T-cell responses through the C3aR.

Costimulation of human T cells *in vitro* through CD3 and CD46 leads to the development of T cells with a regulatory phenotype characterized by synthesis of IL-10 in the absence of other Th2 cytokines (IL-2, IL-4) (Chapter 9).[46] The induction of regulatory T cells (Tregs) was seen in response to both anti-CD46 cross-linking and natural ligands (C3b dimers, streptococcal M protein).

CD55-deficient mice showed enhanced T-cell responses to immunization and increased T cell–dependent autoimmune disease. These effects were complement dependent and apparently involve the loss of CD55 regulation of local complement synthesis by APCs during cognate interactions with T cells. One postulated mechanism is that CD55 inhibits the generation of C3a and C5a by APCs, preventing their interactions with C3aR and C5aR on T cells.[47] Complement anaphylatoxins, C3a and C5a, have many important effects in inflammatory diseases that include attraction and activation of inflammatory cells, as well as regulation of APC and T-cell responses. Examples of these will be discussed in the sections below.

Role of Complement in Clearance of Apoptotic Cells

Damaged tissue and dead and dying cells activate complement through several pathways. This can increase local inflammation and injury, as in I/R injury and hemolytic–uremic syndrome (HUS) (discussed below). Complement activation by apoptotic cells contributes to their opsonization and clearance and may prevent the development of autoimmunity. The deleterious consequences of complement activation following tissue damage are mainly attributable to AP-dependent generation of C5a and the MAC, whereas the beneficial effects are dependent on early CP components and innate recognition molecules.[48,49]

Necrosis, as occurs following ischemic tissue injury, exposes phospholipids and mitochondrial proteins that activate complement directly or indirectly. The pathways are different depending on the tissue involved.[48] For example, renal reperfusion injury appears to be initiated by the AP, possibly secondary to the loss of regulatory proteins on tubular epithelial cells. Intestinal (I/R) injury is initiated by natural IgM Ab and requires both the CP for initiation and the AP for injury. MBL and CRP-initiated complement activation have been proposed to contribute to myocardial reperfusion injury after coronary artery ligation.

Apoptotic cells are recognized by multiple receptors and opsonins.[49,50] The association between early CP deficiencies and

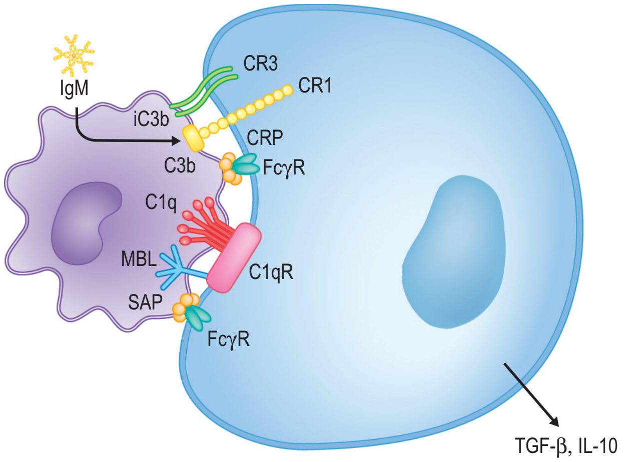


FIG 21.7 Pathways of Opsonization of Apoptotic Cells by Complement. Innate recognition of apoptotic cells by natural immunoglobulin M (IgM), cross-reactive protein (CRP), serum amyloid P (SAP), C1q, and mannose-binding lectin (MBL) is shown. Each reaction activates complement leading to opsonization by C3b and iC3b. In addition, C1q and MBL bind to collectin receptors, and CRP and SAP bind to FcγR on macrophages. Cytokine responses to apoptotic cells opsonized by complement include the antiinflammatory cytokines, transforming growth factor-β (TGF-β), and interleukin-10 (IL-10).

SLE (see below and Chapter 51) has been attributed to a failure of complement-dependent opsonization, resulting in accumulation of apoptotic cells and released autoantigens. Support for this hypothesis is provided by studies of mice deficient in C1q, IgM, or SAP, all of which develop autoantibodies against phospholipid and nuclear antigens characteristic of SLE, and by the therapeutic effect of CRP in mouse models of SLE.[5] The role of complement in apoptotic cell recognition and uptake by macrophages is depicted in Fig. 21.7. MBL, C1q, and surfactant protein-D (SP-D) bind to apoptotic cells and facilitate clearance through direct binding to cellular receptors as well as complement activation.[50] Natural IgM Ab, CRP, and SAP bind to phospholipids exposed on late apoptotic cells. All three proteins can also activate the CP generating C1q, C4b, C3b, and iC3b ligands for complement receptors. CRP and SAP also directly opsonize apoptotic cells for uptake through Fcγ receptors.[51] Phagocytosis of apoptotic cells generally induces antiinflammatory cytokines transforming growth factor-β (TGF-β) and IL-10.[52,53]

Targeted Activation of Complement for Opsonization

Interestingly, CRP and SAP also bind complement regulatory proteins, FH and C4bp, which helps limit complement activation to the deposition of opsonic components with little or no lysis or generation of C5a.[53,54] This type of complement activation was also observed on acrosome-activated spermatozoa. In this case, the CP was activated by CRP from follicular fluid, resulting in bound C3b and iC3b, which are proposed to bind complement receptors on the egg and facilitate fertilization. Riley-Vargas et al. have proposed the acronym TRACS (targeted and restricted activation of the complement system) for this type of limited complement activation that occurs as part of normal processes, such as the acrosome reaction, and the recognition and removal of ischemic tissue and apoptotic cells.[55]

COMPLEMENT DEFICIENCIES

Genetics and Incidence

Complete genetic deficiencies of complement proteins are rare, with an estimated combined prevalence of 0.03% for any inherited complete deficiency (excluding MBL deficiency) in the general population.[2-4,56-58] For most components, inheritance is autosomal and expression is codominant, so complete deficiency is homozygous recessive and heterozygotes express half levels. There are two C4 genes (*C4A* and *C4B*), so a range of partial deficiencies can be observed. All cases of C1-INH deficiency have been heterozygous, and P deficiency is X-linked. MBL is found in multiple allelic forms with different levels of expression ranging from 5 nanograms per milliliter (ng/mL) to more than 5 micrograms per milliliter (μg/mL) in plasma. Deficiencies specific to the LP are not detected by the screening assays described below but can be determined by specific assays.[7] A 10% incidence of MBL deficiency and a single case of MASP-2 deficiency have been described.

The most common clinical presentations of patients with complement deficiencies are recurrent infections with encapsulated bacteria, recurrent neisserial infections, and systemic autoimmune disease (Table 21.3). Populations with these disease manifestations have a much higher incidence of complement deficiency. For example, in Caucasian patients with SLE, the incidence of C2 deficiency is nearly 1%, 100-fold higher than in the general population. Screening of patients with autoimmune disease for complement deficiencies is useful, as these individuals are at higher risk for certain disease manifestations and may be at greater risk for infectious complications. Complement deficiency is found in as many as 20% of patients with recurrent disseminated neisserial infections. Evaluation of complement function is highly recommended in patients with recurrent or disseminated neisserial infections so that appropriate immunization and antibiotic prophylaxis can be initiated.

Complement deficiencies are most readily detected by hemolytic screening assays (the CH_{50} and AH_{50}), which determine the dilution of patient's serum needed to lyse 50% of erythrocytes sensitive to the CP (CH_{50}) or the AP (AH_{50}).[59] Deficiency of any C1 subcomponent, or any of the other CP components (C2–C8), will result in little or no lysis in the CH_{50} (CH_{50} values <5%). C9-deficient patients may have residual activity in this assay (CH_{50} values <30%). Little or no lysis is observed in the AH_{50} assay if factor D, P, or any of the components C3–C9 are deficient. Deficiency of factor B has recently been described.[60] By comparing the results of the two assays, it is possible to narrow down the search for the deficient component (Fig. 21.8). Hemolytic and antigenic assays may be done for each individual component to confirm the deficiency.

CP Deficiencies

Patients with deficiencies of early CP components (C1, C4, C2) are most commonly identified as having systemic autoimmune disease but are also at increased risk of infection.[3,4,56-58] The primary infectious agents in these patients are encapsulated bacteria, *S. pneumoniae, H. influenzae, N. meningitidis,* and *Streptococcus agalactiae,* which rely on Ab and CP opsonization for clearance.

C1 Deficiency

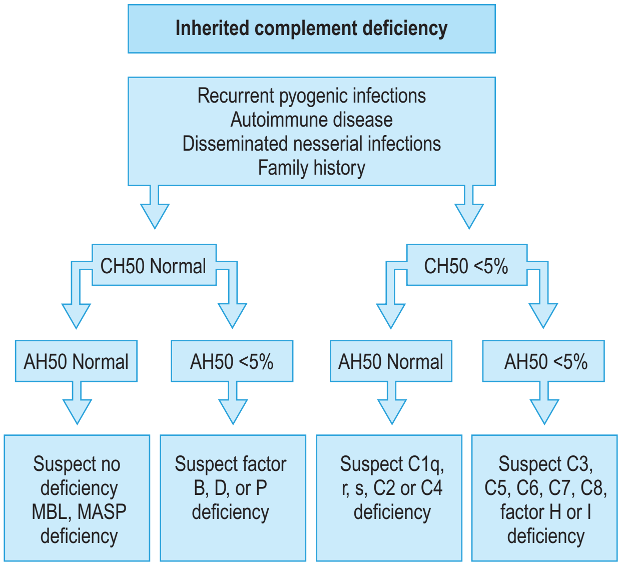
C1-deficient patients most commonly lack C1q, but C1r or C1s deficiency also results in nonfunctional C1 and no CP activity.

TABLE 21.3 Clinical Effects of Genetic Complement Deficiency

Deficient Component	Resulting Defect	Clinical Associations
C1q, C1r, C1s, C4 or C2	Inability to activate the CP	Systemic lupus erythematosus
Factor D, P	Inability to activate the AP	Infections, *Neisseria meningitidis*
MBL, MASP-2	Decreased or absent ability to activate the LP	Recurrent childhood infections, pyogenic bacteria
C3	Opsonization. No MAC. No activation of AP. Decreased inflammation (no C3a).	Recurrent childhood infections, *N. meningitidis, Streptococcus pneumoniae,* other encapsulated bacteria; autoimmune disease (uncommon)
FH, FI, C4 and C3NeFs	Lack of regulation of fluid-phase C3 convertases, severe acquired C3 deficiency	Infections, membranoproliferative glomerulonephritis
C5, C6, C7, C8, C9	Inability to form the MAC	Infection—recurrent, disseminated Neisserial
Serum carboxypeptidase-N	Failure to control C3a, C5a, bradykinin	Recurrent angioedema
C1-INH	Loss of regulation of C1 and bradykinin	Recurrent angioedema (HAE)
FH, FI, CD46 (haploinsufficiency)	Decreased regulation of C3 convertases	Atypical hemolytic–uremic syndrome, age-related macular degeneration*
DAF, CD59	Failure to regulate complement activation on autologous cells (especially red blood cells)	Paroxysmal nocturnal hemoglobinuria (PNH)

C1 INH, C1 esterase inhibitor; DAF, decay-accelerating factor; MAC, membrane attack complex; MASP, MBL-associated serine protease; MBL, mannan-binding lectin; NeF, nephritic factor (stabilizing autoAb to convertase); HAE, hereditary angioedema.

*Heterozygous C3 variants that lead to a gain of function cause atypical hemolytic–uremic syndrome (aHUS), age-related macular degeneration (ARMD), and C3G.


Inherited complement deficiency
Recurrent pyogenic infections
Autoimmune disease
Disseminated nesserial infections
Family history
CH50 Normal
CH50 <5%
AH50 Normal
AH50 <5%
AH50 Normal
AH50 <5%
Suspect no deficiency MBL, MASP deficiency
Suspect factor B, D, or P deficiency
Suspect C1q, r, s, C2 or C4 deficiency
Suspect C3, C5, C6, C7, C8, factor H or I deficiency

C9 deficiency may have up to 30% normal CH50 with low AH50

FIG 21.8 Flow Chart for the Evaluation of Inherited Complement Deficiencies Using Hemolytic Screening Assays for the Classical Pathway (CH_{50}) and the Alternative Pathway (AH_{50}). For each assay, the entire activation pathway including the membrane attack complex (MAC) is required for lysis.

CLINICAL PEARLS

Value of Screening for Complement Deficiencies

1. Patients with recurrent bacterial infections and normal white blood cells and immunoglobulins should be analyzed for a complement deficiency (obtain CH_{50} and AP_{50}).
2. Patients with recurrent or disseminated neisserial infection should be evaluated for deficiency of C3–C9 by CH_{50} and for properdin by AP_{50}.
3. Prophylactic antibiotics and immunization should be considered in complement-deficient individuals, especially for pneumococcus and neisserial species.
4. Patients with systemic lupus erythematosus (SLE) (especially young children and those with familial lupus, and recurrent bacterial infections) should be screened with a CH_{50}.

Absence of C1q is highly associated with the development of SLE, with an incidence of 90%.[56-59] It has been proposed that this association is related to defective clearance of apoptotic cells.[61] Apoptotic cells may be opsonized by IgM, or pentraxins, leading to activation of the CP, which may be initiated by IgM or pentraxin (CRP and SAP). Cells can also be cleared by direct C1q binding, leading to attachment and uptake through other phagocytic receptors (*e.g.*, phosphatidylserine receptor).[52] Other proposed mechanisms to account for the strong association between C1 and C4 deficiency (see below) and SLE include defective immune complex clearance and defective development and maintenance of B-cell tolerance.

C4 Deficiency

There are two C4 genes, C4A and C4B, located within the major histocompatibility complex (MHC) on chromosome 6.[58] The two forms of C4 have similar function, but different substrate preferences for the covalent binding reaction that occurs on activation to C4b. C4A is more efficient in attaching to amino groups on proteins, such as immune complexes, whereas C4B is more efficient in attaching to carbohydrates. Complete C4 deficiency requires four null alleles and is rarely found but is highly associated with SLE (75% incidence). Partial C4 deficiencies with one to three null alleles, however, are relatively common, found in up to 25% of individuals. Complete C4A deficiency is greatly overrepresented in the SLE population. C4A deficiencies are found in about 1% of the general population and 10–15% of patients with SLE. Complete C4B deficiencies are more commonly associated with bacterial infections, suggesting that the functionally different C4 genes contribute differently to host defense and autoimmunity.

C2 Deficiency

The gene for C2 is also located within the MHC. C2 deficiency is the most common complete complement deficiency, with about a 0.01% incidence in the population.[56-58] About half of C2-deficient individuals are clinically normal. The remaining individuals have recurrent pyogenic infections and/or rheumatological diseases. The most common infectious agents are *S. pneumoniae, H. influenzae, N. meningitidis,* and *S. agalactiae.* Infections are invasive and mainly occur in childhood, suggesting that the immune defect may be overcome by development of acquired immune defenses. Rheumatological diseases include SLE (15%), vasculitis, polymyositis, and Henoch-Schönlein purpura. SLE associated with C2 deficiency has some features that distinguish it from other types of SLE; these features include equal expression in males and females, early onset, increased photosensitivity, decreased renal disease, lower frequency of antidsDNA Ab, and higher frequency of anti-SSA/Ro and anti-C1q Ab.[56-58]

LP Deficiencies

MBL deficiency was originally found as a serum defect in the opsonization of yeast in pediatric patients with recurrent infections. There are multiple MBL polymorphisms in the population, in both the promoter and coding regions, and MBL deficiency is common (estimated to be 14% in the normal Swedish population).[56] In addition to the association of MBL deficiency in children with recurrent infections, there is a two- to threefold increased frequency of MBL deficiency in SLE, and these individuals have more frequent and more severe infections during the course of their disease. Serious infectious complications are also more frequent in the subgroups of patients with cystic fibrosis and RA with MBL deficiency.

A single homozygous MASP-2 deficiency has been reported.[7] The patient was asymptomatic until the age of 13 years when he was diagnosed with ulcerative colitis. Additional autoimmune manifestations developed along with recurrent severe infections with *S. pneumoniae.*

AP DEFICIENCIES

Individuals with complete deficiencies of factor D or P have been reported. Patients with factor D deficiency have presented with recurrent infections by *Neisseria* and other organisms. Properdin deficiency is X-linked, and patients most commonly have severe childhood infections with *N. meningitidis.*[18,56,59]

C3 Deficiencies

C3 is central to all three complement activation pathways. Nineteen families with primary inherited deficiency of C3 have been reported.[62] The most common presentation is recurrent life-threatening infections in early childhood (before the age of 2 years), sometimes followed by immune complex disease. The infections observed are primarily respiratory tract infections (48%) and meningitis (34%) with a variety of pathogens, especially encapsulated bacteria. The organisms most often involved are *N. meningitidis* and *S. pneumoniae,* but other encapsulated gram-negative and gram-positive bacteria have also been observed. Recurrent infections are seen in more than 50% of patients with C3 deficiency. This clinical presentation is similar to that seen in hypogammaglobulinemia. Individuals with C3 deficiency may develop renal disease (26%), including membranoproliferative glomerulonephritis (MPGN) and mesangiocapillary glomerulonephritis (MCGN) and autoimmune disease (26%), most commonly SLE.

Acquired C3 Deficiency: Genetic Deficiencies of FH and FI and C3 and C4 Nephritic Factors

Factors H and I are required to control C3 convertase in the fluid phase of the AP. Complete deficiency of either protein results in C3 cleavage and depletion to very low levels. C5, factor B, and P levels may also be reduced. The clinical presentation of patients with FH or FI deficiency resembles that of patients with primary C3 deficiency.[62] The highest disease association is recurrent infection with *N. meningitidis* and *S. pneumoniae,* and there is also an increased incidence of SLE. FH deficiency is more commonly associated with renal disease compared with C3 or FI deficiency (73% of individuals with FH deficiency compared with 13% of individuals with FI deficiency and 26% of those with C3 deficiency).

Nephritic factors (NeFs) are autoantibodies specific to the CP or the AP C3 convertase (C4b2a or C3bBb) or the AP C5 convertase that stabilizes these enzyme complexes and prevents normal regulatory control. The AP C3Nef induces unregulated complement activation, resulting in acquired C3 deficiency. NeFs are often associated with MPGN type II. C3NeF is also associated with partial lipodystrophy, a condition in which fat is lost from the waist upward.

Deficiencies of Complement Receptors

Deficiencies of CR1 (CD35) and CR2 (CD21)

The complete genetic deficiencies of CR1 or CR2 have not been reported. However, partial deficiencies of CR1 on erythrocytes, B lymphocytes, and polymorphonuclear leukocytes and of CR2 on B lymphocytes have been reported in patients with SLE. Decreased CR1 on erythrocytes may be acquired as a result of immune complex clearance.[3,4,56,59]

Leukocyte Adhesion Deficiency: CR3 and CR4 Deficiency

Leukocyte adhesion deficiency (LAD; Chapter 22) is a syndrome caused by mutations of the common β_2-integrin chain, CD18, found in LFA-1, CR3, and CR4. Defects are related to adhesion and activation of phagocytic cells, and the clinical presentation includes childhood infections with pyogenic bacteria.

Deficiencies of Regulatory Proteins

Hereditary Angioedema: C1-INH Deficiency

Hereditary angioedema (HAE) is found in individuals with heterozygous (autosomal dominant pattern of inheritance) deficiency of C1-INH.[24] C1-INH is a serine protease inhibitor (serpin) with regulatory activity for C1r, C1s, MASP-1, and MASP-2 of the complement system; factor XII (Hageman factor) and kallikrein of the contact system; factor XI and thrombin of the coagulation system; and plasmin and tissue plasminogen activator (tPA) of the fibrinolytic system. Although previous studies implicated a C2 product (C2 kinin) as a mediator, more recent data, including studies in a C1-INH deficient mouse model, indicate that bradykinin is the primary biological mediator of angioedema in HAE.[24] In the more common form of HAE (type I, 85% of patients), reduced synthesis of C1-INH is found (5–30% of normal), along with decreased serum C4 and C2. In type II HAE, an abnormal C1-INH is synthesized, making antigenic levels normal or elevated with reduced functional activity and decreased C4 and C2. Clinically, type I and type II HAE are indistinguishable.

HAE presents in childhood or adolescence as recurrent episodes of swelling that are subcutaneous and/or submucosal, nonpainful, nonpruritic, and nonpitting. Urticaria is not present. Episodes are self-limiting, usually peaking at 24 hours and resolving over 2–5 days. Attacks are variable in frequency, severity, duration and location, and initiating factors are poorly understood. The most common areas involved are the extremities, face, genitals, and respiratory and gastrointestinal tracts. Intestinal attacks are often associated with vomiting and diarrhea and are extremely painful (partial obstruction from the bowel wall edema). Laryngeal attacks may result in life-threatening respiratory tract narrowing.

Recurrent attacks continue throughout the life of the patient and may involve multiple sites or progress from one site to another. Diagnosis of HAE is suggested by family history and clinical findings. Confirmation is based on decreased C1-INH functional activity (<10–35% of normal). It is important to note that although C1-INH protein is decreased in type I HAE, it can be normal or even elevated in type II HAE. C4 levels are below normal in 95% of patients with HAE. Acquired forms of C1-INH deficiency have been described, usually in older patients with lymphoproliferative diseases. These are usually caused by autoantibodies to C1-INH and are distinguished from HAE by a lack of family history and decreased C1q as well as C4. The management and treatment of HAE are discussed in Chapter 42.

Paroxysmal Nocturnal Hemoglobinuria: DAF and CD59 Deficiency

Paroxysmal nocturnal hemoglobinuria (PNH) is a rare acquired disorder in which a somatic mutation in the *Pig-A* gene in a clone of bone marrow stem cell results in defective synthesis of GPI-anchored proteins. PNH is characterized clinically by intravascular hemolysis and venous thrombosis. DAF and CD59 are GPI-anchored complement regulatory proteins expressed on erythrocytes, and PNH erythrocytes are highly susceptible to lysis. Studies of individuals with isolated DAF and CD59 deficiencies indicate that hemolysis is more highly associated with CD59 deficiency. The basis for thrombosis in PNH is poorly understood. A mAb to C5 has been approved by the U.S. Food and Drug Administration (FDA) to treat PNH.

Control of Localized Complement Activation: Atypical Hemolytic–Uremic Syndrome, Age-Related Macular Degeneration

HUS is a rare disease characterized by microangiopathic hemolytic anemia, thrombocytopenia, and acute renal failure. “Typical” HUS is found in children and is caused by *E. coli,* mainly O157:H7, producing a shiga-like toxin. Atypical HUS affects older children and adults and is not associated with an enteropathic infection. Recently, mutations in the complement regulatory proteins FH, FI, or CD46 have been identified in approximately 50% of patients with atypical HUS. FH mutations associated with HUS are clustered in the C-terminal end of the molecule in CCP20, a region that is required for FH binding to polyanions and endothelial cells. The ability of FH to regulate fluid-phase AP activation is not affected, and C3 levels are normal. These findings have led to the hypothesis that local complement regulation is essential for preventing renal disease following endothelial cell injury and that FH acts locally after binding to exposed matrix or damaged endothelium. FH and the membrane protein CD46 are both cofactors for FI-mediated cleavage of C3b.

An additional FH polymorphism (Tyr/His402) identified by genetic screening has been shown to be associated with the development of ARMD, a major cause of blindness in older adults.[63] This polymorphism is located in CCP7 in a region of FH that binds heparin and CRP.[28,64,65] As is the case for the mutations associated with atypical HUS (aHUS), this region of FH is not required for regulation of the fluid-phase AP convertase. ARMD develops when abnormal deposits of protein, termed *drusen,* form in the retina. Recent findings support the view that the local inflammatory response, including complement activation with MAC deposition, damages the retina, leading to vision loss. Additional genetic analyses identified protective FH and FI variants, as well as protective and high-risk polymorphisms, in the *FI* and *FB* genes. Although complement factors are not the only genes linked to ARMD, they are estimated to account for more than 50% of cases. Up to 10% of cases of advanced ARMD may carry a rare variation in FH or FI. These variants commonly lead to haploinsufficiency. These findings are driving the development of new complement-based therapeutics that could provide protection from a very common form of age-related visual loss.

COMPLEMENT IN DISEASE

Measurement of Complement in a Clinical Setting

Laboratory tests for complement include functional assays for the CP (CH_{50}), the AP (AH_{50}), and the LP (LP_{50}). Functional and antigenic assays for each of the individual components are available in specialty laboratories. The CH_{50} is a hemolytic assay, in which sheep erythrocytes sensitized with rabbit Ab are incubated with serial dilutions of the patient’s serum. The titer is the dilution at which 50% of the sheep erythrocytes are lysed. The CH_{50} requires all of the CP and terminal components (C1–C9). A comparable assay for the AP uses a buffer that blocks CP activation and employs rabbit erythrocytes in place of sensitized sheep erythrocytes. Rabbit erythrocytes spontaneously activate the human AP and are lysed in the assay. The AH_{50} requires all of the AP and terminal components (factor B, D and P and C3–C9). The combined use of the CH_{50} and AH_{50} is the most effective screening method for genetic deficiencies of complement components. Complete deficiency will generally result in titers of <5% in one or both assays. Because C3–C9 are common to both pathways, the combined results of the two assays can rapidly determine whether the deficiency is one of these shared components, one of the CP components (C1, C2, C4) or one of the AP components (factors B, D, P) (Fig. 21.8). Properdin deficiency results in low, but not absent, lysis in the AH_{50}, and patients with C9 deficiency may have values up to 30% of normal in the CH_{50}. Deficiencies of FH and FI and nephritic factors often result in very low C3 levels, leading to reduced titers in both assays.

LP function (and MBL deficiency) is determined by using a specific ELISA, in which the patient’s serum is placed into wells coated with mannan. Binding MBL and activation of the LP results in the deposition of C4b and C4d that are detected with mAbs. MBL levels may also be determined antigenically.

Heterozygous C1-INH deficiency, as described above, is associated with the clinical syndrome of HAE.[24] The diagnosis can be made on the basis of clinical findings and family history. C1-INH activity is reduced in these patients, and C4 protein is also low in 95% of patients, especially during attacks of edema. In type I HAE (85% of cases), C1-INH protein levels are low, but in type II HAE (15% of cases), an abnormal C1-INH protein is made, and antigenic levels are normal or elevated. There is an acquired form of C1-INH deficiency associated with lymphoma, in which low C1-INH is accompanied by decreased C1q as well as C4 and C2.

Complement levels may also be decreased in diseases or conditions in which complement is activated, leading to consumption. In contrast to genetic deficiencies, complement consumption characteristically results in low, but not absent, functional activity. In addition, multiple components of one or more pathways are expected to be low, and these decreased levels of complement are often correlated with disease activity. The most commonly used and most readily available complement tests are C3 and C4 protein and the CH_{50}. Diseases accompanied by CP activation

TABLE 21.4 Complement Test Interpretation

Pathway	CH_{50}	C4	C3	Related diseases
Classical pathway (CP)	↓	↓	↓	SLE, serum sickness, vasculitis, Subacute Bacterial Endocarditis, MPGN (type I)
Alternative pathway (AP)	↓	N	↓	Poststreptococcal glomerulonephritis, MPGN (type II)
Fluid-phase activation of the CP	↓	↓	N	C4NeF, HAE, cryoglobulinemia
Fluid-phase activation of the AP	↓	N	↓	FH or FI deficiency, C3NeF, MPGN (type II)
Acute-phase response	↑	↑	↑	Acute and chronic inflammation
Decreased CH_{50} (sample collection problems)	↓	N	N	Cryoglobulins, cold activation, sample mishandling; coagulation-associated activation
Decreased CH_{50} (biosynthetic)	↓	N	↓	Severe liver disease; decreased C3, C6, C9

N, normal; SLE, systemic lupus erythematosus; MPGN, membranoproliferative glomerulonephritis; HAE, hereditary angioedema; NeF, nephritic factor.

result in decreased CH_{50}, C4, and C3 levels. The AP is usually spared. These are primarily immune complex–associated diseases, both autoimmune and infectious, and are listed in Table 21.4. In addition, 20% of cases of acute renal allograft rejection are associated with decreased CH_{50} and C2. Another cause of selective CP activation is essentially a laboratory artifact, in which clotting of the blood sample in the cold is associated with consumption of the early CP. Plasma CH_{50} is normal, but the serum CH_{50} value is markedly decreased. C3 and C4 antigenic tests are normal, but their functional activity is lost.

The AP is activated in gram-negative sepsis, poststreptococcal glomerulonephritis, MPGN, IgA nephropathy, FH or FI deficiency, and PNH. Laboratory values may show decreased C3 with decreased or normal CH_{50}, and normal C4 levels (see Table 21.4). AP activation is not always reflected in decreased C3 because C3 is found at the highest concentration of all complement components and is an acute-phase reactant with elevated synthesis during disease states.

In clinical practice, evaluation of complement levels may be useful in a variety of circumstances. Initial consideration of complement deficiency may be appropriate in patients presenting with autoimmune conditions or repetitive pyogenic infections in the setting of a normal white blood cell (WBC) count and immunoglobulinemia (see Table 21.3). The complement profile can also be helpful in differential diagnoses of SLE and its look-alikes (see Table 21.4). Monitoring complement levels is frequently used to follow disease activity in patients with SLE. Complement levels may predict renal disease activity and may reflect a response to therapy in SLE. However, complement levels are rarely useful in isolation and should be taken in the context of clinical assessment and other laboratory values as reflected in disease activity index scores (*e.g.*, the Systemic Lupus Erythematosus Disease Activity Index 2000).

Role of Complement in Specific Immunological Diseases

Complement activation is involved in the pathogenesis of many immunological diseases. A general concept emerging from current research in this area is that most of the pathogenic effects of complement depend on the generation of C5a and the MAC. Further, it is becoming recognized that regardless of the initial activation mechanism, AP amplification is often needed to produce sufficient quantities of these mediators to cause disease. Finally, systems biology approaches are increasingly revealing that members of the complement cascade interact with other inflammatory mediators, resulting in diseases that are a product of complex gene–environment interactions, such as asthma and Alzheimer disease.

Systemic Lupus Erythematosus (Chapter 51)

CLINICAL PEARLS

Complement Tests for Diagnosis and Monitoring of Systemic Lupus Erythematosus (SLE)

A low C4 and C3 assist in the diagnosis of SLE.

Decreased C3 and C4 are associated with increased severity of disease, and especially with lupus nephritis.

On serial observations, decreasing C3 and C4 levels predict and help to establish an SLE flare-up.

Note: Decreases in C4 may precede decreases in C3.

Remission after treatment of lupus often shows return toward normal levels of C4, followed by increases in C3.

Note: Patients with SLE who have partial C4 deficiency may have persistently low C4 levels.

Complete absence of CH_{50} implies the existence of a hereditary deficiency of one of the classical complement pathway components, usually C1q, C4, or C2.

Complement plays a dual role in SLE.[56,57,66,67] There is a strong association of genetic deficiencies of C1q, C1r, C1s, C4, C2, and, to a lesser degree, C3 with SLE, indicating a protective role. Three main complement-dependent mechanisms have been proposed: (1) complement-dependent clearance of immune complexes; (2) modulating the adaptive immune system, particularly through the development and maintenance of self-tolerance in B lymphocytes; and (3) a requirement for complement in the clearance of apoptotic cells and potential autoantigens released from damaged cells. The pathogenesis of SLE results, in large part, from inflammatory response to immune complexes formed by autoantibodies (*e.g.*, Ab to double-stranded DNA [antidsDNA]) binding to antigens from dead and dying cells. However, complement activation is believed to play a pathogenic role in tissue damage induced by autoantibodies in SLE. There is evidence for complement activation in skin and renal lesions of patients with SLE, as well as in autoantibody-mediated hemolytic anemia and thrombocytopenia.

Antiphospholipid Syndrome (Chapter 61)

Antiphospholipid syndrome is characterized by antiphospholipid Ab, recurrent fetal loss, vascular thrombosis, and thrombocytopenia. Antiphospholipid Abs are found in 50% of patients with SLE, and thrombotic events occur in about 50% of them. Antiphospholipid Abs identified in patients without SLE have similar clinical consequences. Disease pathogenesis has been attributed to the procoagulant effects of antiphospholipid Abs. A mouse model of antiphospholipid Ab syndrome has been used to demonstrate that injection of pregnant mice with human IgG antiphospholipid Ab results in fetal loss and wasting. In this model, complement is required for pathogenesis, and treatment with complement inhibitors is protective. Studies in the mouse model are consistent with initial complement activation by antiphospholipid Ab bound to the decidua, followed by C5a generation and recruitment of neutrophils. The AP as well as the CP was required for pathology. Interestingly, C3 deposition in the decidua was decreased if neutrophils were depleted, suggesting an amplification pathway mediated either by tissue damage or by neutrophil release of complement components.

Rheumatoid Arthritis (Chapter 52)

Patients with RA generally have normal or elevated complement values systemically.[67,68] There is, however, evidence for local complement activation in joint fluid, in synovia, and in rheumatoid nodules. In addition to being elevated in the joints of patients with RA, complement activation products are also found in patients with osteoarthritis, SLE, Reiter syndrome, and gout. Concentrations of C3a and C5a in joint fluid are higher in RA than in other types of arthritis. An important role for complement activation in the pathogenesis of RA is suggested by studies in two animal models—collagen-induced arthritis and the K/BxN-derived Ab transfer model. In the first model, inflammatory joint disease was ameliorated by treatment with an Ab to C5 that blocks its cleavage, preventing generation of C5a and the MAC. In the second model, disease was prevented by genetic deficiency of factor B, but not C4, indicating an essential involvement of the AP.

Vasculitis (Chapters 58, 59)

Human vasculitides encompass a spectrum of disease mechanisms and clinical manifestations. Some, such as giant-cell arteritis and the antineutrophil cytoplasmic Ab (ANCA)–associated vasculitides, Wegener granulomatosis, microscopic polyangiitis, and Churg-Strauss syndrome, are not associated with local complement deposition or evidence of systemic complement depletion. Despite that, a "self-fueling inflammatory amplification loop," as a result of the generation of C5a by activated neutrophils and neutrophil priming by C5a, appears to drive necrotizing vascular injury. Additionally, in vasculitides associated with circulating immune complexes, C3b, MAC, and/or AP components are deposited in lesions, and complement profiles consistent with activation are found (see Table 21.4).

Immunological Renal Diseases (Chapter 68)

Complement activation is evident in most types of glomerulonephritis, with the site and pathway of activation dependent on the site of immune complex or autoantibody deposition. AP activation has been identified in IgA nephropathy, poststreptococcal glomerulonephritis, and MPGN type II. More recent results have implicated activation of the LP in IgA nephropathy. Glomerular deposition of MBL has been associated with greater histological damage and higher proteinuria.

MPGN is a chronic progressive form of glomerulonephritis characterized by production of enlarged glomerular tufts by endocapillary proliferation and thickening of glomerular capillaries. MPGN has been historically divided into three histological groups, designated type I, II, and III based on electron microscopy of the glomerular lesions. Complement activation has been detected in all forms of MPGN, with decreases in circulating complement component levels and the presence of C3 on biopsy, in addition to being seen with other glomerulonephritides (*e.g.*, SLE). Given that the historical classification of MPGN did not help delineate disease pathogenesis, MPGN is now increasingly been viewed as immunoglobulin-mediated (associated with CP activation) and non–immunoglobulin-mediated (associated with AP activation). Additionally, the term *C3 glomerulopathy* is being used to describe those cases of glomerular involvement secondary to the AP activation.

In glomerulonephritis secondary to immune complex disease (immunoglobulin-mediated, such as SLE and MPGN type I), complement activation is primarily by the CP, and C4 is detected along with C3 and IgG in glomerular deposits. Complement activation contributes to renal disease by attracting and activating inflammatory cells through the anaphylatoxin C5a and by direct damage to cells through the MAC. Pathology caused by inflammatory cell infiltration is predominant when subendothelial immune complex deposition and complement activation occur.

In contrast, in non–immunoglobulin-mediated glomerulonephritis (*e.g.*, MPGN type II), defects in the tightly regulated AP are believed to result in excessive activity of the C3 convertase. This can occur either in the presence of the C3 nephritic factor (C3Nef), a stabilizing autoantibody, or in the setting of deficient functional FH activity, either through mutations or acquired defects. C3NeF is a pathogenic autoantibody that binds to the AP C3 convertase (C3bBb), preventing its decay and regulation by FH and FI. Absence of functional FH results in unregulated C3 convertase activity, resulting in uncontrolled glomerular inflammation and renal disease. Understandably, such reclassification has helped target treatment, for example, by using plasma infusion or exchange and even the anti-C5 mAb eculizumab in certain cases.

Asthma (Chapter 41)

Asthma is a chronic inflammatory disease of the lung, in which Th2 responses to environmental allergens frequently play a critical role. The development of mice deficient in receptors for C3a and C5a has led to a new understanding of the roles of the complement anaphylatoxins in asthma. Several studies have demonstrated a correlation between C3a and C5a release in asthmatic lungs and the influx of eosinophils and neutrophils. C3-deficient and C3aR-deficient mice were protected from development of acute bronchoconstriction, airway inflammation, and airway hyperresponsiveness. C5a inhibition had similar effects on the response to challenge in an established allergic environment. However, in contradiction to these findings, C5 deficiency was genetically linked to susceptibility to experimental allergic asthma. Further studies found that C5a signaling (most likely through the C5aR on pulmonary DCs) during initial pulmonary exposure to allergen decreased Th2 cytokine and IgE production, thereby preventing the initiation of the allergic response. Thus it appears that both the C3a-C3aR and C5a-C5aR axes contribute to asthma pathogenesis. However, how disruption of their homeostatic roles on different immune cells versus the bronchial epithelium contributes to asthma pathogenesis remains to be understood.

Neurological Disease

Proteins from the complement system are normally found in the central nervous system (CNS) and the peripheral nervous systems. Low levels of hemolytic complement (0.25% of serum levels) can be measured in the cerebrospinal fluid if care is taken to stabilize it with gelatin during storage. Levels of anaphylatoxins are increased in the CNS when the blood–brain barrier is impaired. Complement proteins and regulatory proteins are synthesized by glial cells and astrocytes, and their synthesis is enhanced by inflammatory cytokines, such as IL-6. There is evidence both from human multiple sclerosis (MS) (Chapter 66) and the animal model, experimental allergic encephalitis (EAE), that complement activation with the generation of the MAC contributes to the demyelination in these diseases. Generation of the MAC can lead to oligodendrocyte death, generation of inflammatory mediators, or a repair process in which myelin synthesis is decreased. Complement activation on myelin and oligodendrocytes is initiated by antimyelin Ab or directly by myelin through the CP. There is evidence of MAC formation in the cerebrospinal fluid of patients with MS, and complement depletion, inhibition, and genetic deficiency are protective in rat and mouse models of EAE.

There is also evidence of complement activation in degenerative neurological conditions, such as Alzheimer disease. In Alzheimer disease, neurofibrillary tangles and senile plaques composed of β-amyloid and other proteins develop, resulting in neuronal loss and dementia with progressive loss of cognitive function. Complement activation products C1q, C4, C3, and MAC components, as well as clusterin (ApoJ) and vitronectin (S40), are found deposited in areas of β-amyloid, suggesting CP activation. Peptides derived from β-amyloid were shown to activate C1 directly by binding to the collagen-like domain. SAP, a component of all types of amyloid, including β-amyloid, activates the CP as well. There are limited data on the role of complement in the pathogenesis of Alzheimer disease, with some studies reporting enhanced disease following complement inhibition and another finding decreased inflammatory changes and neuronal degeneration in C1q deficiency. Finally, excessive complement activity, notably C4, has been implicated in the development of schizophrenia and has been associated with reduced numbers of synapses. This suggests that the role of complement proteins in neuropsychiatric illness extends beyond inflammation-mediated tissue damage.[69,70]

Ischemia/Reperfusion Injury

I/R injury refers to injury induced by inflammatory mediators, such as reactive oxygen intermediates produced by activated neutrophils, following the reperfusion of hypoxic tissue. Different pathways of complement activation may be important in different sites of injury, probably because of differences in expression of complement regulatory proteins and the nature of the tissue damage and the antigens exposed to the innate immune system. The primary complement mediators of tissue injury are C5a and the MAC acting locally and, in some cases, C5a acting systemically. In experimental renal I/R injury and in human tubular necrosis, the AP appears to be directly activated and neither Ab nor the CP is required. However, in intestinal I/R injury, the CP as well as the AP are required, and a natural IgM Ab to a newly exposed antigen on damaged endothelium initiates complement activation. In coronary artery ligation/reperfusion models, innate recognition of epitopes of ischemic tissue by MBL and CRP leads to lectin and CP activation, respectively.

COMPLEMENT-BASED THERAPEUTICS

The multiple roles of complement in inflammatory and autoimmune diseases make it an attractive target for therapeutic intervention. Recombinant complement inhibitors, inhibitory mAb, and peptide-based receptor inhibitors have been developed to block the detrimental effects of the complement activation fragments.[70] As described above, products of the complement cascade have many beneficial effects in host defense and the adaptive immune response. The detrimental effects of complement activation are, for the most part, associated with C5a and the MAC. Thus targeting either the generation of C5a or its association with C5aR might be expected to control inflammation while maintaining other important functions, such as opsonization. An anti-C5 mAb that prevents C5 cleavage (eculizumab) has been approved for human use in the treatment of PNH[70] and atypical HUS. It is also being evaluated for treatment of refractory glomerulonephritides with C3 deposition and severe Ab-mediated rejection after organ transplantation. Other drugs targeting the C5 pathway, as well as mAb directed to components of the AP (*e.g.*, factor D), are under investigation for ARMD. Peptides and mAb directed at the C5aR, as well as upstream of C3 convertase production, have shown promise in a number of inflammatory models in animals and are being evaluated for the treatment of sepsis, reperfusion injury, and asthma. Other approaches that are being developed will target complement regulatory proteins to specific cell or tissue targets. As the importance of this system is clarified in a variety of inflammatory diseases, it is likely that further research will establish new complement-based therapeutic agents for additional applications.

ON THE HORIZON

Future Directions in Complement Research

Functional analysis of polymorphisms and rare variants in complement proteins identified in genome-wide association studies (GWAS) and by next generation sequencing of inflammatory and autoimmune diseases will be functionally characterized to provide insight into pathogenesis and treatment.

Genetic sequencing of entire complement activation pathways and their regulators and receptors in patients with inflammatory and autoimmune diseases will reveal novel pathogenic mechanisms and approaches to diagnosis and therapy.

RNA sequence analysis will identify "up and down" regulation of complement proteins in human disease.

Structural analysis of complement protein complexes will lead to targeted small molecules to inhibit or enhance complement activation.

Therapeutic trials of existing agents and those in development will dramatically refine therapy of complement-mediated diseases.

Proteome studies in patients with infectious, inflammatory, and autoimmune diseases will reveal patterns of complement activation and biomarkers for diagnosis, disease activity, and monitoring responses to therapy.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES/BIBLIOGRAPHY

With a few exceptions, the citations for the basics of complement biology are reviews related to the sections on the role of complement in human disease; few references are noted because of space limitations, but each topic is covered in other chapters (as noted for the reader).

1. Ricklin D, Hajishengallis G, Yang K, et al. Complement: a key system for immune surveillance and homeostasis. Nat Immunol 2010;11:785–97.
2. Atkinson JP. Complement system in disease. In: *Goldman-Cecil Medicine*, vol. 1. 25th ed. 2015. p. 240–6.

3. Walport MJ. Complement. First of two parts. N Engl J Med 2001;344:1058–66.
4. Walport MJ. Complement. Second of two parts. N Engl J Med 2001;344:1140–4.
5. Marnell L, Mold C, Du Clos TW. C–reactive protein: Ligands, receptors and role in inflammation. Clin Immunol 2005;117:104–11.
6. Kang YS, Do Y, Lee HK, et al. A dominant complement fixation pathway for pneumococcal polysaccharides initiated by SIGN–R1 interacting with C1q. Cell 2006;125:47–58.
7. Sorensen R, Thiel S, Jensenius JC. Mannan-binding-lectin-associated serine proteases, characteristics and disease associations. Springer Semin Immunopathol 2005;27:299–319.
8. Lachmann PJ. The amplification loop of the complement pathways. Adv Immunol 2009;104:115–49.
9. Zewde N, Gorham RD Jr, Dorado A, et al. Quantitative Modeling of the Alternative Pathway of the Complement System. PLoS ONE 2016;11:e0152337.
10. Janssen BJ, Christodoulidou A, McCarthy A, et al. Structure of C3b reveals conformational changes that underlie complement activity. Nature 2006;444:213–16.
11. Wiesmann C, Katschke KJ, Yin J, et al. Structure of C3b in complex with CRIg gives insights into regulation of complement activation. Nature 2006;444:217–20.
12. Forneris F, Wu J, Xue X, et al. Regulators of complement activity mediate inhibitory mechanisms through a common C3b-binding mode. EMBO J 2016;35:1133–49.
13. Gaboriaud C, Juanhuix J, Gruez A, et al. The crystal structure of the globular head of complement protein C1q provides a basis for its versatile recognition properties. J Biol Chem 2003;278:46974–82.
14. Gaboriaud C, Thielens NM, Gregory LA, et al. Structure and activation of the C1 complex of complement: unraveling the puzzle. Trends Immunol 2004;25:368–73.
15. Dodds AW, Ren XD, Willis AC, et al. The reaction mechanism of the internal thioester in the human complement component C4. Nature 1996;379:177.
16. Matsushita M. Ficolins: complement-activating lectins involved in innate immunity. J Innate Immun 2010;2:24–32.
17. Takahashi M, Ishida Y, Iwaki D, et al. Essential role of mannose-binding lectin-associated serine protease-1 in activation of the complement factor D. J Exp Med 2010;207:29–37.
18. Kemper C, Atkinson JP, Hourcade DE. Properdin: emerging roles of a pattern–recognition molecule. Annu Rev Immunol 2010;28:131–55.
19. Agarwal S, Ferreira VP, Cortes C, et al. An evaluation of the role of properdin in alternative pathway activation on Neisseria meningitidis and Neisseria gonorrhoeae. J Immunol 2010;185:507–16.
20. Bubeck D. The making of a macromolecular machine: assembly of the membrane attack complex. Biochemistry Mosc 2014;53:1908–15.
21. Lambris JD, Ricklin D, Geisbrecht BV. Complement evasion by human pathogens. Nat Rev Microbiol 2008;6:132–42.
22. Cole DS, Morgan BP. Beyond lysis: how complement influences cell fate. Clin Sci 2003;104:455–66.
23. Zipfel PF, Skerka C. Complement regulators and inhibitory proteins. Nat Rev Immunol 2009;9:729–40.
24. Zuraw BL. Clinical practice. Hereditary angioedema. N Engl J Med 2008;359:1027–36.
25. Liszewski MK, Farries TC, Lublin DM, et al. Control of the complement system. Adv Immunol 1996;61:201–83.
26. Giannakis E, Jokiranta TS, Male DA, et al. A common site within factor H SCR 7 responsible for binding heparin, C–reactive protein and streptococcal M protein. Eur J Immunol 2003;33:962–9.
27. Manuelian T, Hellwage J, Meri S, et al. Mutations in factor H reduce binding affinity to C3b and heparin and surface attachment to endothelial cells in hemolytic uremic syndrome. J Clin Invest 2003;111:1181–90.
28. Morgan HP, Schmidt CQ, Guariento M, et al. Structural basis for engagement by complement factor H of C3b on a self surface. Nat Struct Mol Biol 2011;18:463–70.
29. Kim DD, Song WC. Membrane complement regulatory proteins. Clin Immunol 2006;118:127–36.
30. Inal JM, Hui KM, Miot S, et al. Complement C2 receptor inhibitor trispanning: a novel human complement inhibitory receptor. J Immunol 2005;174:356–66.
31. Bohlson SS, Fraser DA, Tenner AJ. Complement proteins C1q and MBL are pattern recognition molecules that signal immediate and long–term protective immune functions. Mol Immunol 2007;44:33–43.
32. Liszewski MK, Atkinson JP. Membrane cofactor protein (MCP; CD46). Isoforms differ in protection against the classical pathway of complement. J Immunol 1996;156:4415–21.
33. Roozendaal R, Carroll MC. Complement receptors CD21 and CD35 in humoral immunity. Immunol Rev 2007;219:157–66.
34. Holers VM. Complement and its receptors: new insights into human disease. Annu Rev Immunol 2014;32:433–59.
35. van Lookeren Campagne M, Wiesmann C, Brown EJ. Macrophage complement receptors and pathogen clearance. Cell Microbiol 2007;9:2095–102.
36. Helmy KY, Katschke KJ Jr, Gorgani NN, et al. CRIg: a macrophage complement receptor required for phagocytosis of circulating pathogens. Cell 2006;124:915–27.
37. Huber–Lang M, Sarma JV, Zetoune FS, et al. Generation of C5a in the absence of C3: a new complement activation pathway. Nat Med 2006;12:682–7.
38. Klos A, Tenner AJ, Johswich KO, et al. The role of the anaphylatoxins in health and disease. Mol Immunol 2009;46:2753–66.
39. Ward PA. Functions of C5a receptors. J Mol Med 2009;87:375–8.
40. Gerard NP, Lu B, Liu P, et al. An anti-inflammatory function for the complement anaphylatoxin C5a-binding protein, C5L2. J Biol Chem 2005;280:39677–80.
41. Rittirsch D, Flierl MA, Nadeau BA, et al. Functional roles for C5a receptors in sepsis. Nat Med 2008;14:551–7.
42. Ojha H, Panwar HS, Gorham RD Jr, et al. Viral regulators of complement activation: structure, function and evolution. Mol Immunol 2014;61:89–99.
43. Yu Q, Yu R, Qin X. The good and evil of complement activation in HIV–1 infection. Cell Mol Immunol 2010;7:334–40.
44. Holers VM, Kulik L. Complement receptor 2, natural antibodies and innate immunity: Inter-relationships in B cell selection and activation. Mol Immunol 2007;44:64–72.
45. Kopf M, Abel B, Gallimore A, et al. Complement component C3 promotes T-cell priming and lung migration to control acute influenza virus infection. Nat Med 2002;8:373–8.
46. Kemper C, Chan AC, Green JM, et al. Activation of human $CD4^+$ cells with CD3 and CD46 induces a T-regulatory cell 1 phenotype. Nature 2003;421:388–92.
47. Hawlisch H, Kohl J. Complement and Toll–like receptors: key regulators of adaptive immune responses. Mol Immunol 2006;43:13–21.
48. Thurman JM, Holers VM. The central role of the alternative complement pathway in human disease. J Immunol 2006;176:1305–10.
49. Taylor PR, Carugati A, Fadok VA, et al. A hierarchical role for classical pathway complement proteins in the clearance of apoptotic cells *in vivo*. J Exp Med 2000;7:359–66.
50. Poon IK, Hulett MD, Parish CR. Molecular mechanisms of late apoptotic/necrotic cell clearance. Cell Death Differ 2010;17:381–97.
51. Du Clos TW, Mold C. Pentraxins (CRP, SAP) in the process of complement activation and clearance of apoptotic bodies through Fcgamma receptors. Curr Opin Organ Transplant 2011;16:15–20.
52. Fadok VA, Bratton DL, Henson PM. Phagocyte receptors for apoptotic cells: recognition, uptake, and consequences. J Clin Invest 2001;108:957–62.
53. Gershov D, Kim S, Brot N, et al. C-Reactive protein binds to apoptotic cells, protects the cells from assembly of the terminal complement components, and sustains an antiinflammatory innate immune response: implications for systemic autoimmunity.[erratum appears in J Exp Med 2001;193:1439]. J Exp Med 2000;192:1353–64.
54. Mold C, Gewurz H, D Clos TW. Regulation of complement activation by C-reactive protein. Immunopharmacology 1999;42:23–30.

55. Riley-Vargas RC, Lanzendorf S, Atkinson JP. Targeted and restricted complement activation on acrosome-reacted spermatozoa. J Clin Invest 2005;115:1241–9.
56. Sjoholm AG, Jonsson G, Braconier JH, et al. Complement deficiency and disease: an update. Mol Immunol 2006;43:78–85.
57. Bryan AR, Wu EY. Complement deficiencies in systemic lupus erythematosus. Curr Allergy Asthma Rep 2014;14:448.
58. Atkinson JP, Yu CY. The complement system in systemic lupus erythematosus. In: Tsokos GC, editor. *Systemic Lupus Erythematosus, Basic, Applied and Clinical Aspects*. 2016. p. 81–112.
59. Wen L, Atkinson JP, Giclas PC. Clinical and laboratory evaluation of complement deficiency. J Allergy Clin Immunol 2004;113:585–93.
60. Slade C, Bosco J, Unglik G, et al. Deficiency in complement factor B. N Engl J Med 2013;369:1667–9.
61. Seelen MA, Roos A, Daha MR. Role of complement in innate and autoimmunity. J Nephrol 2005;18:642–53.
62. Reis E, Falcao DA, Isaac L. Clinical aspects and molecular basis of primary deficiencies of complement component C3 and its regulatory proteins factor I and factor H. Scand J Immunol 2006;63:155–68.
63. Avery RL. The plague and macular degeneration. Ophthalmology 2010;117:2442.
64. Ferreira VP, Pangburn MK, Cortes C. Complement control protein factor H: the good, the bad, and the inadequate. Mol Immunol 2010;47:2187–97.
65. Makou E, Herbert AP, Barlow PN. Functional anatomy of complement factor H. Biochemistry Mosc 2013;2:3949–62.
66. Mayilyan KR. Complement genetics, deficiencies, and disease associations. Protein Cell 2012;3:487–96.
67. Sturfelt G, Truedsson L. Complement in the immunopathogenesis of rheumatic disease. Nat Rev Rheumatol 2012;8:458–68.
68. Ballanti E, Perricone C, Greco E, et al. Complement and autoimmunity. Immunol Res 2013;56:477–91.
69. Sekar A, Bialas AR, de Rivera H, et al. Schizophrenia risk from complex variation of complement component 4. Nature 2016;530:177–83.
70. Morgan BP, Harris CL. Complement, a target for therapy in inflammatory and degenerative diseases. Nat Rev Drug Discov 2015;14:857–77.

MULTIPLE-CHOICE QUESTIONS

1. All of the following are features of the complement system EXCEPT for:
   A. It works in a few minutes and features a feedback loop.
   B. It can interact (bind to) almost any foreign material or self-debris.
   C. It is abundant in plasma but also is secreted locally by many cells.
   D. It has a specific memory system for prior interactions with microbes.

2. All of the following are true about a complete deficiency of a complement component in the activating cascade EXCEPT:
   A. A deficiency of C5, C6, C7, C8, or C9 usually presents with a neisserial infection.
   B. A deficiency of properdin whose gene is on the X chromosome also presents in males with a neisserial infection.
   C. A complete deficiency of C3 commonly presents with early age onset of recurrent viral, fungal, and/or protozoan infections.
   D. In the classical pathway (CP), up to C3 (C1q, C4, C2), recurrent bacterial infections and, particularly, autoimmunity (systemic lupus erythematosus [SLE]) are the two most common clinical scenarios.

3. Which of the following is FALSE relative to measuring C4 and C3 antigenic levels in patients with SLE?
   A. A low factor B antigenic level is commonly observed and is indicative of alternative pathway (AP) activation.
   B. Low C4 and C3 point to activation of the CP by immune complexes.
   C. Reduced C4 and C3 correlate with renal disease, DNA antibodies, and, overall, more severe disease.
   D. A rise in C4 and C3 to normal levels with immunosuppressive therapy is associated with a patient having better short-term and long-term clinical outcomes.

4. Which statement is FALSE about complement regulators/inhibitors?
   A. Abundant in plasma.
   B. Expressed ubiquitously on cell membranes.
   C. Haploinsufficiency of factor H or factor I is well tolerated and does not lead to human disease.
   D. A complete deficiency of plasma protein factor H (FH) or plasma enzyme factor I (FI) leads to overactivation of the alternative pathway (AP) and consumption of C3 (clinical presentation like that of C3 deficiency).

5. Which statement is FALSE about complement receptors:
   A. Complement receptors are abundant on neutrophils, monocytes, and B lymphocytes.
   B. Complement system does not participate in adaptive immunity other than by serving as a “complement” to immunoglobulin M (IgM)– and IgG-mediated events.
   C. Complement receptor one (CR1, CD35) is present on red blood cells, where it carries out immune adherence for C4b/C3b–coated materials and then takes them to the spleen and/or the liver for transfer to macrophages.
   D. Complement receptors and Fcγ receptors work together to mediate opsonization and ingestion of infectious bacteria and viruses.

22

Phagocyte Deficiencies

Steven M. Holland, Gülbü Uzel

We have learned a great deal about phagocytes since their discovery by Metchnikoff in 1905: Neutrophils, monocytes, macrophages, and eosinophils traffic to sites of infection or inflammation and engulf microorganisms and apoptotic cells as the lead players in the innate immune response.

NEUTROPHILS

Neutrophils, also known as *granulocytes* because of their numerous cytoplasmic granules, are crucial for the host defense against bacteria and fungi. They are bone marrow–derived, terminally differentiated cells incapable of further cellular division. They have a short life span in the circulation ($t_{1/2} \approx 7$ hours), but survive an additional 1–2 days in tissue. In peripheral blood, they are normally maintained at 3000–6000 cells/mm^3 and represent 30–50% of the circulating leukocytes. There are four pools of neutrophils *in vivo: (i)* the bone marrow pool (≈90% of the total); *(ii)* the circulating pool (≈3% of the total); *(iii)* the marginated pool (adherent to the endothelium, ≈4% of the total); and *(iv)* those located in the tissues as extravasated or exudative neutrophils. About 55–60% of bone marrow is dedicated to the production of neutrophils, producing around 10^{11} cells daily, but this is upregulatable in times of stress.

Myeloid cell differentiation is a complex stepwise process that typically extends over 2 weeks in bone marrow. The pluripotent stem cell, the precursor for all hematopoiesis, develops into lineage-committed progenitors proceeding to terminally differentiated distinct cells, all the while preserving and regenerating more pluripotent stem cells.[1]

PRODUCTION OF MACROPHAGES AND GRANULOCYTES

The pluripotent stem cell gives rise to the myeloid stem cell from which the colony-forming unit granulocyte/erythrocyte/macrophage/megakaryocyte (CFU-GEMM) is derived. Among the growth factors that are influential at this step are stem cell factor (SCF), interleukin-3 (IL-3), and granulocyte macrophage–colony-stimulating factor (GM-CSF).[2] CFU-GEMM further differentiates into the colony-forming unit–granulocyte macrophage (CFU-GM) under the continuing influence of these growth factors. The colony-forming unit–granulocyte (CFU-G), a neutrophil lineage committed precursor, is derived from CFU-GM under the control of IL-3, GM-CSF, and granulocyte–colony-stimulating factor (G-CSF). The myeloblast is formed from the CFU-G under the influence of GM-CSF and G-CSF and is the first morphologically distinct cell of the neutrophil lineage. Promyelocyte, myelocyte, metamyelocyte, band form, and mature neutrophil formation follow consecutively under the ongoing control of G-CSF and GM-CSF. The maturation process from stem cell to the myelocyte stage takes 4–6 days and an additional 5–7 days for the myelocyte to form the mature neutrophil, all in bone marrow.

Macrophage differentiation is similar to granulocyte differentiation in many respects. CFU-GM differentiates into the colony-forming unit–macrophage (CFU-M) followed by the formation of the monoblast, promonocyte, and monocyte under the influence of macrophage colony-stimulating factor (M-CSF).[3] After monocytes are released into blood, they circulate for 1–4 days before entering tissues, where they further differentiate into macrophages.

EVOLUTION OF NEUTROPHIL GRANULES

During myelopoiesis in bone marrow, the first granules form at about the promyelocyte stage, stain blue with the Wright or Romanowsky stain, and are called *primary granules* or *azurophilic granules*. Their formation ceases at the myelocyte stage, and they are distributed among the daughter cells. These primary granules contain microbicidal enzymes, including defensins, hydrolases, and proteases (Table 22.1). As the granulocyte precursors mature and divide, the number of primary granules per cell decreases. After the promyelocyte stage, secondary or specific granules form. In the mature neutrophil, they comprise about two-thirds of the granules. The secondary granules are less dense and contain cytochrome b_{558}, lysozyme, lactoferrin, and collagenase. The gelatinase-containing tertiary granule probably forms after the metamyelocyte stage and can be detected in the band form and the mature granulocyte.

DISORDERS OF NEUTROPHIL PRODUCTION

Chronic neutropenia refers to conditions with absolute neutrophil counts (ANCs) of less than 500 cells/μL lasting more than 6 months. Chronic neutropenia can have many etiologies, as listed in Table 22.2.

Severe Congenital Neutropenia and Cyclic Neutropenia

Kostmann originally described extensive northern Swedish kindred with both recessive and dominant neutropenia, but subsequently sporadic cases were added, making this a confusing melange of neutropenia syndromes.[4] Severe congenital neutropenia (SCN) is now known to be a heterogeneous group of disorders that present similarly. The genes recognized as mendelian causes of

TABLE 22.1 Neutrophil Granule Components

Granule	Contents	Properties
Primary (azurophilic) granules	*Lysosomal hydrolases* Myeloperoxidase Defensins Lysozyme Elastase Cathepsin G Azurocidin Proteinase 3 Bacterial—permeability increasing protein (BPI) *Acid hydrolases* Cathepsin B Cathepsin D β-Glycerophosphatase granulocyte β-Glucuronidase *N*-acetyl-β-glucosaminidase α-Mannosidase Other Collagenase	• First formed during myelopoiesis at promyelocyte stage • Appear blue when stained with Wright stain • Least mobilizable granule • Measure ≈ 0.8 µm • Defensins constitute 30–50% of granule contents • Augment the microbial damage initiated by reactive oxidants • Help digest dead microbes and host cells • BPI neutralizes gram-negative bacteria
Secondary (specific) granules	*Lysosomal hydrolases* Lysozyme *Other* Collagenase Gelatinase Lactoferrin Vitamin B_{12}–binding proteins Cytochrome b_{558} Histaminidase FMLF receptors C3bi receptors	• Synthesis begins at the myelocyte stage • These granules are specific to phagocytes • Measure ≈0.5 µm • Binding proteins deprive microorganisms of nutrients • Most are positively charged, enhancing cell surface
Tertiary (smaller) granules	*Acid hydrolases* Cathepsin B Cathepsin D β-Glycerophosphatase granulocyte β-Glucuronidase *N*-acetyl-β-glucosaminidase α-Mannosidase *Other* Gelatinase	• Heterogeneous population of organelles, including C-particles and secretory vesicles • Detected in the band form and mature neutrophils

TABLE 22.2 Causes of Neutropenia

Classification	Etiology
Hematological	Kostmann syndrome Severe congenital neutropenia Cyclic neutropenia Myelodysplastic syndrome Aplastic anemia Leukemia
Immunological/inflammatory disorders	Severe combined immunodeficiency (SCID) Hyper-IgM syndrome (CD40L) Chediak-Higashi syndrome Cartilage–hair hypoplasia Reticular dysgenesis Dyskeratosis congenita Autoimmune neutropenia Isoimmune neutropenia
Infections	Human immunodeficiency virus (HIV) Parvovirus Epstein-Barr virus Malaria Cytomegalovirus
Inborn errors of metabolism/ nutritional disorders	Gaucher disease Glycogen storage disease, type Ib Transcobalamin deficiency Vitamin B_{12}, folate deficiency
Other	Schwachman-Bodian-Diamond syndrome Idiopathic neutropenia Chemotherapy Radiation therapy Drugs (*e.g.,* vancomycin, chloramphenicol, sulfamethoxazole, clozapine) Toxins (*e.g.,* benzene) Dialysis Reticuloendothelial sequestration

neutropenia are the neutrophil elastase (*ELANE* or *ELA2*), *HAX1*, *G6PC3*, *GFI1*, *GATA2*, *JAGN1*, *VPS45* genes, and activating mutations in the Wiskott–Aldrich syndrome (*WAS*) gene.

SCN is usually diagnosed in the neonatal period or early infancy because of the occurrence of life-threatening pyogenic infections, cellulitis, stomatitis, peritonitis, perirectal abscess, or meningitis. The most common bacteria isolated are *Staphylococcus aureus*, *Escherichia coli*, and *Pseudomonas aeruginosa.* Patients usually have an ANC below 200 cells/µL, mild anemia, and hypergammaglobulinemia, sometimes with eosinophilia and monocytosis. SCN represents an impairment of myeloid differentiation because of maturational arrest of neutrophil precursors at the level of promyelocytes or myelocytes in bone marrow. A subset of patients with SCN (7.5–10%) subsequently develops a myelodysplastic syndrome (MDS) or acute myeloid leukemia (AML), which has been associated with acquired truncation mutations of the G-CSF receptor (G-CSFR).

The majority of patients with SCN have heterozygous mutations in the neutrophil elastase (*ELANE*, formerly *ELA2*).[4] The clinically fascinating cyclic form of this disorder has oscillating neutrophil counts with 21-day cycles, hence the name cyclic neutropenia (CN). These mutations are transmitted as autosomal dominant mutations but also occur spontaneously. There is no clear genotype–phenotype correlation between specific *ELANE* mutations that lead to CN as opposed to SCN. These typically missense mutations lead to intracellular accumulation of mutant proteins that are inappropriately trafficked into azurophilic granules.[5] This mutated, aberrantly folded protein is thought to contribute to neutrophil precursor apoptosis and the clinical phenotype of neutropenia, but the mechanisms by which this occurs are still somewhat obscure. Treatment with subcutaneous G-CSF can typically increase the ANC above 1000 cells/µL, with a decrease in the frequency of infections and significant clinical improvement overall.[6] Patients with SCN who have received long-term G-CSF therapy are at an increased risk of developing AML or MDS, which correlates with overall G-CSF responsiveness.

Homozygous loss-of-function mutations in *HAX1* account for the majority of recessive cases of SCN,[7] some of which were in original pedigree described by Kostmann. Patients may have isolated SCN or associated neurological problems (cognitive impairment, developmental delay, or epilepsy), depending on which isoform of *HAX1* is mutated. Patients with mutations affecting isoform A have only SCN, as opposed to patients with mutations affecting both isoforms (A and B), who develop neurological problems in addition to SCN.[8]

Dominant zinc finger mutations disabling transcriptional repressor activity of the *growth factor independent 1* (*GFI1*) gene have been described in a few patients with SCN.[4] *GFI1* is a transcriptional repressor prooncogene controlling normal hematopoietic cell differentiation and also regulates *ELANE* as well as several of the CAAT enhancer binding proteins (C/EBP). Mutations in *GFI1* are also associated with aberrations in lymphoid and myeloid cells, leading to a circulating population of immature myeloid cells. *Gfi1* knock-out mice have impaired T-helper type 2 (Th2) regulation and B-cell, Th17, and dendritic-cell (DC) differentiation.

Mutations in the *glucose-6-phosphatase catalytic subunit 3 (G6PC3)* gene complex cause another form of SCN along with developmental and somatic problems.[9] *G6PC3* encodes glucose-6-phosphatase-β, which hydrolyzes glucose-6-phosphate (G6P) in the final step of gluconeogenesis and glycogenolysis. It is coupled to a glucose transporter (G6PT) that facilitates G6P transport from the cytoplasm to the endoplasmic reticulum. Mutations in the *G6PT* gene lead to glycogen storage disease type Ib, which has variable neutropenia and infections and other complications, such as liver adenomas, growth retardation, osteoporosis, polycystic ovaries, and inflammatory bowel disease (IBD). Children with these complications have increased susceptibility to bacterial infections and cardiovascular abnormalities, including prominent ectatic superficial veins.

Shwachman-Bodian-Diamond Syndrome

Shwachman-Bodian-Diamond syndrome (SBDS) was first described in 1964 as a disorder with pancreatic exocrine insufficiency and bone marrow dysfunction. Currently, it is recognized as the second most common cause of inherited exocrine pancreatic insufficiency after cystic fibrosis. It is autosomal recessive (located at 7q11) with an estimated incidence of 0.5–1/100 000.[10] The SDBS protein belongs to a highly conserved protein family involved in RNA metabolism. Mutations cause defects in the development of the exocrine pancreas, hematopoiesis, and chondrogenesis. Recurring mutations result from gene conversion caused by recombination with a pseudogene in 89% of unrelated patients; 60% carry two converted alleles. (Pseudogene conversion is also the cause of the majority of cases of $p47^{phox}$-deficient chronic granulomatous disease [CGD].)

Patients present with recurrent infections, failure to thrive, hematopoietic dysfunction, metaphyseal dysostosis, growth retardation, and fatty replacement of the pancreas. Most patients have mild neutropenia, and a few have neutrophil counts below 500 cells/μL, which can be intermittent or chronic.[11] Anemia and thrombocytopenia are associated with neutropenia. Congenital aplastic anemia is an unusual presentation. Upper and lower respiratory tract pyogenic infections are common and related to neutropenia. Short ribs with broadened anterior ends are common radiological findings, along with metaphyseal dyschondroplasia of the femoral head. The diagnosis is suggested by neutropenia, radiological findings, and abnormal pancreatic exocrine function. It is confirmed by gene sequencing.

Autoimmune Neutropenia

Autoimmune neutropenia (AIN) is caused by peripheral destruction of neutrophils as a result of granulocyte-specific autoantibodies.[12]

Primary Autoimmune Neutropenia

Primary AIN is seen in infancy unassociated with other systemic immune-mediated disorders and is the most common form of neutropenia, equally affecting boys and girls at around 1/100 000.[12] The average age at diagnosis is 8 months. The majority present with mild skin and upper respiratory tract infections; some patients remain asymptomatic despite low neutrophil counts. The majority of patients have a neutrophil count greater than 500 cells/μL at the time of diagnosis, but the ANCs may transiently increase two- to threefold during severe infection. Bone marrow shows normal to increased cellularity. Myeloid precursors typically reach the myelocyte/metamyelocyte stage. Phagocytosed granulocytes in bone marrow may indicate removal of sensitized granulocytes there. Granulocyte-specific antibodies are detected by direct granulocyte immunofluorescence testing (D-GIFT), the vast majority of which are immunoglobulin G (IgG) against glycoproteins of the granulocyte membrane designated neutrophil antigens (NAs). NAs are located on IgG receptor IIA or IIIB (FcγRIIa and FcγRIIIb).

AIN is generally self-limiting. Disappearance of the antibodies from the circulation precedes normalization of neutrophil counts. Prophylactic antibiotic treatment should be reserved for those with recurrent infections. Alternative treatment strategies for severe infections and in the setting of emergency surgical interventions include high-dose intravenous immunoglobulin (IVIG), corticosteroids, and G-CSF, with the latter being the most effective at increasing the ANC.

Secondary Autoimmune Neutropenia

Secondary AIN can be seen at any age and has a more variable clinical course. Hepatitis, systemic lupus erythematosus (SLE), or Hodgkin disease may underlie it and cause other autoimmune problems as well. These antineutrophil antibodies (ANAs) have pan-FcγRIII specificity. CD18/CD11b antibodies have been detected in a subset of patients with secondary AIN. This neutropenia responds poorly to most therapies.

Alloimmune Neonatal Neutropenia

First described by Lalezari in 1966, alloimmune neonatal neutropenia (ANN) is caused by the transplacental transfer of maternal antibodies against the fetal neutrophil antigens NA1, NA2, and NB1, leading to immune destruction of neonatal neutrophils.[13] These complement-activating antineutrophil IgG antibodies can be detected in about 1/500 live births. Antibody coated neutrophils in ANN are phagocytosed by the reticuloendothelial system and removed from circulation, leaving the neutropenic neonate at risk for infections. Omphalitis, cellulitis, and pneumonia typically occur within the first 2 weeks of life in association with neutropenia. Diagnosis can be made by detection of neutrophil-specific alloantibodies in the maternal serum. ANN responds to G-CSF or high-dose IVIG, but most patients improve without specific treatment in a few weeks to 6 months with waning of maternal antibody.

DEFECTS OF LEUKOCYTE ADHESION

Migration of circulating leukocytes from the bloodstream into tissues depends on complex bidirectional interactions between leukocytes and endothelial cells (Chapter 11). The initial steps involve the activation of circulating leukocytes by signal molecules released from inflamed tissues or from the bacteria themselves.

After activation by chemotactic factors, such as the complement fragment C5a, IL-8, leukotriene B4 (LTB4), or the bacterial product formyl-methionyl-leucyl-phenylalanine (fMLF), leukocytes rapidly become adhesive to the endothelium, other leukocytes, or laboratory surfaces. The activation process involves translocation of subcellular granules containing adhesion molecules (CD18/CD11b) to the polymorphonuclear leukocyte (PMN) surface and qualitative alterations in the adhesion molecules constitutively expressed on the plasma membrane. Adhesion and transmigration of leukocytes occur as a result of interactions between three groups of molecules: leukocyte integrins, endothelial intercellular adhesion molecules (ICAMs, members of the immunoglobulin supergene family), and glycosaminoglycans or selectins (Fig. 22.1). The first step in targeting PMNs to inflamed tissues is the rolling or tethering of PMNs on the endothelium of postcapillary venules.[14] This is attributed to the interactions between CD15s (sialyl Lewis^{X} or SLe^{X}) expressed on the leukocyte surface and P-selectin or E-selectin—members of the selectin family of adhesion molecules—expressed on the vascular endothelium. In addition, L-selectin on the leukocyte surface interacts with its counterligands P-selectin, CD34, glyCAM-1, and other glycoproteins located on the endothelial surface. *Rolling*, a relatively low affinity interaction mediated by selectins, is followed by *firm adhesion*, which is a high affinity interaction between integrins on the neutrophil and ICAMs on the endothelium. Adhesion is followed by the transmigration of neutrophils between endothelial cells out to the extracellular matrix (ECM).

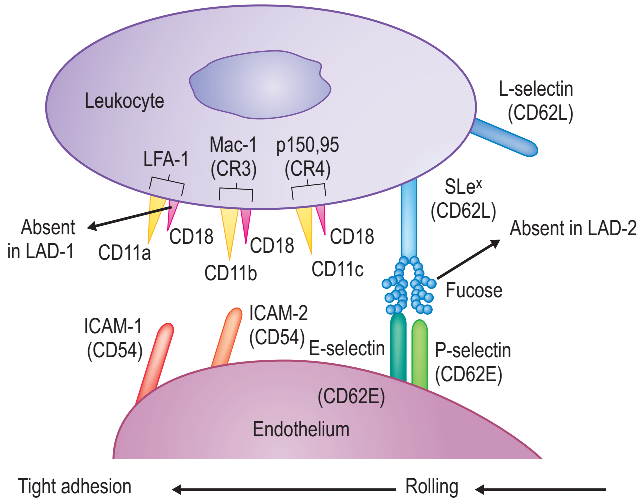


FIG 22.1 Leukocyte Adhesion to Nonlymphoid Endothelium. Selectins (L-selectin/CD62L, P-selectin/CD62P, and E-selectin/CD62E), integrins (CD18/CD11a or LFA-1, CD18/CD11b or Mac-1, and CD18/CD11c or p150,95), and intercellular adhesion molecules (ICAMs) are involved in leukocyte adhesion to the nonlymphoid endothelium. Rolling, the initial step of leukocyte adhesion, is mediated by the interactions of E-selectin and P-selectin on endothelial surfaces with the sialyl Lewis^{x} (SLe^{x} or CD15s) of leukocytes as well as L-selectin on the leukocyte surfaces with its counterligands CD34 or glyCAM-1. This low-affinity tethering or rolling facilitates tight adhesion as a result of the interactions of leukocyte function–associated antigen-1 (LFA-1) with ICAM-1 or ICAM-2 and Mac-1 with ICAM-2. CD18 is the molecule that is missing or dysfunctional in leukocyte adhesion defect-1 (LAD-1); SLe^{x} is the missing molecule in LAD-2.

Leukocyte Adhesion Defect-1

In the 1970s, infants and children were recognized with severe, recurrent life-threatening bacterial infections affecting the skin, gingiva, and lungs. A common clinical feature was delayed separation of the umbilical stump with severe omphalitis. These patients were shown to have defects in membrane expression of the leukocyte adhesion glycoproteins of the integrin superfamily.[15,16]

Integrins are noncovalently associated, heterodimeric cell surface receptors, comprising one α subunit (CD11a, CD11b, or CD11c) and a common β chain (CD18), which is required for surface expression of the CD11 chains. These proteins mediate leukocyte adhesion to the endothelium and other leukocytes. Leukocyte adhesion defect-1 (LAD-1) results from mutations in the CD18 gene (*ITGB2*), located on chromosome 21q22. Patients with LAD-1 have defective polymorphonuclear cell adherence, leading to defective chemotaxis and trafficking, as well as low natural killer (NK) and cytotoxic T-lymphocyte (CTL) activity. The absence of CR3 leads to loss of complement-mediated phagocytosis and bacterial killing. LAD-1 is often manifested by delayed umbilical cord separation, omphalitis, persistent leukocytosis, destructive periodontitis, and recurrent infections with *S. aureus*, *Pseudomonas aeruginosa*, and *Klebsiella* spp. Patients with some residual CD18 expression and function (*i.e.*, hypomorphic mutations) live beyond childhood with less frequent or severe infections and do not typically have delayed umbilical cord separation. Persistent neutrophil leukocytosis (usually >15 000 cells/μL) in the absence of infection is common in all patients, driven by both low-level ongoing infection and impaired exit of neutrophils from the circulation. Oral ulcers, severe periodontitis, gingivitis with apical bone loss (Fig. 22.2), and eventual loss of permanent teeth are major problems in LAD-1 and reflect excessive IL-17 expression by CD4 T cells as a result of uninhibited IL-23 production by tissue macrophages.[17] Necrotizing cutaneous ulcers with delayed wound healing and lingering eschar formation are common (Fig. 22.3). Defective chemotaxis and adhesion mean that leukocytes fail to migrate to sites of infection, accounting for the inability to form pus and erythema at the site of infection. Biopsies of the ulcers

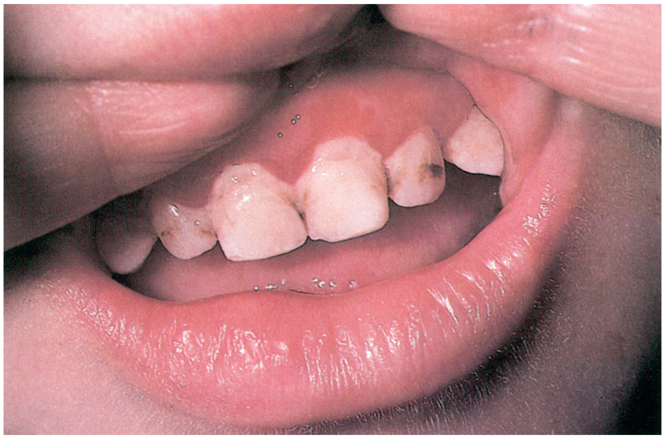


FIG 22.2 Oral Pathology in a Patient With Leukocyte Adhesion Defect-1 (LAD-1). Gingivitis and severe periodontitis are hallmarks of LAD-1.

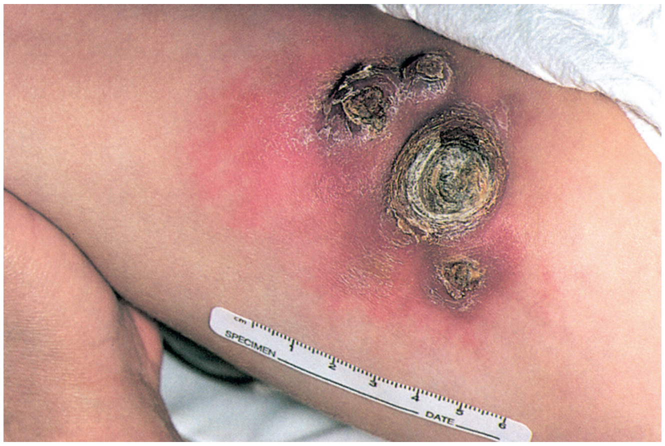


FIG 22.3 **Skin Infection of a Patient With Leukocyte Adhesion Defect-1 (LAD-1).** Failure to form pus, inability to demarcate the fibrotic skin debris, and limited inflammation.

characteristically show poorly formed granulation tissue and scant fibrinous exudate without neutrophils. Ulcerative gastrointestinal (GI) disorders resembling idiopathic IBD are also recognized in LAD-1, especially as patients grow older.

Although most cases of CD18 deficiency are homozygous, compound heterozygotes also occur.[18] The diagnosis is usually made by flow cytometric analysis of neutrophils showing decreased or absent CD18 and its associated heterodimers: CD11a, CD11b, and CD11c, and confirmed by mutational analysis of *ITGB2*. More subtle phenotypes can be detected by testing for mobilization of CD18 complexes, such as CD18/CD11b from neutrophils upon cellular stimulation. Definitive therapy of LAD-1 is bone marrow transplantation. Infections must be managed aggressively, since inflammatory responses and clinical signs are unreliable in these patients with profoundly impaired innate immune responses. Surgery is often essential for debridement of nonhealing ulcers, which frequently need tissue grafts, but immunomodulation of the affected cytokine pathways may also be helpful. Although not correcting the underlying gene defect, but reflecting the excessive production of IL-23 and IL-17 at sites of inflammation that is associated with severe oral ulcers, periodontitis and bone loss, a patient with moderate LAD-1 (34% of control CD18 activity) was treated with ustekinumab, which blocks IL-23-dependent production of IL-17. After one year of treatment the patient's severe chronic periodontitis and a deep sacral ulcer had resolved without serious infections or adverse reaction. This approach to therapy offers additional insight into the complex pathophysiology of LAD-1 inflammation and suggests the possibility of novel therapeutic approaches to symptomatic management.[18a]

KEY CONCEPTS

Leukocyte Adhesion Defect (LAD)

- Three types of adhesion defect are known: LAD-1, -2, and -3. There are two phenotypes for LAD-1: moderate and severe.
- LAD-1 results from mutations in CD18; LAD-2 is caused by mutations in sialyl Lewis^X (CD15s); LAD-3 results from mutations in *FERMT3*.
- High white blood cell count, delayed umbilical cord separation, recurrent bacterial infections, skin ulcers, defective wound healing, gingivitis, and periodontitis are the hallmarks of LAD-1.

Leukocyte Adhesion Defect-2

A distinct defect of leukocyte adhesion with susceptibility to infection was described by Etzioni et al. in 1992 and named LAD-2.[15,19] It is characterized by growth retardation and mental retardation, hypotonia, seizures, dysmorphic features, strabismus, and persistent periodontitis. In contrast to LAD-1, wound healing is not impaired, nor is the susceptibility to bacterial infections as severe. Hypofucosylation of the protein CD15s (sialyl Lewis^X, SLe^X) on neutrophils impairs the rolling step of neutrophil adhesion. The underlying defect is in guanosine diphosphate (GDP)–fucose biosynthesis, resulting from mutations in the GDP–fucose transporter-1 (*FUCT1* or *SLC35C1*), hence the designation of this disease as a congenital disorder of glycosylation IIc (CDGIIc). In addition to severe impairment in neutrophil migration as in LAD-1, lymphocyte homing to skin is also defective. Patients with LAD-2 have had relatively mild courses of infections with several pneumonias and superinfection of varicella lesions,[19] and some have reportedly improved with fucose supplementation. In addition to SLe^X, fucosylated blood group antigens are also affected, leading to the Bombay blood group phenotype (lack of the H antigen) and Lewis a^-b^- in these patients. Absence of CD15 on patient neutrophils can be detected by flow cytometry. Effective and prompt treatment of infections is central to the management of LAD-2.

Leukocyte Adhesion Defect-3

Following the elucidation of LAD-1 and LAD-2, another leukocyte adhesion deficiency was recognized, initially named LAD-1/variant (LAD-1v), now LAD-3. It had a distinct infantile bleeding diathesis similar to Glanzmann-type thrombasthenia along with defective leukocyte adhesion.[20] Although CD18/CD11a (lymphocyte function–associated antigen-1 [LFA-1] or $\alpha_1\beta_2$) is the main integrin on leukocytes, $\alpha_{IIb}\beta_3$ (also called *GPIIb-IIIa*) allows platelets to bind fibrinogen to promote clotting. This was initially described in Turkish patients and ascribed to *FERMT3*, which encodes KINDLIN3, an adaptor protein expressed in hematopoietic cells that regulates integrin activation.[21,22] KINDLIN3 activates integrins through binding to distinct motifs on the short tails of the integrin β subunits. Phenotypically, leukocytes and platelets in LAD-3 have defective β_3, β_2, and β_1 integrin activation as a result of loss of "inside out" or chemokine-mediated LFA-1 activation and intrinsic LFA-1/$\alpha_1\beta_2$ adhesiveness. In addition, these cells have decreased adherence to endothelial cells and reduced expression of the Rap-1 guanine nucleotide exchange factor, CalDAG-GEFI (CDGI). Based on the location and severity of mutation, LAD-3 leukocytes may also display loss of adhesion to vascular cell adhesion molecule-1 (VCAM-1). LAD-3 platelets have decreased binding to soluble fibrinogen, do not respond properly to thrombin via thrombin receptors (PARs), and therefore have poor platelet granule secretion through integrin activation. Bone marrow transplantation is necessary and curative.

CHRONIC GRANULOMATOUS DISEASE

CGD characterized by recurrent infections and hypergammaglobulinemia was first described in 1954. CGD results from defective phagocyte superoxide production leading to impaired microbial killing. It comprises five genotypes with a common phenotype (Table 22.3), characterized by recurrent severe bacterial and fungal infections and tissue granuloma formation.[23] CGD occurs at around 0.5–1/100 000 births. It is inherited in X-linked

TABLE 22.3 Genotype–Phenotype Correlations in X-Linked Chronic Granulomatous Disease (CGD)

	X91^0	X91$^-$	X91$^+$
gp91^{phox} protein levels	Undetectable	Normal to low	Normal
Residual superoxide production	Undetectable	Undetectable	Low
Cytochrome b$_{558}$ spectrum	Absent	Low	Low or normal
Type of mutations in *CYBB*	Deletions, insertions, splice site mutations, missense mutations, nonsense mutations	Missense mutation, especially involving amino acids 310–587	Missense mutations, especially involving amino acids 1–309

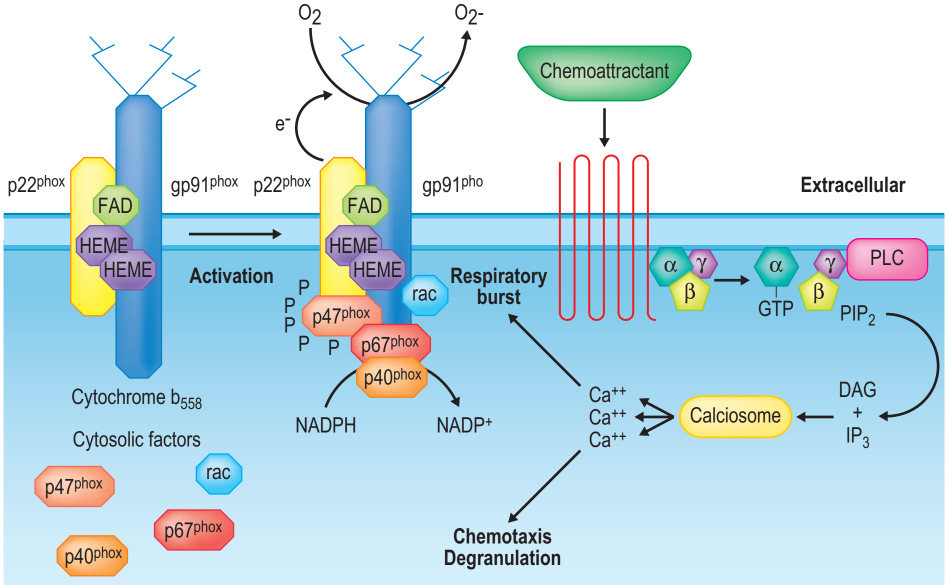


FIG 22.4 **Schematic Representation of the Nicotinamide Adenine Dinucleotide Phosphate (NADPH) Oxidase System.** Chemoattractants interact with their receptors on the neutrophil surface, leading to an increase in intracellular calcium concentration. This activation results in the assembly of the NADPH oxidase complex following phosphorylation of cytosolic factors. This, in turn, leads to superoxide production. DAG, diacylglycerol; PIP$_2$, phosphatidylinositol bisphosphate; IP$_3$, inositol triphosphate; α, β, γ, subunits of the guanosine triphosphate (GTP)–coupled receptors.

and autosomal recessive patterns, with the relative frequencies of recessive disease depending on the rates of local consanguinity. In the United States, the X-linked form accounts for about 65% of cases and the autosomal recessive p47^{phox} (phagocyte oxidase) deficiency for about 25%.

Patients with CGD often present with pneumonia, liver abscess, skin infections, lymphadenitis, or osteomyelitis; bacteremia is relatively uncommon. Initial presentation with IBD is not unusual. Exuberant tissue granuloma formation at the sites of infection, at surgical wounds, and in hollow viscera is a frequent problem seen more often in patients with the X-linked form of CGD.

The NADPH Oxidase and Its Activity

The nicotinamide adenine dinucleotide phosphate (NADPH) oxidase is a multicomponent system that transfers an electron to molecular oxygen by way of FAD and heme to form superoxide (O_2^-) (Fig. 22.4). Cytochrome b$_{558}$ is a membrane-bound heterodimer lodged in the wall of the secondary granules; the large glycosylated β subunit is gp91^{phox} and the small nonglycosylated α subunit is p22^{phox}. The cytoplasmic tail of gp91^{phox} binds FAD, heme, and NADPH, which are required for electron transfer to oxygen (O_2). Neutrophil stimulation leads to aggregation and phosphorylation of p47^{phox}, p67^{phox}, p40^{phox}, and the small guanosine triphosphate (GTP)–binding proteins RAC1/RAC2, which dock with the cytochrome at the membrane through binding of p47^{phox} and p22^{phox}.

Segal et al. showed that much of the killing effect of neutrophils is, in fact, carried out by proteases, enhanced by NADPH oxidase activity.[24] Charge created by electron flux across the membrane is compensated mostly by K^+ flux, which enhances microbial killing. Papayannopoulos and Zychlinksy identified neutrophil extracellular traps (NETs; extruded DNA with attached antimicrobial peptides), which depend on superoxide generation and are deficient in CGD.[25]

Mutations Leading to CGD

X-Linked CGD

The most common form of CGD is caused by mutations in *CYBB*, which encodes gp91^{phox} (located at Xp21.1) (see Table 22.3). Mutations include deletions (22.2%), insertions (7%), deletion/insertions (1.5%), nonsense (29.8%), missense (19.4%), splice sites (19.5%), and promoters (0.6%).[26] The spontaneous mutation rate is approximately 11%. With large interstitial deletions adjacent genes may be deleted as well, leading to complex phenotypes. Telomeric deletions cause McLeod syndrome (*KX*, or Kell antigen deletion), Duchenne muscular dystrophy (*DMD*), and X-linked retinitis pigmentosa (*RPGR*) along with CGD.[27] McLeod syndrome includes absent erythrocyte Kx protein and diminished levels of Kell blood group antigens. Patients may eventually develop progressive neurodegenerative symptoms, such as areflexia, dystonia, and choreiform movements. In patients with McLeod syndrome, anti-Kx and anti-Km antibodies are formed when transfusions are given, making future transfusions extremely difficult. Deletions centromeric from *CYBB* may cause ornithine decarboxylase deficiency along with CGD.[28]

Autosomal Recessive CGD

Mutations in p47*phox* (*NCF1*, located at 7q11.23) cause the majority of the recessive cases of CGD, around 25%, usually caused by homozygous deletions of the canonical GT splice site at the start of exon 2.[29] p22*phox* (*CYBA*, located at 16q24) and p67*phox* (*NCF2*, located at 1q25) are responsible for less than 5% of CGD cases each. p40*phox* (*NCF4*, located at 22q13.1) deficiency has been reported in a boy with early-onset severe granulomatous fistulizing colitis without a significant infectious phenotype.[30] No autosomal dominant cases of CGD have been identified.

Clinical Manifestations of CGD

The first severe infection usually occurs in infancy or childhood but can also occur in adulthood. Later diagnoses usually are seen in patients with residual superoxide production, either hypomorphic gp91*phox* or p47*phox* deficiency.[23,31] The constellation of signs and symptoms that suggest CGD range from failure to thrive, to IBD, to visceral abscesses, to recurrent sinopulmonary infections, to characteristic infections. Infections are most commonly pneumonia, lymphadenitis, liver abscess, skin abscess, perianal abscess, and osteomyelitis. As in other neutrophil defects, the most common pathogen is *S. aureus.* Characteristic infections are caused by catalase-positive organisms such as *S. aureus*, *Burkholderia cepacia* complex, *Serratia marcescens*, *Nocardia* spp., and *Aspergillus* spp. (Fig. 22.5).

Staphylococcal liver abscesses in CGD are dense and necrotic and cause significant morbidity. Their fibrocaseous consistency means that percutaneous drainage was rarely successful, and open surgery was required. However, combined steroid and antibiotic therapy of CGD liver abscess has now become the preferred approach in many centers.[32] Liver involvement leading to portal hypertension is the likely cause of the splenomegaly commonly seen in CGD and is also closely tied to mortality.[33]

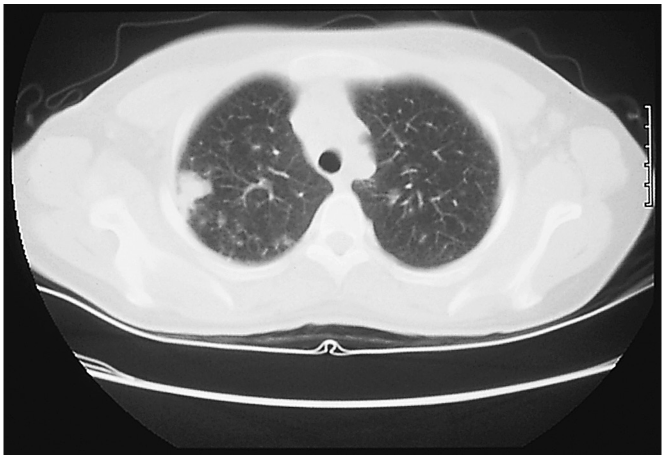


FIG 22.5 **Computed Tomography (CT) Scan of the Lungs of a Patient With Chronic Granulomatous Disease (CGD) and *Aspergillus* Pneumonia.** *Aspergillus* pneumonia is often a peripheral consolidation in the lung parenchyma.

Pulmonary aspergillosis remains a major cause of death in CGD.[23] *Aspergillus fumigatus* is the most commonly isolated fungus, but it is now easily and successfully treated with azole antifungals. In contrast, *A. nidulans* and *A. viridinutans*, species with low pathogenicity in the normal host, cause severe disease in CGD.[34,35] Surgical resection of these infections is often required. *Aspergillus* infections in CGD are often unaccompanied by fever and leukocytosis.[34] Acute diffuse pulmonary fungal infection in CGD is referred to as "mulch pneumonitis," characterized by fever, hypoxia, and diffuse pulmonary infiltrates caused by inhalation of fungi, typically during mulching, leaf raking, or gardening.[36] This syndrome can be the initial presentation of CGD in older children and adults and is important to recognize, since it best responds to a combination of antifungals and steroids.

Septicemia is relatively uncommon but may occur with *B. cepacia* complex and *Chromobacterium violaceum*. *Granulibacter bethesdensis* is a pathognomonic gram-negative rod that causes chronic necrotic lymph node and spleen involvement in CGD.[37]

Inflammatory granuloma formation is one of the hallmarks of CGD. Pyloric outlet obstruction, bladder outlet obstruction, and ureteral obstruction are common. Crohn-like IBD affects between 30–50% of patients, predominantly X-linked, and may involve the esophagus (Fig. 22.6), jejunum, ileum, cecum, rectum, and perirectal area.[38] GI manifestations can include diarrhea,

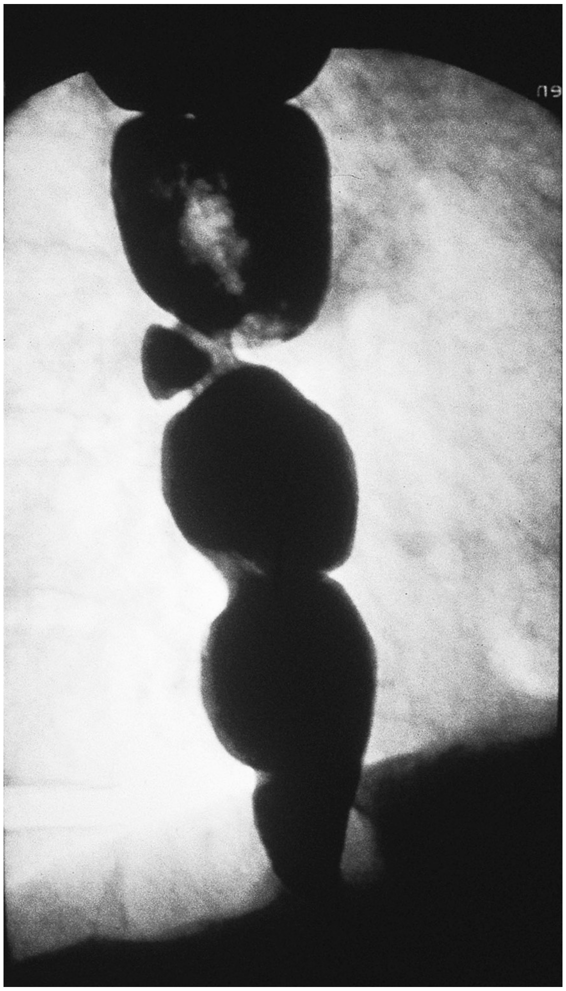


FIG 22.6 **Esophageal Involvement in Chronic Granulomatous Disease (CGD).** Esophageal strictures caused by granuloma formation as shown by barium swallow.

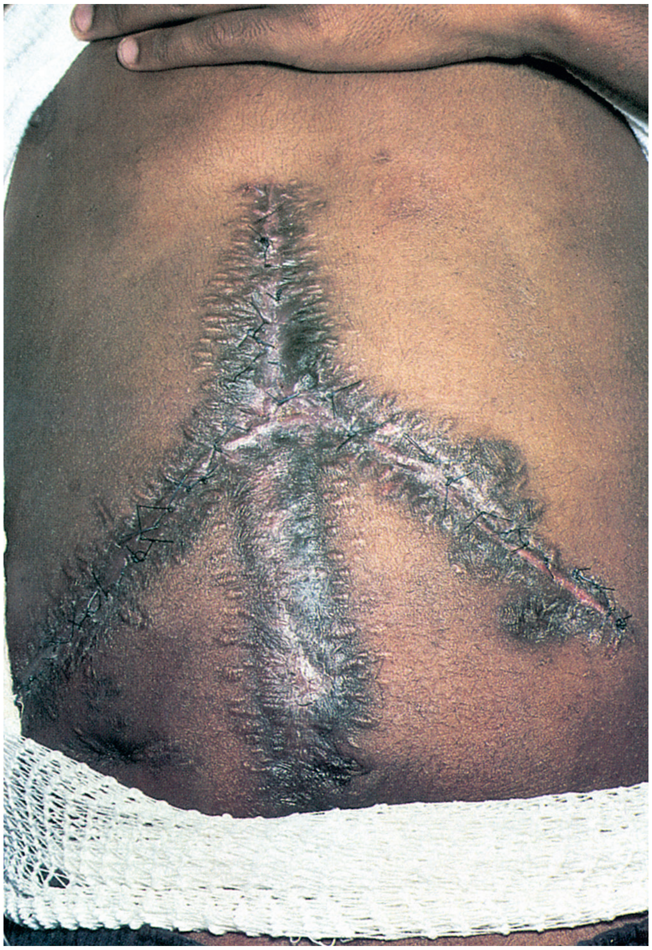


FIG 22.7 **Exuberant Granuloma Formation in Chronic Granulomatous Disease (CGD).** Wound dehiscence and impaired wound healing at surgical incision sites as a result of dysregulated inflammatory responses in a patient with X-linked CGD.

malabsorption, abdominal pain, growth delay, and hypoalbuminemia. The median age of initial GI manifestations is 5 years, and abdominal pain is common. Interestingly, GI involvement has no effect on mortality, is not associated with liver disease, and is unaffected by the use of interferon-γ (IFN-γ).[31,33,38]

Granulomata respond very well to steroids and often require a slow taper over several weeks to months. Exuberant formation of granulation tissue and dysregulated cutaneous inflammatory responses lead to wound dehiscence and impaired wound healing (Fig. 22.7). Autoimmune and rheumatological problems have been reported at higher than normal levels in patients with CGD.[39]

A comprehensive study of 287 patients with CGD from 244 kindred correlated the production of reactive oxygen intermediates with survival.[31] Patients with residual superoxide production had significantly better long-term survival compared with patients without residual superoxide production. Confirming the importance of this association, there was a direct correlation between the degree of superoxide and survival. Consistent with their previously recognized milder disease and better survival, patients with mutations in p47^{phox} had significant residual superoxide production. For those with gp91^{phox} mutations, the findings were more surprising. Patients with X-linked CGD with residual superoxide production were those with missense or splice mutations in the first 309 amino acids of gp91^{phox}. Those with missense mutations involving amino acids 310–587 had no residual superoxide production, regardless of protein levels. Therefore identification of the specific molecular subtype of CGD and specific mutation has important implications for severity and survival. Interestingly, mortality curves did not diverge until after age 20 years, suggesting that residual superoxide production determines later toxicities, such as liver dysfunction, not early childhood mortality from infection. It is critical to keep in mind that this comprehensive study included data from patients followed for up to 30 years—that is, a significant number of patients were born before the advent of modern antimicrobials. Therefore survival of a child born in the current age and who receives ideal management may well exceed that in the reported population.

CLINICAL PEARLS

Chronic Granulomatous Disease (CGD)

- CGD comprises a group of five inherited disorders with a common phenotype.
- Infections with catalase-positive bacteria and fungi and granuloma formation in the gastrointestinal and urinary tract are major problems in CGD.
- Oral prophylactic antibiotics and subcutaneous interferon (IFN)-γ injections three times a week are currently recommended for CGD.
- Diagnosis can be made via nitroblue tetrazolium (NBT) test or dihydrorhodamine (DHR) assay, the latter being a more sensitive diagnostic tool.
- Bone marrow transplantation is highly effective and curative.

Diagnosis of CGD

The diagnosis of CGD is most easily established by the dihydrorhodamine (DHR) assay, which measures the hydrogen peroxide–dependent conversion of DHR 123 to rhodamine 123, which is accompanied by fluorescence. This test is relatively reproducible and can be quantized on a flow cytometer allowing the determination of a DHR index, which correlates with residual superoxide production capacity. Other assays include NBTR and dichlorofluorescein (DCF), but these are somewhat more complicated or more prone to reader effects (Fig. 22.8). One important false positive to keep in mind in DHR testing is myeloperoxidase (MPO) deficiency. MPO deficiency gives a DHR result consistent with CGD, but when superoxide production is measured by NBTR or the more specific ferricytochrome c reduction, it is normal to increased.

Treatment of CGD

Prophylactic trimethoprim–sulfamethoxazole (TMP-SMX) significantly reduces the frequency of bacterial infections in CGD, especially those caused by *S. aureus*. TMP-SMX prophylaxis is ineffective against fungal infections but does not encourage them. Prophylactic itraconazole prevents fungal infections. IFN-γ is beneficial as a prophylactic treatment in CGD. In a multicenter, placebo-controlled trial of IFN-γ, the number and severity of infections were significantly reduced by IFN-γ. The exact mechanism of action of IFN-γ is not known, but it has multiple effects, including stimulation of components of NADPH oxidase in partial deficiencies, increased bactericidal activity through neutrophil granule components, and Fc receptor expression. Subcutaneous administration of recombinant IFN-γ three times

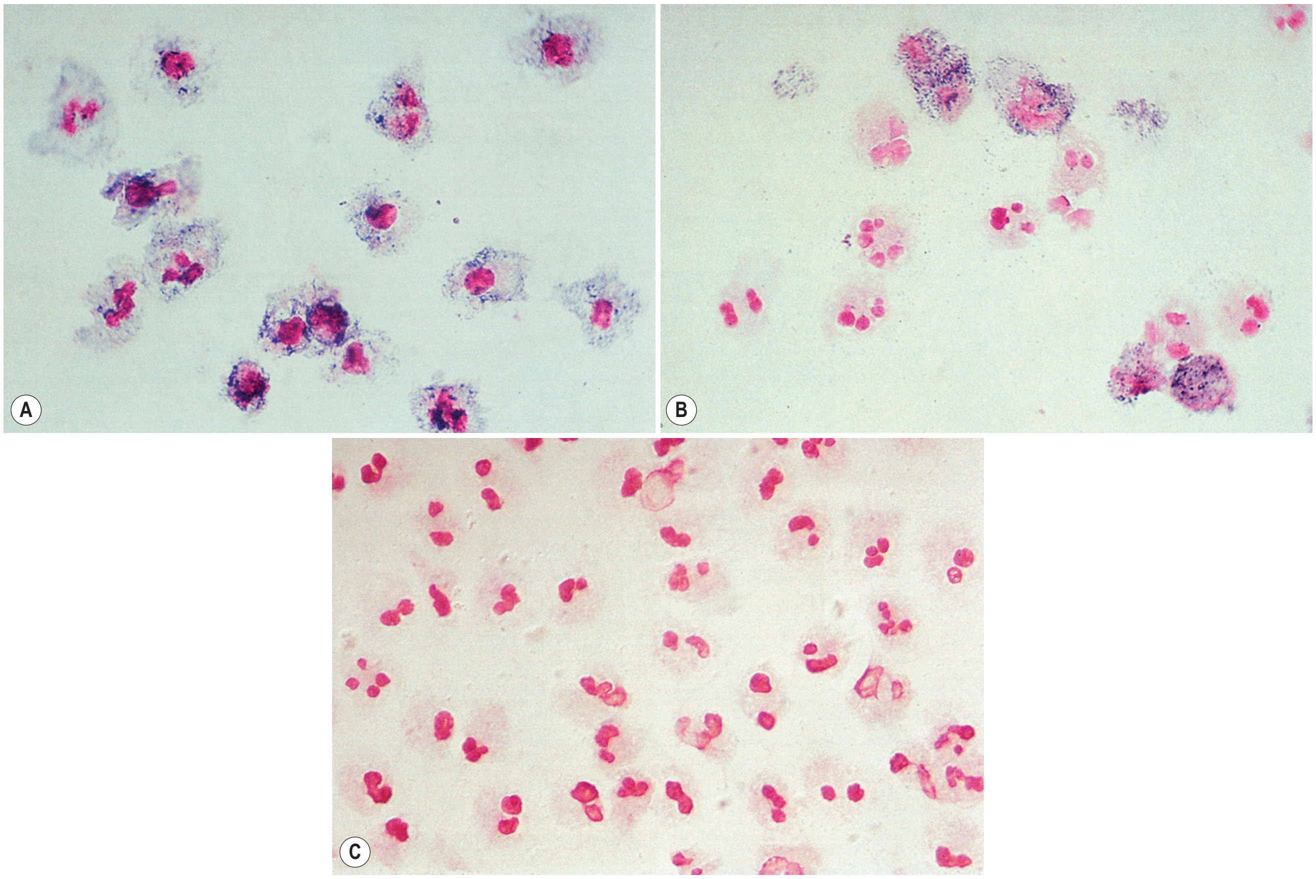


FIG 22.8 **Laboratory Diagnosis of Chronic Granulomatous Disease (CGD) with the Nitroblue Tetrazolium (NBT) Test.** (A) Nitroblue tetrazolium reduction (NBTR) reduction by purified normal neutrophils following stimulation with phorbol esters and calcium ionophore. NBT is reduced by all neutrophils, showing a blue/purple deposit. (B) NBTR by purified neutrophils from an X-linked CGD carrier; two different populations of cells are seen. Normal (unaffected cells) reduce the NBT dye and stain blue/purple, whereas affected cells fail to reduce the NBT dye and appear clear. (C) Neutrophils from a patient with CGD fail to reduce the NBT dye and appear clear. (Courtesy of Dr. Douglas B. Kuhns, Leidos, Frederick, MD.)

a week at a dose of 50 μg/m^2 (for those with body surface area >0.5 m^2) is recommended. The adverse effects of recombinant IFN-γ in patients with CGD have been limited to fever, chills, headache, flu-like illness, and diarrhea. During severe infections, leukocyte transfusions are sometimes used in addition to antibiotics, but this approach may lead to alloimmunization, compromising future bone marrow transplantation opportunities.

Because CGD is predominantly a hematopoietic disorder, bone marrow transplantation can cure CGD, even in the setting of active infection.[40] The type of transplant that is used in patients with CGD varies among centers, but both fully myeloablative and partially myeloablative transplants (reduced intensity conditioning) have been effective. Although active infection is a relative contraindication for bone marrow transplantation overall, there are certain infections in CGD, especially those caused by atypical *Aspergillus* spp. infections, that are not curable with standard antifungal therapy. Bone marrow transplantation prevents not only recurrent life-threatening infections, but also GI disease and growth retardation and is currently successful in about 90% of cases.

MYELOPEROXIDASE DEFICIENCY

MPO is a heme-containing enzyme necessary for the conversion of hydrogen peroxide (H_2O_2) to hypochlorous acid (HOCl). MPO is expressed early in myeloid differentiation and resides in the azurophilic granules of neutrophils and the lysosomes of monocytes.[41] Mature MPO is a symmetrical molecule of four peptides, with each half consisting of a heavy–light chain heterodimer. Neutrophils of individuals with MPO deficiency fail to produce HOCl upon stimulation, whereas the NADPH oxidase system remains unaffected. Prolonged supranormal levels of superoxide and H_2O_2 production follow stimulation in MPO-deficient neutrophils. This may result from lack of negative feedback regulation of HOCl on the NADPH oxidase, although the exact mechanism is unknown. MPO deficiency can be primary (congenital) or secondary (acquired).

Primary MPO Deficiency

Primary MPO deficiency is the most common phagocyte defect with a frequency of 1/4000. Both total and partial MPO deficiencies have been described. Patients with primary MPO deficiency do not usually have increased infections, probably because MPO-independent mechanisms compensate for the lack of MPO-dependent microbicidal activity. Visceral candidiasis occurring with concurrent diabetes has been reported in some patients. However, the frequency of such cases is very low. Affected individuals may develop nonfungal infections, malignancies, and certain skin disorders. In several cohorts of patients with complete MPO deficiency, an increased incidence of solid or hematological tumors has been observed.[40] MPO-deficient neutrophils have no apparent defect in the phagocytosis of bacteria or fungi, but microbicidal activity is slower than normal. MPO-deficient neutrophils are severely impaired in killing *Candida* spp. or *Aspergillus* spp. *in vitro* despite the fact that most patients with MPO deficiency do not develop significant fungal infections. This suggests that the mucosal barrier to fungal infection is independent of MPO activity and is able to prevent invasive infection.

The most common mutation is a missense replacement of arginine 569 with tryptophan (R569W), causing maturational arrest of the MPO precursor and preventing heme incorporation. Most patients are compound heterozygotes. The diagnosis of MPO deficiency is made by using anti-MPO monoclonal antibodies (mAbs) in flow cytometric analysis of neutrophils. No MPO expression is seen in congenital deficiency, whereas near-normal antigenic reactivity may be seen with the acquired form.[41] Maintenance antibiotic or antifungal therapy is not routinely recommended. Prompt and prolonged therapy is advised in patients with diabetes mellitus and congenital MPO deficiency, as they may develop localized or systemic infections.

Secondary or Acquired MPO Deficiency

In the majority of patients, MPO deficiency is partial and transient. Secondary MPO deficiency occurs under certain clinical conditions, such as some hematological malignancies or disseminated cancers, exposure to cytotoxic agents or antiinflammatory medications, iron deficiency, lead intoxication, thrombotic diseases, renal transplantation, and pregnancy. MPO activity in bone marrow myeloid precursors as well as peripheral blood cells may vary from cell to cell. Treatment of the underlying condition typically corrects the defect. This deficiency is most likely linked to somatic mutations in the case of malignancy or toxic–metabolic effects on MPO activity.[41]

Specific Granule Deficiency

Neutrophil-specific granule deficiency (SGD) is a rare disorder of leukocyte maturation in which neutrophil secondary granules and some primary granule proteins are absent as a result of mutations in CCAAT/enhancer binding protein epsilon (*C/EBPε*, located at 14q11.2), a member of the leucine zipper family of transcription factors.[42] SGD is characterized by frequent, severe pyogenic infections, a paucity or absence of neutrophil-specific granule proteins and defensins, and atypical neutrophil nuclear structure with mostly bilobed nuclei. *In vitro*, these patients' cells show diminished neutrophil migration, reduced staphylococcal killing, reduced phagocytosis, and increased cell surface-to-volume ratio. Eosinophils and platelets are also affected in SGD. Platelets lack high-molecular-weight (HMW) von Willebrand factor multimers and have reduced platelet fibrinogen and fibronectin due to diminished platelet α granules. Bleeding diatheses and neutrophil phagocytosis of platelets are seen in SGD. In addition, SGD eosinophils are deficient in the eosinophil-specific granule proteins eosinophil cationic protein (ECP), eosinophil-derived neurotoxin (EDN), and major basic protein (MBP) despite the presence of messenger RNA (mRNA) transcripts for these proteins. Few patients have been reported to have survived beyond adolescence except for those with a milder dominant form. Bone marrow transplantation should be considered early in the course of the disease.

KEY CONCEPTS

Specific Granule Deficiency (SGD)

- SGD is caused by promyelocyte–myelocyte transition block as a result of a mutation in the *C/EBPε* gene.
- Absence of secondary granule proteins, and a selective loss of the primary granule defensins are the pathological findings in SGD granulocytes.
- The prognosis is very poor in recessive forms of SGD.

CHEDIAK-HIGASHI SYNDROME

Chediak-Higashi syndrome (CHS) is a rare, autosomal recessive disorder characterized by partial oculocutaneous albinism, increased susceptibility to infections, deficient NK-cell activity, and abnormal giant primary granules in neutrophils.[43] This immunodeficiency was first reported by Beguez-Cesar in 1943 and then further elaborated by Chediak and Higashi a decade later. The hallmark of CHS is giant abnormal granules in all granule-containing cells, including melanocytes (melanosomes are members of the lysosomal lineage of organelles), neutrophils, central and peripheral nerve tissues, fibroblasts, and hair. The problem is the inability to form appropriate lysosomes and cytoplasmic granules. CHS granulocytes lack cathepsin G and elastase, but the defensin content is normal. The giant granules of CHS are derived predominantly from azurophilic granules. CHS is classically described as a biphasic immunodeficiency, in which susceptibility to infection marks the first phase, and an accelerated lymphoproliferative syndrome with histiocytic infiltration of various tissues marks the second. Rarely, the accelerated phase may be the initial presentation. The giant organelles are derived from the late compartments of the endocytic pathway, affecting specifically late endosomes and lysosomes with minimal or no effect on early endosomes. *CHS1* encodes a 3801 amino acid peptide (lysosomal transporter [LYST]) that has a vital role in lysosomal trafficking. Lysosomal exocytosis triggered by membrane wounding is impaired in Chediak-Higashi fibroblasts. The reduced survival of fibroblasts after wounding indicates that impaired lysosomal exocytosis inhibits membrane resealing. Inability of cells to repair plasma membrane lesions may contribute to the pathology of CHS. The degree of albinism can vary from a slightly diluted skin pigment to hypopigmented skin and hair, photophobia, nystagmus, strabismus, macular hypoplasia, and reduced visual acuity. Skin biopsy shows large irregular melanin granules in melanocytes. Microscopic analysis of hair also shows poor distribution of melanin. Pancytopenia, neutropenia, and lack of NK-cell activity result in frequent pyogenic infections, usually caused by staphylococci or streptococci. Hepatosplenomegaly and lymphadenopathy are common. A mild bleeding diathesis results from platelet storage pool deficiency. Neurological dysfunction, including mental retardation, seizures, cranial nerve palsies, and progressive peripheral neuropathy, has been noted in CHS.

The lymphoma-like lymphohistiocytic accelerated phase is characterized by increased hepatosplenomegaly, lymphadenopathy, and worsened pancytopenia, which may resemble the virus-associated hemophagocytic syndromes or familial hemophagocytic lymphohistiocytosis. Although chemotherapy can induce transient remissions, relapses are common. Bone marrow transplantation prevents the accelerated phase and restores NK-cell function, but it does not resolve the central or peripheral nervous system abnormalities. Demonstration of giant azurophilic cytoplasmic inclusions on peripheral blood smear is very suggestive of the diagnosis of CHS; mutation analysis confirms the diagnosis.

HYPER-IGE RECURRENT INFECTION, OR JOB'S SYNDROME

Davis et al. first described hypoinflammatory recurrent infections with severe eczema in 1966. This was further refined and expanded by Buckley et al. in 1972, who recognized the characteristic IgE elevation. We now consider this to be a multisystem autosomal dominant disorder caused by heterozygous mutations in signal transducer and activator of transcription 3 (*STAT3*, located at 17q21).[44] Mutations in *STAT3* are mostly missense and clustered in either the DNA-binding domain or Src homology 2 (SH2) domains. Hyper-IgE recurrent infection (HIES, or Job's syndrome) is characterized by recurrent infections of the lower respiratory system and skin, chronic eczema, arterial anomalies, including coronary arterial tortuosity and aneurysms, extremely elevated IgE levels, and eosinophilia (Table 22.4). HIES occurs in all racial and ethnic groups.

Facial, Skeletal, and Dental Abnormalities

Facial abnormalities seen in the majority of the patients are a protruding, prominent mandible and forehead, apparent ocular hypertelorism, a broad nasal bridge, and a wide, fleshy nasal tip with increased interalar distance (Fig. 22.9). Midline anomalies common in this disorder are high-arched palate.

Skeletal abnormalities are common. Grimbacher et al. noted pathological fractures in 57% and scoliosis in 76% (Fig. 22.10). Low bone density and cortical bone loss are also common but not clearly correlated with the fracture rate. Other infrequent skeletal abnormalities reported in HIES are craniosynostosis, spina bifida, bifid rib, wedge-shaped lumbar vertebra, hemivertebra, and pseudoarthritis of the hip. Hyperextensibility of joints is common. A unique dental abnormality seen in this syndrome is retained primary teeth causing delayed eruption of permanent teeth.

TABLE 22.4 Clinical and Laboratory Findings in Patients With the Hyper-IgE Syndrome

Findings	Incidence (%)
Eczema	100
High IgE levels (>2000 IU/mL)	97
Eosinophilia (>2 SD above the mean for normals)	93
Boils	87
Pneumonia	87
Mucocutaneous candidiasis	83
Characteristic facies (in those ≥16 years)	83
Lung cysts	77
Scoliosis (for those ≥16 years)	76
Hyperextensible joints	68
Delayed shedding of primary teeth	72
Bone fractures	57

Adapted from Grimbacher B, Holland SM, Gallin JI, et al. Hyper-IgE syndrome disorder. N Engl J Med 1999; 340: 692.

KEY CONCEPTS

Hyper-IgE Recurrent Infection Syndrome (HIES) or Job's Syndrome

- Recurrent infections of the lower respiratory system and skin, chronic eczema, extremely elevated immunoglobulin E (IgE) levels, and eosinophilia are the hallmarks of the syndrome.
- Facial, skeletal, and dental abnormalities are very common.
- Lung abscesses and pneumatoceles following pneumonias caused by *Staphylococcus aureus* and *Haemophilus influenzae* are the major factors for morbidity.

Infections and Immunological Characteristics

Moderate to severe eczema presenting within the first hours to weeks of life is almost universal in HIES. Mucocutaneous candidiasis involving finger and toenails, mouth, vagina, and intertriginous areas is seen in most patients. Primary pulmonary infections are caused by *S. aureus*, *Haemophilus influenzae*, and *Streptococcus pneumoniae*. These pneumonias are often associated with abscess formation and usually lead to the development of pneumatoceles (see Fig. 22.10). Once lung cavities are formed, they provide an attractive environment for superinfection with *Pseudomonas* or *Aspergillus* spp. The clinical morphotype suggests abnormal tissue remodeling. *Pneumocystis jiroveci* pneumonia, cryptococcosis, histoplasmosis, and coccidioidomycosis have been reported. IgE levels are usually above 2000 IU/mL, but substantial fluctuations in IgE levels have been recorded over time, and the IgE levels do not correlate with disease activity or eosinophilia. Total serum IgG levels are usually within the normal range. Eosinophilia is common; the white blood cell (WBC) count is usually normal to low.

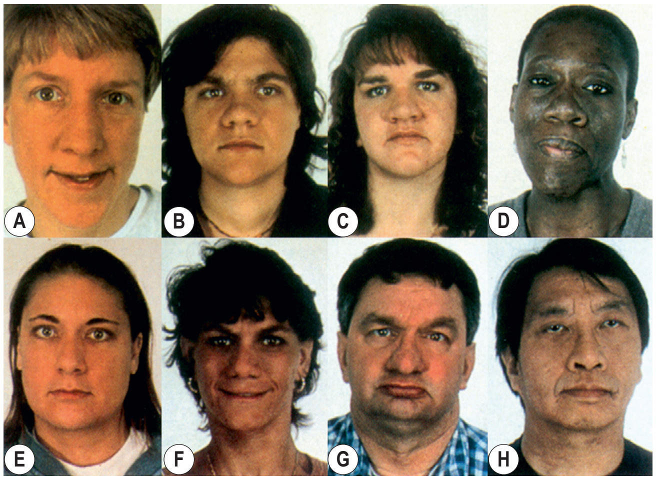


FIG 22.9 Facial Abnormalities Seen in Patients With Hyper-IgE Recurrent Infection Syndrome (HIES). Prominent mandible and forehead, apparent hypertelorism, broad nasal bridge with a wide nasal tip, and increased interalar distance are commonly seen facial features of HIES. (With permission from Grimbacher B, Holland SM, Gallin JI, et al. Hyper-IgE syndrome with recurrent infections—an autosomal dominant multisystem disorder. N Engl J Med 1999; 340: 692.)

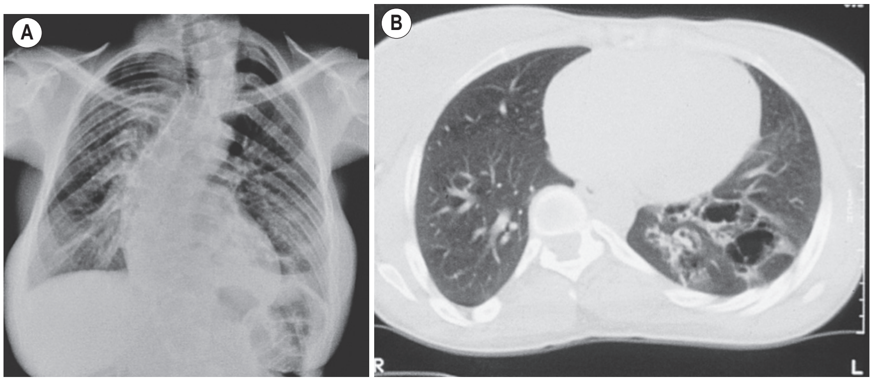


FIG 22.10 **Thoracic Pathology in Hyper-IgE Recurrent Infection Syndrome (HIES).** (A) Chest X-ray of a patient with scoliosis. (B) Computed tomography (CT) scan of the lungs in the same patient demonstrates multiple pneumatoceles caused by prior infections.

Mutations in *STAT3* lead to disruption of cytokine signaling, including IL-6, IL-10, IL-11, and IL-17. Signal transducer and activator of transcription 3 (STAT3) deficiency leads to elevated tumor necrosis factor-α (TNF-α) and IFN-γ, but reduced IL-17 producing T cells (Th17 cells). This latter defect may explain the predisposition to mucocutaneous candidiasis.[45] Memory B and T cells are low. Shingles vesicles are increased in Job's syndrome despite the fact that primary varicella-zoster virus (VZV) infections are normally eliminated. Lymphomas, but not epithelial malignancies, are increased in STAT3 deficiency.

Recently gain-of-function mutations in STAT3 associated with high IL-6 response have been identified in children characterized by failure to thrive, arthritis, livedo, and lung disease.

DOCK8 Deficiency (Autosomal Recessive Hyper-IgE Syndrome)

Several distinct diseases have elevated IgE, eczema, and eosinophilia. Dedicator of cytokinesis 8 (*DOCK8*, located at 9p24) deficiency includes food allergies, asthma, herpesvirus infections, human papilloma virus (HPV), and molluscum contagiosum infections, which are not part of dominant negative STAT3 deficiency. In addition to the infection susceptibility, patients with DOCK8 deficiency are predisposed to cutaneous and lymphoid malignancies.[46] Although Th17 cells are markedly diminished in patients with STAT3 deficiency, they are less severely reduced in those with DOCK8 deficiency (patients with HIES-like disease[45]). The specific contributions of these relatively IL-17–deficient states are areas of active investigation.

GATA2 Deficiency (Monomac Syndrome)

GATA2 (located at 3q21.3) is an early hematopoietic transcription factor most active in myeloid development.[47] The syndrome of monocytopenia and mycobacterial disease (monoMAC) is characterized by late childhood or adult-onset disseminated nontuberculous mycobacterial disease or disseminated fungal disease. Patients have absolute circulating monocytopenia, NK-cell cytopenia, and B-cell lymphopenia. Despite these circulating cytopenias, there are tissue macrophages and plasma cells, and Ig levels are normal to elevated. Neutrophils are variably affected. Other infections in this syndrome include HPV infection, molluscum contagiosum, histoplasmosis, and aspergillosis. Progressive pulmonary alveolar proteinosis is common, as are cytogenetic abnormalities of bone marrow, such as trisomy 8 and monosomy 7. Diagnosis is suspected on the basis of the infections and the abnormal results on hemography. Since most routine hematological laboratory studies allow quite low ranges of normal monocyte percentages, it is necessary to look at absolute monocyte numbers. Other presentations of this syndrome include aplastic anemia, acute myelogenous leukemia, lymphedema, and chronic myelomonocytic leukemia. Sequencing of *GATA2* is required for diagnosis.

ASSESSMENT OF NEUTROPHIL FUNCTION

Discrete abnormalities in neutrophil function lead to recurrent bacterial or fungal infections. Assays have been developed to interrogate those functions. However, since neutrophils cannot be viably stored or frozen, samples are usually examined fresh with simultaneous normal volunteer controls. The techniques discussed here are reviewed in detail by Elloumi et al.[48] and in Chapter 94.

Isolation of Neutrophils

Most assays require neutrophils to be purified away from other leukocytes and blood components. Blood is usually anticoagulated using either citrate or heparin (10 units/mL) tubes and maintained at 20–25°C (room temperature) in polypropylene containers. Most protocols use differences in the cell density as the basis for the separation by sedimentation or centrifugation, or both. Typically, $1–2 \times 10^6$ neutrophils can be isolated per milliliter of whole blood.

Neutrophil Adherence

The adhesive function of phagocytes is commonly assessed by passage of 1 mL of whole blood through a column filled with nylon wool. Adherence is measured as the difference in the absolute neutrophil count of the precolumn sample and of the sample after passage through the nylon wool column. Alternatively, isolated neutrophils can be induced to bind to plastic using a 96-well plate either uncoated or coated with fetal bovine serum, a ligand like ICAM-1, or a specific ECM protein, such as fibrinogen or fibronectin. Endothelial cell monolayers harvested from human umbilical veins can serve as a more physiological substrate for the measurement of cell adhesion. Isolated neutrophils are preloaded with the cell permeant, acetoxymethyl ester derivative of the fluorescent dye, calcein (calcein-AM). Nonspecific esterases in the cytosol cleave the ester linkage, trapping the fluorescent probe in the cytosol. The labeled neutrophils are added to each well and incubated in the absence or presence of phorbol myristate acetate (PMA) to promote adherence through activated integrins. At the end of the incubation, the wells are washed to remove nonadherent cells. The fluorescence of each well is determined with a fluorescent microplate reader and compared with the fluorescence of a control well with a fixed number of fluorescent cells. Under control conditions, fewer than 10% of the neutrophils adhere to plastic or to plastic coated with fetal bovine serum. Slightly more neutrophil adherence is observed on wells coated with fibrinogen. Treatment of normal neutrophils with PMA for 30 minutes results in the adherence of 100% of the neutrophils under all conditions. This adherence is abnormal in patients with LAD. Neutrophils isolated from patients with typical LAD-1 generally exhibit markedly reduced adherence under both unstimulated and PMA conditions.

Neutrophil Chemotaxis

Neutrophil chemotaxis *in vivo* can be evaluated by using skin windows. Skin blisters are gently raised on the volar surface of the forearm using a vacuum pump and a blister device, with little hemorrhage or vascular damage. The roof of the blister is removed, and the exposed dermis is bathed with autologous serum with the use of a skin window chamber. In 24 hours, exudative neutrophils accumulate in the autologous serum bathing the skin lesion. The skin chamber provides a mechanism for characterizing the immune cells as well as the soluble immune mediators that accumulate in the autologous serum during the evolution of the inflammatory response. Chemotaxis *in vitro* is generally measured by using a Boyden chamber. The Boyden chamber includes three components: a lower (chemoattractant) chamber, a nitrocellulose or polycarbonate filter layer, and an upper cell chamber. The lower compartments of the Boyden chamber are filled with a chemoattractant, such as fMLF; 10^{-8} M) or IL-8 (10 ng/mL). Recently an improved rapid fluorescence-based measure of neutrophil chemotaxis, which uses a new 96-well disposable chemotaxis chamber that can be read in a fluorescence microplate reader, has been developed.

Expression of Surface Antigens

The expression of cell surface antigens on neutrophils relies on labeled mAbs analyzed by flow cytometry. The panel may include the β_2 integrins (CD11a, CD11b, CD11c, and CD18), selectins (CD62L), Fcγ receptors I, II, and III (CD64, CD32 and CD16), leukosialin (CD43), the common leukocyte antigen (CD45), and markers for the specific granules (CD67), and azurophilic granules (CD63). The expression of surface antigens can be used to assess the responsiveness of neutrophils to particular ligands, such as fMLF and lipopolysaccharide (LPS).

Neutrophil Degranulation

The proteases, acid hydrolases, and inflammatory mediators released from storage granules in neutrophils can mediate bacterial killing, tissue damage, healing, and immune regulation. Lactoferrin from specific granules can chelate iron, resulting in a bactericidal or bacteriostatic effect. Stimulation of neutrophils with various secretagogues can release granular enzymes into the extracellular fluid. Treatment of the neutrophils with cytochalasin b (5 μg/mL) disrupts microfilament assembly and facilitates the release of both specific and azurophilic enzymes. To differentiate degranulation from cell lysis, release of the cytosolic enzyme lactate dehydrogenase should be monitored simultaneously. The release of azurophilic granules can be assessed by determination of β-glucuronidase activity. Supernatant fluids or cell extracts obtained from stimulated neutrophils are incubated with 4-methylumbelliferyl-β-ᴅ-glucuronide. Alternatively, myeloperoxidase can be determined by using commercially available enzyme-linked immunoassays. CD63 is also found in the membrane of azurophilic granules and migrates to the neutrophil surface after stimulation with fMLF in the presence of cytochalasin b. The release of specific granules can be assessed by determination of lactoferrin levels using an enzyme-linked immunoassay. The carcinoembryonic antigen CD66b (formerly CD67) is found on the neutrophil surface and in the specific granules, and its expression on the surface of the neutrophils is increased after stimulation with fMLF or LPS. Detection of the constituents of secretory granules can be assessed by flow cytometric analysis of the change in expression of surface proteins, such as adhesion molecules, and cytochrome b_{558} of the NADPH oxidase.

Generation of Reactive Oxygen Species

The release of reactive oxygen intermediates, such as O_2^- and H_2O_2, is an important component of the neutrophil bactericidal machinery. Neutrophils isolated from patients with CGD are unable to generate superoxide, leading to their oxygen-dependent bactericidal defect. The production of O_2^- can be detected by using the reduction of cytochrome c. Because O_2^- causes a one-to-one stoichiometric reduction of ferricytochrome c to ferrocytochrome c, the resultant increase in the absorption spectrum at 550 nM can be used to quantitate the production of O_2^-. Superoxide dismutase is added to an identical tube to control for the nonspecific reduction of cytochrome c. However, since cytochrome is not permeable to the cells, the detection of O_2^- is limited to that released into the extracellular milieu. Neutrophils isolated from patients with CGD produce little O_2^- in response to PMA in 10 minutes. However, some patients with forms of CGD associated with residual superoxide production have low, but detectable, O_2^- production in 60 minutes. Neutrophils isolated from X-linked heterozygous carriers of CGD can yield a full spectrum of O_2^- production, whereas neutrophils from autosomal recessive carriers of CGD generally yield a normal response. Although the detection of O_2^- by reduction of cytochrome c is useful in the diagnosis of patients with CGD, it cannot be used in the diagnosis of carriers because of the wide spectrum of responses that result from the degree of X-chromosome lyonization.

The extracellular release of H_2O_2 can be measured by using horseradish peroxidase–induced oxidation of either phenol red or Amplex red. PMN suspensions in the presence of horseradish peroxidase and one of the chromophores are exposed to either PMA or buffer alone. Changes in optical density of phenol red at 600 nm can be determined with a standard microplate reader. Amplex red is a much more sensitive fluorescent chromophore, and H_2O_2-dependent changes in fluorescence can be determined with a fluorescence microplate reader.

The NBT test is a qualitative assay of O_2^- production. Either whole blood or isolated neutrophils are mixed with NBT in a chamber slide and stimulated with PMA for 15–30 minutes at 37°C. The slide is counterstained with 0.1% safranin and examined under a microscope. The NBT test yields a visual record of the reduction of the NBT dye to the insoluble, blue-black deposits of formazan. Normal neutrophils, but not neutrophils from patients with CGD, reduce the yellow dye to black-brown-blue aggregates in the cells. The NBT test can be used to diagnose

X-linked carriers of CGD but cannot differentiate autosomal carriers from normal subjects.

An alternative to the NBT test is a flow cytometric assay, in which the dye, dihydrorhodamine-1,2,3 is used. Neutrophils are loaded with the nonfluorescent dye and then stimulated with PMA for 15 minutes at 37°C. The H_2O_2 produced oxidizes the dye and results in increased fluorescence, detectable with a flow cytometer. Catalase is added to prevent cell-to-cell diffusion of H_2O_2. Since dye is localized to the cytoplasm, and catalase is present in the extracellular fluid, the dihydrorhodamine-1,2,3 assay detects the intracellular production of reactive oxygen metabolites. Stimulation of normal neutrophils with PMA results in a two-log shift in the fluorescence intensity. Neutrophils from an X-linked carrier of CGD exhibit mosaicism with a negatively stained (abnormal) population and a brightly stained positive population. Neutrophils from a patient with X-linked CGD that lack $gp91^{phox}$ express little increase in fluorescence, whereas neutrophils from a patient with a deficiency in $p47^{phox}$ exhibit a slight increase in fluorescence. The major advantages of the dihydrorhodamine-1,2,3 assay are the sensitivity, the signal-to-noise ratio, and the ease of counting a large number of cells.

Western Blot for Determination of NADPH Oxidase Defect

Determination of the specific protein defect in CGD by Western blot analysis provides direction for the genetic defect. A validated normal control and a typical $gp91^{phox}$ CGD are included on each blot to ensure adequate development of $p22^{phox}$. Patients with $p47^{phox}$ CGD are Western blot negative. Patients with $p67^{phox}$ CGD are generally Western blot negative. Because $p22^{phox}$ and $gp91^{phox}$ exist as a membrane complex, patients with a defect are $p22^{phox}$, in general, are Western blot negative for both $p22^{phox}$ and $gp91^{phox}$. In contrast, defects in $gp91^{phox}$ yield more variable results. In general, patients with nonsense defects in $gp91^{phox}$ exhibit low but detectable levels of $p22^{phox}$. Patients with missense mutations in $gp91^{phox}$ that yield detectable $gp91^{phox}$ protein exhibit proportionately higher levels of $p22^{phox}$.

TRANSLATIONAL RESEARCH

The progress made over the past 60 years in understanding and managing phagocyte defects has been remarkable, spanning initial identification, phenotyping, and molecular characterization. The last 20 years have seen the advent of oral antifungals, potent oral antibiotics, oral antivirals, reduced intensity conditioning for bone marrow transplantation, and gene therapy. These advances have transformed the quality of life and longevity of all patients with immune deficiencies. We now need further in-depth study of the mechanisms that drive disease pathophysiology to gain novel insights. The advent of therapeutic cytokines, small molecule inhibitors and agonists, RNA inhibitors, and effective means of gene transfer should make directed approaches to disease modification a reality. However, we must be sure we know exactly what functions need addressing. The downstream effects of genetic defects are surprisingly complex and not always as straightforward as has been anticipated. More detailed clinical and functional phenotyping, careful examination of developmental and gene expression effects, and continued longitudinal studies of large cohorts will convert these severe diseases punctuated by acute, life-threatening infections into chronic conditions that are successfully treated medically or successfully managed until successful bone marrow transplantation becomes available.

ON THE HORIZON

- Early recognition and molecular diagnosis of all phagocyte defects, leading to prophylactic antimicrobial treatment, where indicated
- Improvement in bone marrow transplantation technology to allow for early, safe, successful, fertility-preserving transplantation in all cases
- Understanding the mechanisms of the hepatic complications of chronic granulomatous disease (CGD) that correlate with mortality
- Characterizing the pathways that converge on signal transducer and activator of transcription 3 (STAT3) signaling causing the complex somatic and immune disorder, hyper-IgE recurrent infection syndrome (HIES).

ACKNOWLEDGMENT

This work was supported by the intramural program of the National Institute of Allergy and Infectious Diseases.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Kaushansky K. Lineage-specific hematopoietic growth factors. N Engl J Med 2006;354:2034–45.
2. Berliner N. Lessons from congenital neutropenia: 50 years of progress in understanding myelopoiesis. Blood 2008;111:5427–32.
3. Perdiguero EG, Geissmann F. The development and maintenance of resident macrophages. Nat Immunol 2016;17:2–8.
4. Boztug K, Klein C. Genetics and pathophysiology of severe congenital neutropenia syndromes unrelated to neutrophil elastase. Hematol Oncol Clin North Am 2013;27:43–60.
5. Kollner I, Sodeik B, Schreek S, et al. Mutations in neutrophil elastase causing congenital neutropenia lead to cytoplasmic protein accumulation and induction of the unfolded protein response. Blood 2006;108:493–500.
6. Dale DC, Welte K. Cyclic and chronic neutropenia. Cancer Treat Res 2011;157:97–108.
7. Klein C, Grudzien M, Appaswamy G, et al. HAX1 deficiency causes autosomal recessive severe congenital neutropenia (Kostmann disease). Nat Genet 2007;39:86–92.
8. Germeshausen M, Grudzien M, Zeidler C, et al. Novel HAX1 mutations in patients with severe congenital neutropenia reveal isoform-dependent genotype-phenotype associations. Blood 2008;111:4954–7.
9. Boztug K, Appaswamy G, Ashikov A, et al. A syndrome with congenital neutropenia and mutations in G6PC3. N Engl J Med 2009;360:32–43.
10. Boocock GR, Marit MR, Rommens JM. Phylogeny, sequence conservation, and functional complementation of the SBDS protein family. Genomics 2006;87:758–71.
11. Huang JN, Shimamura A. Clinical spectrum and molecular pathophysiology of Shwachman-Diamond syndrome. Curr Opin Hematol 2011;18:30–5.
12. Capsoni F, Sarzi-Puttini P, Zanella A. Primary and secondary autoimmune neutropenia. Arthritis Res Ther 2005;7:208–14.
13. Maheshwari A, Christensen RD, Calhoun DA. Immune neutropenia in the neonate. Adv Pediatr 2002;49:317–39.
14. Nauseef WM, Borregaard N. Neutrophils at work. Nat Immunol 2014;15:602–11.
15. van de Vijver E, van den Berg TK, Kuijpers TW. Leukocyte adhesion deficiencies. Hematol Oncol Clin North Am 2013;27:101–16.
16. Kishimoto TK, Hollander N, Roberts TM, et al. Heterogeneous mutations in the beta subunit common to the LFA-1, Mac-1, and p150,95 glycoproteins cause leukocyte adhesion deficiency. Cell 1987;50:193–202.
17. Moutsopoulos NM, Konkel J, Sarmadi M, et al. Defective neutrophil recruitment in leukocyte adhesion deficiency type I disease causes local IL-17-driven inflammatory bone loss. Sci Transl Med 2014;6:229–40.

18. Roos D, Meischl C, de Boer M, et al. Genetic analysis of patients with leukocyte adhesion deficiency: genomic sequencing reveals otherwise undetectable mutations. Exp Hematol 2002;30:252–61.
18a. Moutsopoulos NM, Zerbe CS, Wild T, et al. Interleukin-12 and interleukin-23 blockade in leukocyte adhesion deficiency type 1. New Engl J Med 2017;376:1141–6.
19. Etzioni A, Gershoni-Baruch R, Pollack S, et al. Leukocyte adhesion deficiency type II: long-term follow-up. J Allergy Clin Immunol 1998;102:323–4.
20. Alon R, Etzioni A. LAD-III, a novel group of leukocyte integrin activation deficiencies. Trends Immunol 2003;24:561–6.
21. Malinin NL, Zhang L, Choi J, et al. A point mutation in KINDLIN3 ablates activation of three integrin subfamilies in humans. Nat Med 2009;15:313–18.
22. Svensson L, Howarth K, McDowall A, et al. Leukocyte adhesion deficiency-III is caused by mutations in KINDLIN3 affecting integrin activation. Nat Med 2009;15:306–12.
23. Marciano BE, Spalding C, Fitzgerald A, et al. Common severe infections in chronic granulomatous disease. Clin Infect Dis 2015;60:1176–83.
24. Segal AW. How neutrophils kill microbes. Annu Rev Immunol 2005;23:197–223.
25. Papayannopoulos V, Zychlinsky A. NETs: a new strategy for using old weapons. Trends Immunol 2009;30:513–21.
26. Roos D, Kuhns DB, Maddalena A, et al. Hematologically important mutations: X-linked chronic granulomatous disease (third update). Blood Cells Mol Dis 2010;45:246–65.
27. Royer-Pokora B, Kunkel LM, Monaco AP, et al. Cloning the gene for an inherited human disorder–chronic granulomatous disease—on the basis of its chromosomal location. Nature 1986;322:32–8.
28. Deardorff MA, Gaddipati H, Kaplan P, et al. Complex management of a patient with a contiguous Xp11.4 gene deletion involving ornithine transcarbamylase: a role for detailed molecular analysis in complex presentations of classical diseases. Mol Genet Metab 2008;94:498–502.
29. Roos D, Kuhns DB, Maddalena A, et al. Hematologically important mutations: the autosomal recessive forms of chronic granulomatous disease (second update). Blood Cells Mol Dis 2010;44:291–9.
30. Matute JD, Arias AA, Wright NA, et al. A new genetic subgroup of chronic granulomatous disease with autosomal recessive mutations in p40 phox and selective defects in neutrophil NADPH oxidase activity. Blood 2009;114:3309–15.
31. Kuhns DB, Alvord WG, Heller T, et al. Residual NADPH oxidase and survival in chronic granulomatous disease. N Engl J Med 2010;363:2600–10.
32. Leiding JW, Freeman AF, Marciano BE, et al. Corticosteroid therapy for liver abscess in chronic granulomatous disease. Clin Infect Dis 2012;54:694–700.
33. Feld JJ, Hussain N, Wright EC, et al. Hepatic involvement and portal hypertension predict mortality in chronic granulomatous disease. Gastroenterology 2008;134:1917–26.
34. Segal BH, DeCarlo ES, Kwon-Chung KJ, et al. Aspergillus nidulans infection in chronic granulomatous disease. Medicine (Baltimore) 1998;77:345–54.
35. Vinh DC, Shea YR, Jones PA, et al. Chronic invasive aspergillosis caused by Aspergillus viridinutans. Emerg Infect Dis 2009;15:1292–4.
36. Siddiqui S, Anderson VL, Hilligoss DM, et al. Fulminant mulch pneumonitis: an emergency presentation of chronic granulomatous disease. Clin Infect Dis 2007;45:673–81.
37. Greenberg DE, Shoffner AR, Zelazny AM, et al. Recurrent Granulibacter bethesdensis infections and chronic granulomatous disease. Emerg Infect Dis 2010;16:1341–8.
38. Marciano BE, Rosenzweig SD, Kleiner DE, et al. Gastrointestinal involvement in chronic granulomatous disease. Pediatrics 2004;114:462–8.
39. De Ravin SS, Naumann N, Cowen EW, et al. Chronic granulomatous disease as a risk factor for autoimmune disease. J Allergy Clin Immunol 2008;122:1097–103.
40. Güngör T, Teira P, Slatter M, et al. Reduced-intensity conditioning and HLA-matched haemopoietic stem-cell transplantation in patients with chronic granulomatous disease: a prospective multicentre study. Lancet 2014;383:436–48.
41. Lanza F. Clinical manifestation of myeloperoxidase deficiency. J Mol Med 1998;76:676–81.
42. Gombart AF, Koeffler HP. Neutrophil specific granule deficiency and mutations in the gene encoding transcription factor C/EBP(epsilon). Curr Opin Hematol 2002;9:36–42.
43. Kaplan J, De Domenico I, Ward DM. Chediak–Higashi syndrome. Curr Opin Hematol 2008;15:22–9.
44. Freeman AF, Holland SM. Clinical manifestations, etiology, and pathogenesis of the hyper-IgE syndromes. Pediatr Res 2009;65:32R–37R.
45. Milner JD, Sandler NG, Douek DC. Th17 cells, Job's syndrome and HIV: opportunities for bacterial and fungal infections. Curr Opin HIV AIDS 2010;5:179–83.
46. Zhang Q, Davis JC, Lamborn IT, et al. Combined immunodeficiency associated with DOCK8 mutations. N Engl J Med 2009;361:2046–55.
47. Spinner MA, Sanchez LA, Hsu AP, et al. GATA2 deficiency: a protean disorder of hematopoiesis, lymphatics, and immunity. Blood 2014;123:809–21.
48. Elloumi HZ, Holland SM. Diagnostic assays for chronic granulomatous disease and other neutrophil disorders. Methods Mol Biol 2014;1124:517–35.

MULTIPLE-CHOICE QUESTIONS

1. A 12-year-old boy with X-linked chronic granulomatous disease (CGD) presents to your office with progressive fever, dyspnea, and hypoxia a week after a Halloween hayride at a local farm. His chest X-ray shows extensive bilateral military infiltrates. The MOST likely diagnosis and management:
   - **A.** *Burkholderia cepacia* complex pneumonia; initiate antibacterials
   - **B.** *Nocardia spp.* pneumonia; initiate antibacterials
   - **C.** *Serratia marcescens* pneumonia; initiate antibacterials
   - **D.** Mulch pneumonitis; initiate antifungals and corticosteroids
   - **E.** *Staphylococcus aureus* bacteremia; initiate antibacterials and obtain echocardiogram

2. Leukocyte adhesion defect-1 (LAD-1) leads to low numbers of neutrophils in the tissues (tissue neutropenia) and recurrent severe infections. Progressive severe periodontitis is characteristic, with most patients losing all their teeth by late adolescence. The immunological mechanism behind the periodontal disease is:
   - **A.** Tissue neutropenia leading to severe bacterial infection
   - **B.** Excessive interferon (IFN)-γ production
   - **C.** Poor root maintenance by defective macrophages
   - **D.** Excessive interleukin (IL)-17 production
   - **E.** Severe tooth decay caused by impaired neutrophil trafficking

3. GATA2 deficiency is a complex disease with manifestations involving hematopoietic and lymphatic development. It is often identified in adolescence or adulthood with cytopenias and/or infections. Signs and symptoms that should lead to early consideration of GATA2 deficiency include:
   - **A.** Human papilloma virus (HPV) infection, mycobacterial infection, monocytopenia
   - **B.** Respiratory syncytial virus (RSV) infection, staphylococcal pneumonia, monocytosis
   - **C.** Influenza infection, *Aspergillus* infection, thrombocytopenia
   - **D.** Adenovirus infection, *Bordetella* infection, anemia
   - **E.** Cytomegalovirus (CMV) infection, streptococcal infection, lymphocytosis

23

Mast Cells, Basophils, and Mastocytosis

Joshua Boyce, Laura Fanning

Mast cells and basophils are functionally related effector cells of hematopoietic origin that are implicated in allergy, type 2 immune responses to parasites, and innate immunity. Both cell types derive from bone marrow progenitors and express high-affinity immunoglobulin E (IgE) Fc receptors (FcεRI) on their surface. Both release histamine and generate a variety of other inflammatory mediators in response to both IgE- and non–IgE-mediated activation. However, there are important differences between mast cells and basophils in terms of their developmental origins, morphology, distribution, and mechanisms and responses to stimulation[1] (Table 23.1, Fig. 23.1). This chapter will deal with the functional characteristics of both cell types in host defense and disease.

KEY CONCEPTS

Origins of Basophils and Mast Cells

- Mast cells and basophils represent distinct hematopoietic lineages.
- Basophils are granulocytes that mature in the bone marrow, circulate in blood, and can be recruited into peripheral tissues at sites of inflammatory responses.
- Mast cell progenitors arise in bone marrow, but mast cells mature in peripheral tissues.
- Mast cells are present throughout connective tissues and mucosal surfaces and are particularly abundant near epithelial surfaces exposed to the environment.
- CD117 and its ligand stem cell factor (SCF) are critical for mast cell development and survival.
- Basophil development is promoted in the context of T_H2 inflammatory responses and regulated by cytokines, including interleukin-3 (IL-3) and thymic stromal lymphopoietin (TSLP).

DEVELOPMENT AND DISTRIBUTION OF MAST CELLS

Mast Cell Development and Survival

Mast cells are evolutionarily ancient, predating the emergence of adaptive immunity. A mast cell–like cell containing histamine and heparin has been identified in *Styela plicata*, a urochordate species that first appeared around 500 million years ago.[1] Studies in rodents indicate that mast cells derive from hematopoietic stem cell (HSC)–derived granulocyte–monocyte progenitor cells (GMPs) in bone marrow, which can give rise to basophils and other lineages.[2] In mice, the transcription factors C/EBPα and MITF interact antagonistically and specify basophil versus mast cell fate, respectively, in GMPs. Basophils undergo full differentiation in bone marrow, whereas mast cells differentiate from committed precursors that migrate to their ultimate tissue destinations before acquiring the morphological and biochemical characteristics of mature mast cells.[3] Moreover, circulating basophils are short-lived and terminally differentiated, whereas mast cells can survive for long periods, and retain the capacity for proliferation and substantial inducible gene expression. This likely explains the high degree of heterogeneity among mast cell subpopulations.

In mice, committed mast cell progenitors (MCPs) can be identified cytofluorographically as a Kit (CD117)^{hi,} FcεRI$^+$, β7 integrin$^+$ mononuclear cell population. A similar cell population was recently identified in human blood.[4] In *ex vivo* colony forming assays, both mouse and human CD117$^+$, FcεRI$^+$, β7 integrin$^+$ cells give rise preferentially to mast cells. In studies of cells sorted from various organs BALB/c mice, committed basophil precursors (BaPs) were found exclusively in bone marrow, whereas committed mast cell progenitors were most abundant in the small intestine.[2] This study identified a splenic population of cells capable of giving rise to both mast cells and basophils with *ex vivo* culture (bipotent basophil–mast cell precursor [BMCP]). However, neither mature mast cells nor mast cell progenitors show significant overlap in gene expression profiles with basophils in either mice or humans.[4,5] Thus despite some functional similarities, the bulk of evidence suggests that the developmental pathways of mast cells and basophils diverge early in hematopoiesis and that the two cell types likely serve correspondingly very different functions.

CD117 and its ligand, stem cell factor (SCF), are absolutely required for mast cell development. SCF exists as both a membrane-bound protein expressed by fibroblasts and other stromal cells, as well as a soluble form. Mice with mutations in the loci for CD117 (W/W^v or W^{sh}/W^{sh} mice) or SCF (Sl/Sl^d mice) lack mast cells, among other hematopoietic and nonhematopoietic defects.[1] Although loss-of-function mutations in CD117 or SCF have not been identified in humans, gain-of-function mutations in CD117 are associated with systemic mastocytosis,[6] suggesting that the SCF-KIT axis is important for human mast cell development as well. Although SCF can stimulate mast cell development from either mouse or human progenitor cells *in vitro,* mouse mast cells can be derived *in vitro* from unfractionated bone marrow cells by using interleukin-3 (IL-3) or thymic stromal lymphopoietin (TSLP) as the sole exogenous growth factors.[7] Mice lacking IL-3 have wild-type levels of mast cells under steady-state conditions but show diminished capacity to develop a reactive mast cell hyperplasia in the intestine in response to helminth infections.[8] TSLP knock-out mice display modest steady-state reductions in mast cell numbers in all organs studied.[7] Additionally, transgenic overexpression of IL-9 results in marked

TABLE 23.1 **Natural History, Major Mediators, and Surface Membrane Structures of Human Mast Cells and Basophils**

Characteristic	Basophils	Mast Cells
Natural History		
Origin of precursor cells	Bone marrow	Bone marrow
Site of maturation	Bone marrow	Peripheral tissues (a few in bone marrow)
Mature cells in the circulation	Yes (usually <1% of blood leukocytes)	No
Mature cells recruited into tissues from circulation	Yes (during immunological, inflammatory responses)	No
Mature cells normally residing in connective tissues	No (not detectable by microscopy)	Yes
Proliferative ability of morphologically mature cells	None reported	Yes (in certain circumstances)
Life span	Days (like other granulocytes)	Weeks to months (based on studies in rodents)
Mediators		
Major mediators stored preformed in cytoplasm	Histamine, chondroitin sulfates-glucuronidase, elastase, cathepsin G-like enzyme, eosinophil cationic protein, major basic protein, Charcot-Leyden crystal protein, eosinophil peroxidase, some positive for tryptase, chymase, carboxypeptidase A	Histamine, heparin and/or chondroitin sulfates, chymase and/or tryptase, many acid hydrolases, cathepsin G, carboxypeptidase A
Major lipid mediators produced upon appropriate activation	LTC$_4$, PAF	PGD$_2$, LTC$_4$, LTB$_4$, PAF, TXA$_2$
Cytokines and growth factors released upon appropriate activation	IL-2, IL-3, IL-4, IL-5, IL-6, IL-9, IL-13, IL-15, TSLP	IL-1β, IL-2, IL-3, IL-4, IL-5, IL-6, IL-8, IL-9, IL-10, IL-11, IL-13, IL-15, IL-16, IL-17A, IL-18, IL-22, IL-24, IL-25, IL-33, TNF-α, SCF, NGF, TGF-β, FGF-2, VEGF, IFN-α, GM-CSF, TSLP, amphiregulin
Chemokines produced	CCL3, CCL4, CCL12, CXCL2	CCL1, CCL2, CCL3, CCL4, CCL5, CCL7, CCL8, CCL13, CCL17, CCL18, CCL19, CCL20, CCL22, CCL23, CCL25, CCL28, CXCL2, CXCL3, CXCL4, CXCL5, CXCL7, CXCL8, CXCL10, CXCL14, CXCL16, CXCL17, XCL1, CX3CL1
Surface Structures		
Ig receptors	FcεRI, FcγRII (CDw32), FcγRIIB, FcγRIIIA	FcεRI, FcγRI, FcγRIIA, FcγRIIB, FcγRIIIA
Cytokine/growth factor receptors	IL-1RIIb (CD121b), IL-2R (CD25), IL-3R, IL-4Rα, IL-5Rα, IL-18R, IL-33R, TSLPR, GM-CSFRα, TrkA	KIT (SCF receptor), IL-3R, IL-4Rα, IL5Rα, IL-6R, IL-18R, IL-33R, IFN-γRα, TGF-β type-I and type-II β receptors, TrkA, T1/ST2/IL-1R4
Chemokine receptors	CCR1, CCR2, CCR3, CCR5, CXCR1, CXCR2, CXCR4	CCR1, CCR3, CCR4, CCR5, CXCR1, CXCR2, CXCR3, CXCR4, CX3CR1
Other GPCRs	CRTH2, C3aR, C5aR	EP3, BLT1, BLT2, cysLT1, PAF receptor, C3aR, C5aR

LT, leukotriene; PG, prostaglandin; PAF, platelet-activating factor; TXA$_2$, thromboxane A$_2$; IL, interleukin; TSLP, thymic stromal lymphopoietin; SCF, stem cell factor; NGF, nerve growth factor; TGF-β, transforming growth factor-β; FGF-2, fibroblast growth factor-2; VEGF, vascular endothelial growth factor; IFN-α, interferon-α; GM-CSF, granulocyte macrophage–colony-stimulating factor; CCL, chemokine (C-C motif) ligand; CXCL, chemokine (C-X-C motif) ligand; XCL, chemokine (C motif) ligand; CX3CL, chemokine (C-X3-C motif) ligand; TrkA, tropomyosin receptor kinase A; CCR, C-C chemokine receptor; CXCR, C-X-C chemokine receptor; CX3CR, C-X3-C chemokine receptor; GPCRs, G protein–coupled receptors; CRTH2, chemoattractant receptor-homologous molecule expressed on Th2 cells; C3aR, complement component 3a receptor; C5aR, complement component 5a receptor.

hyperplasia of mast cells in multiple organs and tissues.[9] However, the profound reduction of mast cells in W/W^v, W^{sh}/W^{sh}, and *Sl/Sld* mice argue that none of these cytokines can substitute for the absence of SCF or CD117.

In contrast to mice, SCF is absolutely required to derive mast cells *in vitro* from human bone marrow, fetal liver, or cord blood progenitors. IL-3 cannot induce mast cell development from human hematopoietic progenitor cells, although it does can enhance survival of human cord blood mast cells (as do IL-4, IL-5, or IL-6).[10] Human intestinal mast cells proliferate in response to IL-4 when it is provided in combination with SCF. Notably, IL-4 primes human cord blood and/or intestinal mast cells for FcεRI-dependent degranulation, cytokine generation, and cysteinyl leukotriene (cysLT) generation,[11] the latter reflecting strong induction of leukotriene C$_4$ synthase (LTC$_4$S) expression. Thus cytokines other than SCF are likely to play context-specific roles in mast cell development and may help to account for the substantial heterogeneity of anatomically and functionally distinct mast cell subpopulations (see below). Tissue mast cells vary widely in their capacity for cysLT generation, which may reflect cytokine-induced modification of their effector phenotype *in vivo*.

Homing, Distribution, and Heterogeneity of Mast Cells

Mature mast cells are distributed throughout connective tissues and mucosal surfaces, often adjacent to blood or lymphatic vessels or peripheral nerves. Much of the current understanding of mast cell tissue homing comes from mouse models. MCPs home

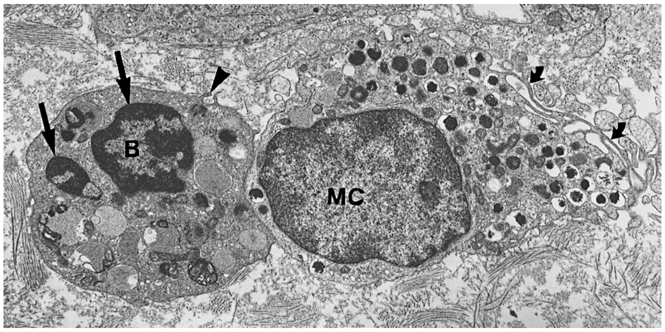


FIG 23.1 **Ultrastructures of Mast Cell and Basophil.** A basophil (B) adjacent to a mast cell (MC) in the ileal submucosa of a patient with Crohn disease. The basophil exhibits a bilobed nucleus (solid arrows) whose chromatin is strikingly condensed beneath the nuclear membrane. The basophil surface is relatively smooth with a few blunt processes (arrowhead). The mast cell nucleus is larger and its chromatin less condensed than that of the basophil. The mast cell's granules are smaller, more numerous, and more variable in shape and content than those of the basophil. The mast cell surface has numerous elongated, thin folds (curved arrows). (Original magnification ~ × 9000). (From Dvorak AM, Monahan RA, Osage JE, Dickersin GR. Crohn's disease: transmission electron microscopical studies. Hum Pathol 1980;11:606–619, with permission from Ann M. Dvorak.)

constitutively to the mouse small intestine, requiring both $\alpha_4\beta_7$ integrin, as well as the chemokine receptor CXCR2.[12] Although the lungs of naïve mice contain few mast cells, MCP recruitment can be experimentally induced in response to allergen sensitization and challenge. In this context, MCP recruitment to the lung is controlled by $\alpha_4\beta_7$ as well as $\alpha_4\beta_1$ and its counterligand, vascular cell adhesion molecule-1 (VCAM-1).[12] To date, no studies have addressed the requirements for homing or recruitment of MCPs to skin, the central nervous system, or other characteristic destinations. Cultured human mast cells express the chemokine receptors CCR3 and CXCR4[13] and respond chemotactically to their corresponding ligands. Mast cells localizing to asthmatic airway smooth muscle express CXCR3, and lung mast cell migration can be induced by airway smooth muscle cultures via activation of CXCR3.[14] However, definitive information regarding the homing requirements for human mast cells *in vivo* is lacking because of technical constraints.

Once in the tissues, MCPs give rise to anatomically and biochemically distinct subpopulations of mast cells. Rodent mast cell subsets are often referred to as *mucosal mast cells* (MMCs) and *connective tissue mast cells* (CTMCs), respectively, based on differences in anatomical distribution and staining characteristics. Early studies of Kanakura et al. used adoptive transfers to demonstrate that both MMCs and CTMCs arose in vivo from a single population of progenitor cells.[15] The differences between MMCs and CTMCs in staining characteristics relate to the relative content of heparin versus chondroitin sulfate glycosaminoglycans (GAGs) bound to the serglycin core of mast cell granules. These GAGs impart the basophilic staining characteristics of mast cell granules and are essential for the storage and stability of highly charged serine proteases and histamine that comprise the major preformed products of mast cells. The mouse genome encodes four mast cell–restricted tryptases and several chymases (at least five of which can be expressed by mast cells), as well as carboxypeptidase A3 (CPA3). Accordingly, the anatomically distinct mouse mast cell subpopulations express very different profiles of proteases (and correspondingly different GAG compositions), dependent on the microenvironment.[16]

In contrast to the mouse, human mast cells express only a single mast cell–specific chymase gene and thus exhibit less heterogeneity compared with their rodent counterparts. Human mast cells can express up to four tryptase genes (α, β, γ and δ), and CPA3. All human mast cells express tryptase β, and some also express chymase, CPA3, and cathepsin G. Human mast cells expressing tryptase in the absence (MC_T) or presence (MC_{TC}) of chymase, CPA3, and cathepsin G differ in anatomical distribution, with MC_T dominating in normal lung, intestinal mucosa, and bronchial and nasal epithelium and MC_{TC} dominating in skin, heart, uterus, and deep connective tissues.[16] These distributions can change with disease and inflammation. For example, the fraction of MC_{TC} among bronchial submucosal mast cells increases markedly with disease severity in asthma,[17] as do the numbers of MC_{TC} infiltrating the airway smooth muscle.[18] MC_{TC}s accumulate selectively in atherosclerotic plaques.[19] In allergic rhinitis and eosinophilic esophagitis, MC_T selectively accumulate in the inflamed mucosal surfaces. Thus while the extent and defining features of mast cell heterogeneity differ between human and rodent mast cells, there is considerable lineage plasticity in both species that reflect both tissue localization and context.

One of the most strongly conserved aspects of mast cell heterogeneity is the dependency, or lack thereof, on T cell–derived accessory growth factors. Athymic nude mice altogether lack MMCs in the intestinal mucosa despite the presence of wild-type levels of CTMCs in the same organ.[20] MC_T are lacking in the intestines of humans with T-cell immunodeficiencies, but MC_{TC} numbers are unaffected.[21] The differential requirement of mast cell subsets for T-cell help likely reflects the capacity of T cell–derived growth factors to synergize with SCF to promote mast cell proliferation and/or survival. This may be functionally important, since T cell–driven expansion of mast cells in the intestine is essential for the elimination of certain helminthic infections. Moreover, the capacity of IL-4 and IL-9 to alter expression of key mast cell effectors suggests that effector systems can be modulated by T cells to suit the needs of the host during immune responses.[9,11]

DEVELOPMENT AND DISTRIBUTION OF BASOPHILS

Basophils are the rarest of granulocytes in mammals (<1% of peripheral blood leukocytes) and take their name from the basophilic granules in their cytoplasm. The scarcity and short life span (~2.5 days in mice) of basophils hindered the study of their functions until recently, but there is now a growing appreciation for their roles in various immune processes that are distinct from the mast cell. These include FcεRI-dependent antigen capture and a potent and rapid capacity to generate and secrete IL-4 and IL-13.

Basophils are myeloid cells that originate and develop in bone marrow. Like mast cells, they develop from GMPs, which become BMCPs or BaPs, each of which has the capacity to develop into a mature basophil.[2,3] The basophil lineage commitment is regulated by the transcription factors GATA-2 and C/EBPα. Unlike mast cells, basophils exit bone marrow only after they have fully matured. Basophil development is promoted in the context of T-helper cell-2 (Th2) inflammatory responses. Cytokines, including IL-3 and the epithelial cell–derived cytokine TSLP, regulate basophil development as well.[3]

Ex vivo developmental studies with human progenitor cells have suggested a common origin for basophils and eosinophils. Denburg et al. showed that a fraction of clonal colonies arising from human peripheral blood progenitor cells contained both eosinophils and basophils, with some cells showing an intermediate phenotype with granules of both lineages.[22] Saito et al. showed that recombinant IL-3 stimulated the growth of both eosinophils and basophils from cord blood progenitors.[23] Human basophils express eosinophil cationic protein, major basic protein, Charcot-Leyden crystal protein, and eosinophil peroxidase, all of which are eosinophil granule markers,[24] consistent with a common origin of both cell types. A recent transcriptional profiling study confirmed that mouse blood basophils are more similar to eosinophils in their gene expression pattern than to any other cell type.[5] The shared lineage pathway for basophils and eosinophils explains why blood basophil counts are frequently elevated in conditions associated with increased blood eosinophil counts.

Basophil recruitment to the tissues is likely governed by a combination of protein and lipid chemoattractants. Human basophils express CCR1, CCR2, CCR3, CCR5, CXCR1, and CXCR2 and respond *ex vivo* to the corresponding ligands. Basophils strongly express chemoattractant receptor homologue expressed by Th2 cells (CRTH2), a specific receptor for prostaglandin D_2 (PGD_2). PGD_2 is an abundant product of activated mast cells, and CCR3 ligands (CCL11, CCL24, CCL26) are abundantly expressed by structural cells in allergic inflammation.[1] Thus all of these are good candidates to account for the accumulation of basophils that is frequently observed in rhinitis and asthma, and that is characteristic of allergen-induced late-phase responses.

BIOLOGICAL MEDIATORS PRODUCED BY MAST CELLS AND BASOPHILS

Mast cells and basophils can elaborate a broad array of potent biologically active mediators (see Table 23.1). Some of these products are stored preformed in cytoplasmic granules (*e.g.*, histamine, proteases, proteoglycans, certain cytokines), and others are synthesized *de novo* upon activation of the cell (*e.g.*, products of arachidonic acid oxidation via the cyclooxygenase [COX] or lipoxygenase pathways, platelet-activating factor [PAF], cytokines, chemokines, growth factors).[19] These mediators enable mast cells and basophils to orchestrate various functions in inflammation and host defense. Some, such as histamine and cysLT, are targets of the best-established drugs for the treatment of allergic diseases, validating the pathogenetic importance of mast cells, basophils, and their associated effector systems.

Preformed Mediators

Mediators that are stored in secretory granules can be rapidly (seconds to minutes) released into the extracellular environment upon stimulation. These mediators include histamine, proteoglycans, and (in the case of mast cells) proteases.[19]

Histamine is stored in the granules of mast cells and, to a lesser extent, basophils. Histamine is released within minutes of cell activation and mediates increased vascular permeability, smooth muscle contraction, mucus production, and increased heart rate and cardiac output, gastric acid secretion, itching, and sneezing. These actions are mediated via four types of histamine receptors (H1–4) on target cells.[1,3,16] H1 receptor antagonists are among the oldest antiallergy medications, and their efficacy supports the role of histamine in mediating the clinical signs and symptoms of rhinitis, allergic conjunctivitis, and urticaria.

KEY CONCEPTS

Mast Cell and Basophil Mediators

- Mast cells and basophils produce distinct but overlapping panels of mediators with diverse biological effects.
- Some mast cell and basophil mediators are preformed and stored in cytoplasmic granules (*e.g.*, histamine, proteoglycans, proteases). These can be released rapidly (seconds to minutes) upon cellular activation and degranulation.
- Mast cells and basophils produce lipid mediators (*e.g.*, PGD_2, LTC_4), which are derived from arachidonic acid and are newly synthesized upon cell activation.
- Mast cells and basophils can transcribe and secrete many cytokines, chemokines, and growth factors (see Table 23.1).

Proteases, including tryptases, chymase, and CPA3, comprise the major protein component of mast cell secretory granules. Tryptases are the most abundant proteases in the human mast cell and are found in basophils as well. Human tryptases are encoded by several different alleles, including α, β, γ, and δ. As noted previously, only a single mast cell–specific chymase gene exists in humans. CPA3 is expressed strongly by mast cells and weakly by basophils.[1,16] Tryptase β can activate protease-activated receptors, inducing fibroblast proliferation and collagen secretion. Both human tryptase β and its murine orthologue, mouse mast cell protease 6, can induce neutrophil recruitment in models of bacterial sepsis, suggesting prominent functions in innate immunity. Human mast cell chymase converts angiotensin 1 to angiotensin 2, suggesting a potential role in the control of blood pressure and cardiovascular homeostasis. CPA3 can neutralize endothelin 1 and the snake venom protein sarafotoxin.[25] These vasoprotective and antivenom functions are consistent with a role for resident cutaneous and perivascular mast cells vascular homeostasis and host defense. The functions of the mast cell proteases are reviewed comprehensively elsewhere.[16,26]

Newly Synthesized Mediators

Both mast cells and basophils utilize membrane phospholipid-derived arachidonic acid to synthesize eicosanoid mediators and can use the resultant lysophospholipids to generate PAF. These mediators are released within 15 minutes of activation. Both mast cells and basophils express 5-LO and LTC_4S, permitting them to generate LTC_4, the parent cysLT, from endogenous arachidonic acid. Mast cells, but not basophils, can generate smaller quantities of the dihydroxy leukotriene, LTB_4, which is a potent neutrophil chemoattractant. LTC_4 is converted extracellularly to LTD_4 (the most potent known bronchoconstrictor) and then to the stable metabolite LTE_4.[3,19] CysLTs are abundant in allergic inflammation, increasing microvascular permeability and causing bronchoconstriction. Mouse models suggest they play important roles in amplifying Th2 responses through several cellular targets. There are three cysLT receptors, respectively termed $CysLT_1R$, $CysLT_2R$, and GPR99.[27] Both $CysLT_1R$ antagonists and 5-LO inhibitors show efficacy in asthma, supporting the pathogenic role of the cysLT in this disease.

In contrast to basophils, mast cells generate substantial quantities of PGD_2, a product of successive metabolism of arachidonic acid by COX-1 or COX-2 and hematopoietic PGD_2 synthase, an enzyme that is expressed more strongly by mast cells than any other cell type. PGD$_2$ binds to two specific receptors, termed DP$_1$ and CRTH2 (also known as DP$_2$). Once released, PGD$_2$ elicits vasodilation (via DP$_1$) and bronchoconstriction (via the actions of a putative metabolite at the T prostanoid receptor). Additionally, PGD$_2$ is potently chemotactic for basophils, eosinophils, Th2 cells, and group 2 innate lymphoid cells (ILC2), each of which express CRTH2.[19] PGD$_2$ induces cytokine generation (IL-4, IL-5, IL-13) from both human Th2 cells and ILC2 and could, therefore, amplify type 2 inflammation in a polyclonal fashion. There are several CRTH2-selective antagonists in clinical development for the treatment of asthma and other allergic diseases.

Cytokines, Chemokines, and Growth Factors

Basophils and mast cells produce a wide variety of cytokines and chemokines that recruit and activate other cells (see Table 23.1). Mast cells generate tumor necrosis factor-α (TNF-α), IL-6, and IL-1β, which is consistent with their sentinel role in innate antimicrobial immunity (Chapter 3). Mast cells can generate the Th2 cytokines IL-5 and IL-13, a functional property that is induced when the cells are exposed to IL-4. TNF-α can be stored in granules constitutively as well as produced *de novo* in response to activation.[19] Activated mouse and human mast cells express a large range of both CC and CXC chemokines, likely reflecting their role in facilitating the recruitment of bloodborne effector cells. A recent study demonstrated that mast cell progenitors in the intestinal mucosa are an important source of IL-9 and IL-13 in a model of food allergy.[28] Basophils can secrete large quantities of IL-4 (and IL-13) rapidly, contributing to Th2 immune responses in a variety of settings. These observations suggest that basophils are more specialized than mast cells for a role in type 2 immunity.

MECHANISMS OF ACTIVATION OF MAST CELLS AND BASOPHILS

KEY CONCEPTS

Mechanisms of Mast Cell and Basophil Activation

- Mast cells and basophils have cell-surface high-affinity immunoglobulin E (IgE) receptors (FcεRI) that are tetrameric and thus fully functional, unlike the trimeric FcεRI found on other cells (*e.g.,* dendritic cells [DCs] and eosinophils).
- IgE-mediated activation of mast cells and basophils occurs when multivalent antigen cross-links IgE antibodies bound to adjacent FcεRI receptors, resulting in degranulation and *de novo* mediator production and release.
- Mast cells and basophils can be activated by non–IgE-mediated mechanisms (anaphylatoxins, Toll-like receptor [TLR] ligands, certain cytokines, drugs, venoms, etc.).
- Some types of non–IgE-mediated activation cause a more compartmentalized activation (*e.g.,* cytokine production without degranulation caused by activation of mast cells or basophils by interleukin-33 [IL-33] and thymic stromal lymphopoietin [TSLP]).

FcεRI-Mediated Activation (and Inhibition of IgE-Dependent Activation)

Both mast cells and basophils express the high-affinity IgE receptor, FcεRI. The FcεRI expressed on these cells is unique in that it has a tetrameric structure consisting of a single IgE-binding αα chain, a single β chain, and two identical disulfide-linked γ chains. Human hematopoietic cells other than mast cells and basophils (*e.g.,* eosinophils and dendritic cells [DCs]) express FcεRI in the absence of the β chain. The β chain amplifies signaling through the receptor, and the tetrameric form on mast cells and basophils is thus fully functional.[29] Activation occurs when adjacent FcεRI receptors, occupied by receptor-bound IgE, are cross-linked by multivalent antigen, stimulating degranulation and release of preformed mediators as well as *de novo* production and release of additional mediators described above.[1,3,19,29] The signaling cascades that are initiated by FcεRI cross-linking are reviewed in detail elsewhere.[30]

At the ultrastructural level, stimulation of sensitized human basophils with specific antigen provokes fusion of the membranes enveloping individual cytoplasmic granules with the plasma membrane (Fig. 23.2). This results in the release of the granules' contents via multiple narrow communications between single granules and the cell surface. IgE-dependent degranulation of human lung mast cells results in the fusion of granule membranes with the plasma membrane (Fig. 23.3). However, in this cell type the first ultrastructural changes detectable in the stimulated cells are granule swelling, followed by fusion of individual granule membranes forming interconnecting chains of swollen granules; histamine release is initiated by the opening of these channels to the exterior through multiple narrow points of fusion with the plasma membrane.[31]

Non–IgE-Mediated Activation

Mast cells and basophils can be activated through a variety of non–IgE-mediated mechanisms. Both mast cells and basophils express receptors for complement fragments C3a and C5a and release histamine in response to these complement components. The C5a receptor (CD88), in particular, is expressed in human lung on MC$_{TC}$s, and elevated C5a concentrations have been measured in induced sputum from subjects with asthma.[2,19] Mast cells and basophils can be activated by microbial constituents via Toll-like receptors (TLRs), many of which are expressed on their surface and internal membranes.[19,29] Both cell types respond to the barrier-derived cytokines IL-33 and TSLP, both of which drive Th2-type immune responses. Each cytokine can directly activate mast cells and, in some cases, amplify IgE-mediated responses.[29] In contrast to the complement fragment-dependent activation, IL-33 and TSLP fail to induce degranulation but are powerful inducers of cytokine production, particularly when they are provided together. Thus non–IgE-dependent activation mechanisms tend to elicit a response from mast cells that is more compartmentalized than FcεRI-mediated activation.

A variety of drugs, such as opiates and muscle relaxants, can directly activate mast cells.[19,29] Recently, the murine counterpart of the human G protein-coupled receptor, MAS-related G protein-coupled receptor X2 [MRGPRX2], was found to be necessary to trigger mast cell degranulation *in vivo* response to a range of basically charged drugs, as well as compound 48/80, a classic secretagogue for mast cells.[32] The potential that certain drugs possess in their ability to mimic components of innate immunity to trigger mast cells represents a major conceptual step forward toward understanding drug-induced pseudoallergic reactions that do not involve specific IgE.

Basophil activation is promoted by IL-3, and other cytokines, including IL-18 and IL-33, can activate basophils and enhance their effector functions.[3] Basophils can be directly activated by protease allergens, including the house dust mite protease Derp1.[3] Finally, both mast cells and basophils respond to PAF with

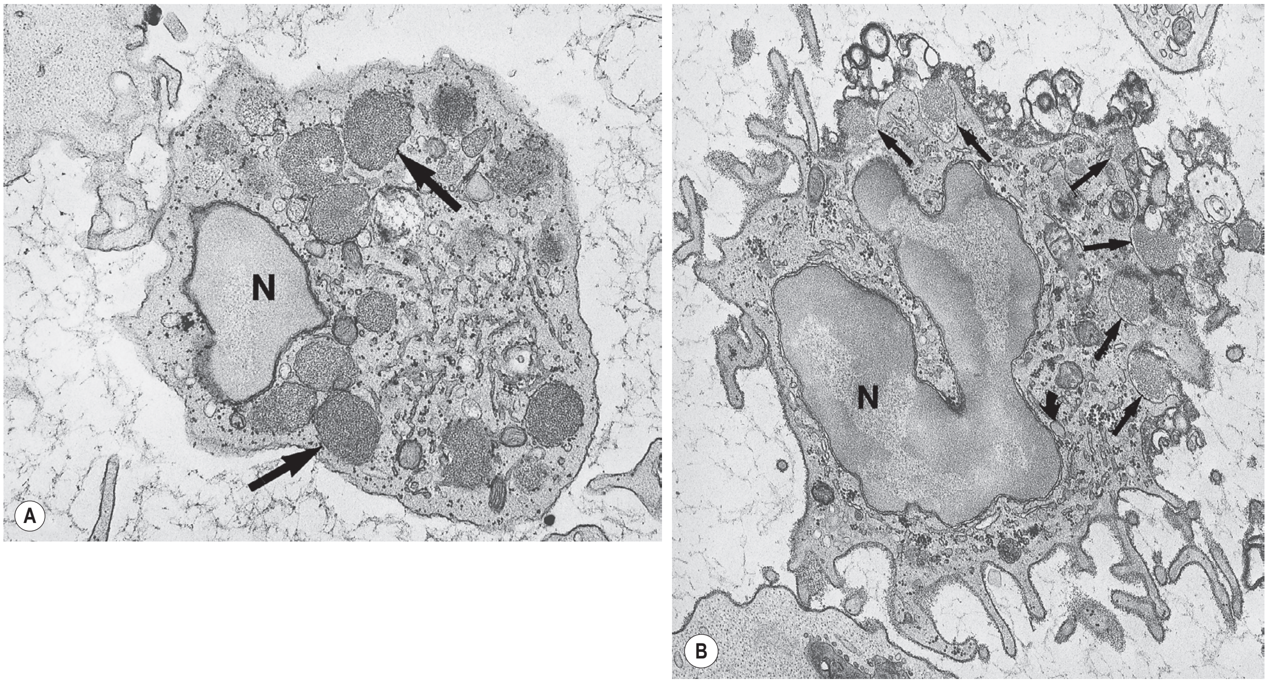


FIG 23.2 (A) Transmission electron micrograph of a human basophil in a preparation of peripheral blood leukocytes obtained by separation over Ficoll-Hypaque. All of the cytoplasmic granules (some indicated by solid arrows) contain particulate electron-dense material. N, nucleus. (Original magnification ~ × 19800). (B) A human basophil 2 minutes after exposure to antigen *in vitro.* The cell exhibits extrusion of granules from six separate sites on the plasma membrane (small arrows). At this time after cell stimulation, particle-filled granules retain their shape and characteristic structure even after exposure to extracellular milieu. Cationized ferritin coats the cell surface and enters culs-de-sac that contain exteriorized granules. The cell exhibits no fully intracytoplasmic typical basophilic granules, but one of the smaller kind of granules (curved arrow) can be observed in the perinuclear region. N, nucleus. (Original magnification ~ × 19200). (From Dvorak AM, Newball HH, Dvorak HF, Lichtenstein LM. Antigen-induced IgE-mediated degranulation of human basophils. Lab Invest 1980;43:126–139, with permission from Nature Publishing Group Ltd.)

activation and histamine release, and PAF concentrations are elevated in the plasma after anaphylactic reactions.[33] These findings suggest a mechanism by which endogenous lipid metabolism can magnify the physiological effects of mast cell activation in anaphylaxis.

MAST CELLS AND BASOPHILS IN DISEASE AND HOST DEFENSE

CLINICAL RELEVANCE

Mast Cells and Basophils in Health and Disease

- Mast cells and basophils are primary effector cells in atopic disorders (anaphylaxis, asthma, allergic rhinitis, atopic dermatitis).
- Antigen-specific, immunoglobulin E (IgE)-mediated mast cell activation results in an immediate clinical reaction and, in some cases, contributes to a late-phase reaction.
- Medications that block mast cell and basophil mediators are mainstays of treatment for atopic diseases.
- Mast cells and basophils contribute to host defense against parasites and bacterial pathogens.
- Mast cells may play a role in other human diseases, including atherosclerosis, autoimmune diseases, and burns and in responses to venom toxins.

Allergic Disease

Anaphylaxis

Anaphylaxis is an acute, life-threatening, multisystem reaction that results from the precipitous release of mast cell and basophil mediators into the circulation (Chapter 42). These mediators cause the cardinal manifestations of hypotension, bronchoconstriction, intestinal colic, urticaria, and, in some instances, disseminated intravascular coagulation. Anaphylactic reactions can be triggered by foods, insect stings, and medications, among other factors. Anaphylaxis, like mast cell and/or basophil activation, can be immunological or nonimmunological.

The classic mechanism of anaphylaxis in human disease is IgE mediated, with an interaction between antigen and antigen-specific IgE bound to FcεR1 on mast cells and/or basophils. In this instance, anaphylaxis can be considered a severe form of type I hypersensitivity. Production of histamine and leukotrienes in response to IgE-mediated activation of both mast cells and basophils induces smooth muscle contraction and contributes to the symptoms and physiology of systemic anaphylaxis.[3] Besides the classically described IgE-mediated type of anaphylaxis, IgG-driven, PAF-mediated anaphylaxis has also been described in mouse models. Mast cells are dispensable for this type of anaphylaxis, whereas basophils (among other lineages) have been reported as PAF-producing effector cells.[3] PAF levels are elevated

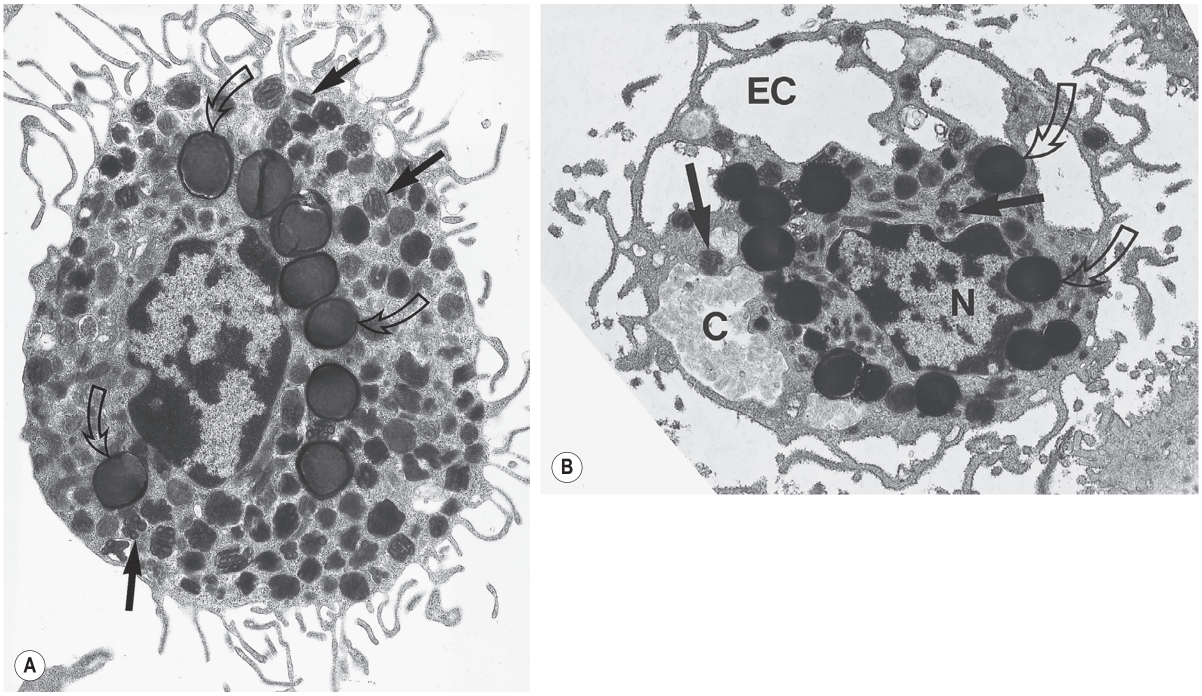


FIG 23.3 (A) Mast cell purified from human lung. The cell contains many cytoplasmic granules with scroll-like substructural elements (solid arrows) and eight large non–membrane-bound lipid bodies (open arrows). The plasma membrane has prominent folds. N, nucleus. (Original magnification ~ × 12 900). (B) An isolated human lung mast cell 10 minutes after exposure to anti-IgE *in vitro.* Some degranulation channels (C), formed by fusion of membranes surrounding individual cytoplasmic granules, contain altered granule matrix; others (EC) are empty. Cationized ferritin stains the plasma membrane and the membranes of some empty degranulation channels (EC). The membranes lining other channels (C) are unstained. A few unaltered scroll-containing granules (solid arrows) remain. Numerous lipid bodies are also present (open arrows). The cytoplasm contains prominent filaments. N, nucleus. (Original magnification ~ × 9 100). (From Galli SJ, Dvorak AM, Dvorak HF. Prog Allergy 1984;34:1, with permission from Ann M. Dvorak.)

in the serum of patients with severe peanut anaphylaxis, suggesting that PAF may serve to amplify the physiological manifestations of IgE-dependent anaphylaxis. Immune complex–mediated complement activation has been proposed as a mechanism for some anaphylactic reactions to drugs. Nonimmunological anaphylaxis can be caused by complement activation in the absence of immune complexes or by direct activation of mast cells, often by drugs, such as opiates or vancomycin.[34] As noted above, some of these (particularly muscle relaxants) may result from drugs that activate MRGPRX2 or related receptors.[32]

Asthma

Asthma is a chronic inflammatory disease of the airways characterized by variable airflow obstruction and airway hyperresponsiveness (AHR) (Chapter 41). The most direct evidence for a mechanistic contribution of mast cells to allergic asthma comes from allergen challenge studies. In sensitized individuals, allergen challenge causes an increase in mast cell granule contents in bronchoalveolar lavage (BAL) fluid, including histamine, tryptase, and cysLTs. About 60% of patients having an immediate, early-phase asthmatic response will have a late-phase reaction beginning 3–5 hours later, peaking around 8 hours, and resolving in about 24 hours. Mast cells contribute to both the early-phase and the late-phase asthmatic response, via the release of histamine and production of proinflammatory cytokines, leukotrienes, and prostanoids.[19] The involvement of mast cells in early- and late-phase responses is supported by earlier studies using mast cell stabilizing cromone drugs, as well as anti-IgE.

Immunohistological studies provide additional support for the role of mast cells in asthma. Bronchial biopsies from subjects with asthma exhibit mast cells localizing to the airway smooth muscle bundles, a region generally lacking mast cells in biopsies from healthy controls and from subjects with nonasthmatic eosinophilic bronchitis.[19] Smooth muscle mast cell density correlated strongly with disease severity and with hyperresponsiveness to methacholine,[35] a cardinal feature of asthma. Notably, these increases were observed irrespectively of atopic status of the study subjects. In biopsy samples from subjects with severe asthma, mast cells exhibit substantially higher proportions with positive immunostaining for chymase than in specimens from those with mild asthma. Thus both the distribution and phenotype of mast cells is altered in the airways of patients with severe asthma. The mechanisms responsible for these alterations, and the factors responsible for driving mast cell activation in severe disease, particularly in individuals without atopy remain to be determined.

Mast cells may play an especially important role in aspirin-exacerbated respiratory disease (AERD), a distinctive disorder in which asthma, severe rhinosinusitis, and nasal polyposis are associated with idiosyncratic respiratory reactions to aspirin and other drugs that inhibit COX-1. Although tissue mast cell burden in AERD does not exceed that seen in aspirin-tolerant control patients with asthma, baseline levels of serum tryptase and urinary PGD_2 metabolites are higher, potentially reflecting a loss of homeostatic PGE_2 from the surrounding tissue. These values increase markedly with aspirin challenges, and treatment with cromolyn or nedocromil blocks the characteristic reactions, supporting their mast cell dependency.[36] Notably, a mouse model suggests that mast cell activation in response to aspirin challenge depends on IL-33 rather than the allergen.[36] It is tempting to speculate than innate cytokines could account for mast cell activation in other nonatopic disorders.

There is growing evidence that basophils play a role in the pathogenesis of asthma as well. Basophils are enriched in postmortem human lung tissue from patients who have died as a result of asthma as well as in bronchial biopsy specimens of patients with asthma.[1,3] In mouse models of allergic airway inflammation, basophils have been shown to promote Th2 cytokine responses and pathological airway inflammation.[3] The specific functions of basophils relative to asthma are not yet known, though it seems likely that they could contribute Th2 cytokines and cysLTs.

Allergic Rhinitis

Allergic rhinitis is an inflammatory disease of the nasal mucosa that occurs as a reaction to airborne allergens (Chapter 41). Increased numbers of activated MC_T mast cells are found in the nasal epithelium of patients with allergic rhinitis, and these numbers increase further with allergy provocation challenge or during seasonal disease exacerbations.[34] Basophils are present in nasal washes from patients with allergic rhinitis.[3] In addition, nasal challenges with allergens lead to the release of histamine and cysLTs, which are preformed mediators of both mast cells and basophils.[34]

Atopic Dermatitis

Atopic dermatitis (AD) is a chronic, relapsing inflammatory skin condition that is commonly diagnosed in young children and often precedes the development of allergic rhinitis and asthma (Chapter 44). The pathogenesis of AD remains incompletely understood but is likely multifactorial with contributions from skin barrier dysfunction, innate and adaptive immune dysregulation, and environmental and genetic factors. Increased numbers of mast cells can be seen in skin lesions from human AD and mouse AD models, but their role in the pathophysiology of the disease is not known.[34]

Basophils have been identified in lesional skin of patients with AD, but their precise role in human disease remains unclear. In murine models of AD-like skin disease, significant TSLP-dependent basophilia has been noted in lesional skin. TSLP is also associated with human AD, so it is possible that TSLP-induced basophil populations could play a role in human disease as well.[1,3]

Mast Cells and Basophils in Immunity

Although mast cells are found throughout most tissues, they are notably located at interfaces with the environment, such as skin and the mucosal surfaces. This suggests a potential role in pathogen recognition and host defense. Additionally, mast cells have the capacity to increase in numbers during mucosal inflammatory responses, suggesting that their effector capabilities can be magnified to suit the context. The first associations of mast cells with host defense included observations of mast cells clustering around sites of parasitical infestation. Subsequent studies have demonstrated a role for mast cells in mouse models of helminth infection in the gut as well as in parasitical skin infections.[37]

Production of IL-4 and IL-13 by mast cells and basophils may play an important role in early parasitical infections, leading to IgE production, secretion of effector molecules, activation of smooth muscle cells, and increased mucus production, among other effects.[1] Furthermore, although expulsion of hookworm during primary infection in mice depends on mast cells, basophils are necessary to expel hookworms during secondary infection, possibly because of activation by helminth-specific antibodies.[1] Intestinal helminth infections typically elicit a T cell-dependent expansion of mast cells and basophils in the involved mucosal site.[8] In that regard, mast cells and basophils can be viewed as components of the effector arm of adaptive immunity.

Mast cells are important in the generation of immune responses to bacterial pathogens. For example, mast cell deficient mice have increased mortality in a model of *Escherichia coli*–induced peritonitis. Mast cells have shown protective roles in *K. pneumoniae* lung infections and *Pseudomonas* skin infections in mice.[37] Mice lacking mast cells as a result of loss-of-function *Kit* and *SCF* mutations are significantly more susceptible to lethal gram-negative sepsis than are congenic mast cell–sufficient controls. Reconstitution of the tissues with wild-type mast cells restores protective immunity. Both proteases and TNF-α are linked to the protection from bacterial infections conveyed by mast cells.

Mast Cells in Other Diseases

Mast cells may play a role in other disease processes as well. Increased numbers of mast cells have long been noted in human atherosclerotic lesions compared with healthy control tissue. Mast cell functions within these plaques are not completely understood, although mouse models suggest that atherosclerotic changes are promoted by release of proinflammatory cytokines by mast cells.[38] Mouse models of autoimmunity, including autoantibody-mediated arthritis[39] and multiple sclerosis,[40] have demonstrated a pathological role for mast cells. Skin mast cells are activated rapidly following thermal trauma, and the release of mast cell proteases is crucial for development of the inflammatory response in burn lesions.[41]

Recent work in mice has shown that mast cells protect from the detrimental effects of arthropod or reptile venoms via the degradation of venom toxins by mast cell proteases. Furthermore, acquired Th2 immune responses associated with IgE production can protect mice from challenge with potentially lethal doses of venoms from honeybee or the Russell viper.[42]

Mast Cell Roles in Normal Physiology

In addition to their immune functions, mast cells are thought to have normal physiological functions as well. Data from mouse studies suggest a role for mast cells in intestinal barrier function and homeostatic intestinal epithelial migration.[43] Human uterine tissue has increased density of mast cells during pregnancy, where they may mediate contractility,[44] in addition to contributing to local immunity.

MASTOCYTOSIS AND OTHER MAST CELL DISORDERS

A variety of human disorders are associated with excessive numbers of mast cells, increased activation of mast cells, or both.

Mastocytosis can be limited to skin or have widespread systemic effects. Mast cell disorders cause a broad range of symptoms that can mimic a variety of other conditions, often making diagnosis a challenge.

Epidemiology

Mastocytosis can occur at any age, although it tends to have a more benign and transient course in children and often is limited to cutaneous manifestations. In a Polish case series of 100 children with cutaneous mastocytosis (CM), 73% had onset of the disease by age 6 months and 94% within the first year of life.[45] Most cases of systemic mastocytosis (SM) are diagnosed in middle age. The exact prevalence of mastocytosis is unknown, but an estimate from a recent Danish population-based study is approximately 1:10 000.[46] With rare exceptions, mastocytosis does not appear to be inherited.

Pathogenesis

The manifestations of mastocytosis can result from mast cell mediator release (both chronic and episodic) and from excessive accumulation of mast cells in one or more tissues. Although the molecular pathogenesis is incompletely understood, mastocytosis is frequently associated with somatic gain-of-function mutations in KIT (CD117). The most common of these mutations is Asp816Val, or D816V.[47] KIT-activating mutations lead to SCF-independent activation. Mast cells are thought to accumulate in tissues because of clonal expansion and apoptotic defects of KIT-mutated mast cells in mastocytosis.[6]

Clinical Features

The clinical signs and symptoms of mastocytosis are varied and can affect many different organ systems (Fig. 23.4). Clinical features can be broadly categorized as cutaneous, related to mast cell mediator release, or resulting from organ infiltration by mast cells.

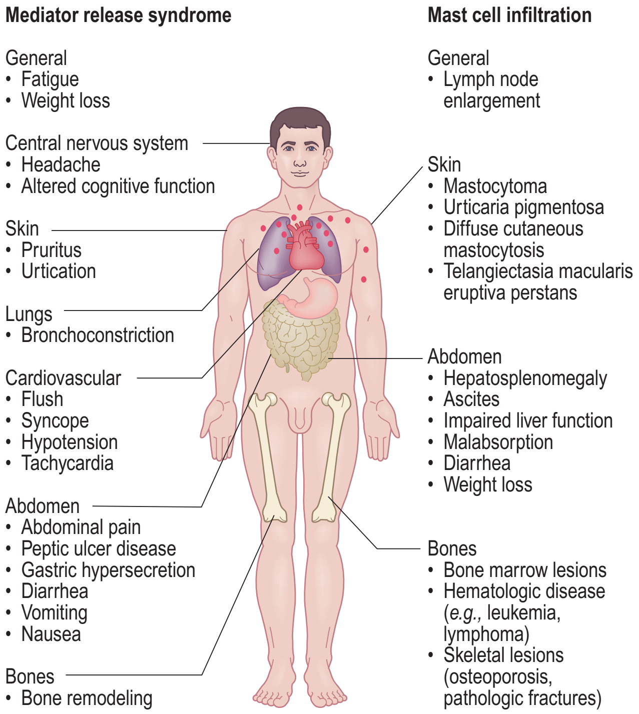

Mediator release syndrome
General
• Fatigue
• Weight loss
Central nervous system
• Headache
• Altered cognitive function
Skin
• Pruritus
• Urtication
Lungs
• Bronchoconstriction
Cardiovascular
• Flush
• Syncope
• Hypotension
• Tachycardia
Abdomen
• Abdominal pain
• Peptic ulcer disease
• Gastric hypersecretion
• Diarrhea
• Vomiting
• Nausea
Bones
• Bone remodeling
Mast cell infiltration
General
• Lymph node enlargement
Skin
• Mastocytoma
• Urticaria pigmentosa
• Diffuse cutaneous mastocytosis
• Telangiectasia macularis eruptiva perstans
Abdomen
• Hepatosplenomegaly
• Ascites
• Impaired liver function
• Malabsorption
• Diarrhea
• Weight loss
Bones
• Bone marrow lesions
• Hematologic disease (e.g., leukemia, lymphoma)
• Skeletal lesions (osteoporosis, pathologic fractures)

FIG 23.4 Clinical manifestations of mastocytosis.

Skin findings may be seen in both CM and SM, and the most common skin-related symptom in patients with mastocytosis is pruritus. Accumulations of mast cells in skin can present as a variety of lesions. Urticaria pigmentosa (UP), also called *maculopapular cutaneous mastocytosis* (MPCM), is the most common skin manifestation of mastocytosis in both adults and children. Lesions of UP appear as small, tan to reddish-brown macules or slightly elevated papules (Fig. 23.5A–C). UP is present more commonly in the less aggressive forms of SM.

Diffuse cutaneous mastocytosis (DCM) is characterized by edema and increased skin thickness, with or without a yellowish-brown coloration. DCM is rare, accounting for 1–3% of CM. Cutaneous mastocytomas are well-demarcated flat or slightly elevated lesions that may have a yellow or red-brown coloration. Mastocytomas can be solitary or multiple and are typically 2–5 cm in size. The least frequent form of CM (<1% of cases) is telangiectasia macularis eruptive perstans (TMEP). TMEP typically presents in adulthood and consists of tan/brown macules with telangiectasias.

Lesions of UP, DCM, and cutaneous mastocytomas display the Darier sign: localized urtication and redness of lesions following rubbing, scratching, or stroking of skin. Mastocytomas can also precipitate severe systemic symptoms when rubbed. In both CM and SM, mast cell mediator release can occur either chronically or episodically, resulting in a variety of symptoms, including flushing, pruritus, shortness of breath, nausea, vomiting, abdominal pain, diarrhea, hypotension, syncope, fatigue, and headache. Anaphylaxis can be seen in both CM and SM. Interestingly, urticaria and angioedema are uncommon in mastocytosis.

Gastrointestinal (GI) symptoms are particularly common in SM, with diarrhea and abdominal pain reported in up to 80% of patients with SM.[48] Neuropsychiatric symptoms, including depression, irritability, increased somnolence, and problems with memory and concentration, are often observed in mastocytosis, although the mechanisms are not well-understood.[49] Osteopenia and osteoporosis are observed in a subset of patients with SM. This may be attributed to the effects of mast cell mediators, including histamine, tryptase, heparin, IL-6, TNF-α, and transforming growth factor-β (TGF-β), on bone turnover.[50]

Mast cell mediator release can be triggered by a variety of factors, including medications (including nonsteroidal antiinflammatory drugs [NSAIDs], opiates, muscle relaxants, and contrast media), alcohol, surgical or endoscopic procedures, infections (viral, bacterial, parasitical), physical factors (exercise, friction, extremes of temperature), and emotional stress. Patients with mastocytosis are particularly susceptible to anaphylaxis during allergic reactions to *Hymenoptera* stings. Severe allergic reactions to *Hymenoptera* stings should prompt an evaluation for mastocytosis, and patients with mastocytosis should be offered venom immunotherapy.[51]

Infiltration and/or proliferation of mast cells in organs other than the skin distinguish SM from CM. The organ systems most often affected in SM include bone marrow, GI tract, lymph nodes, liver, spleen, and cortical bone. The most frequent hematological abnormality in SM is anemia, and eosinophilia is also common.[52] The lymph nodes and spleen are commonly infiltrated in all subtypes of SM, manifesting as lymphadenopathy or splenomegaly in many patients.

Classification

The World Health Organization (WHO) has defined seven variants of mastocytosis (Table 23.2). CM is limited to the skin and more

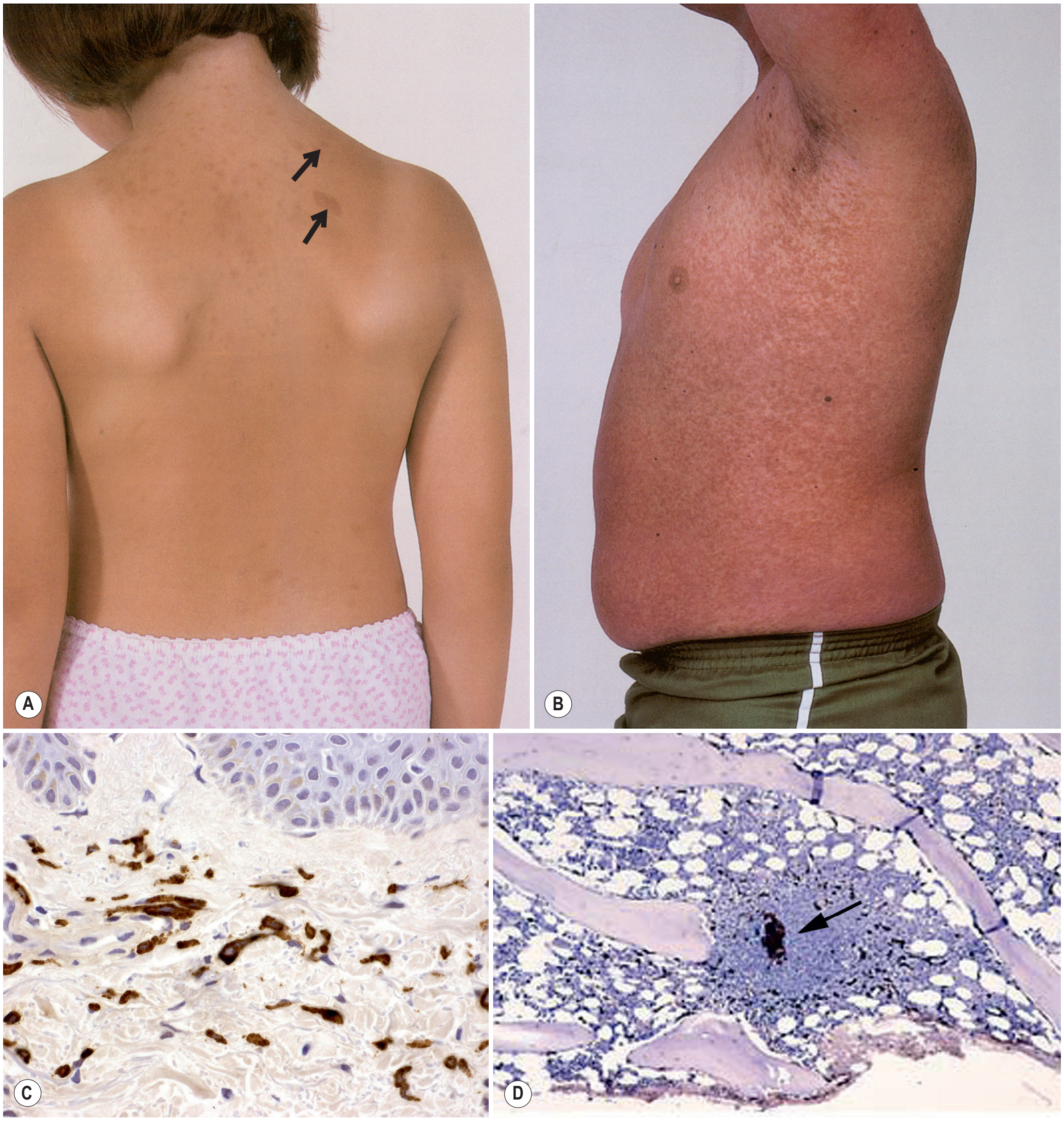


FIG 23.5 **Cutaneous Mastocytosis and Histopathology.** Urticaria pigmentosa is the most common form of cutaneous mastocytosis. In childhood, lesions are disseminated and consist of well-demarcated hyperpigmented macules (*e.g.,* arrows) (A). In adults, lesions may be numerous with less well-demarcated brownish-red macules and papules (B). Histopathology of cutaneous mastocytosis shows many mast cells containing abundant tryptase immunoreactive cytoplasmic granules in the papillary dermis (antitryptase, AA1 clone, Dako; original magnification ~ × 400) (C). Bone marrow pathology of indolent mastocytosis is characterized by paratrabecular lymphoid nodule containing small, well-differentiated lymphoid cells around substantial numbers of fusiform cells with prominent granules in the tryptase stain (arrow) (antitryptase, AA1 clone, Dako; original magnification ~ × 250) (D). (Courtesy of Cem Akin.)

commonly found in children. SM can include findings in a variety of organ systems, with or without skin involvement, and is defined by a set of major and minor diagnostic criteria[53] (Table 23.3). The different categories of SM are distinguished by clinical signs and symptoms and/or pathological findings, including mast cell burden and involvement of non–mast cell lineages,[52] and have different prognostic implications, as described below.

The term *monoclonal mast cell activation syndrome* (MMAS) refers to the condition in patients who have clinical symptoms consistent with mast cell activation and meet one or two of the minor criteria for SM but do not fully meet the diagnostic criteria. The term *mast cell activation syndrome* (MCAS) describes patients presenting with clinical symptoms of mast cell activation affecting at least two organ systems, but not fulfilling the criteria for

TABLE 23.2 World Health Organization (WHO) Classification of Mastocytosis

Cutaneous mastocytosis
Indolent systemic mastocytosis
Systemic mastocytosis with an associated clonal hematologic non–mast cell lineage disease
Aggressive systemic mastocytosis
Mast cell leukemia
Mast cell sarcoma
Extracutaneous mastocytoma

TABLE 23.3 Diagnosis of Systemic Mastocytosis Based on World Health Organization (WHO) Criteria[a]

Major Criterion

Multifocal, dense infiltrates of mast cells (≥ 15 mast cells in aggregates) detected in sections of bone marrow and/or other extracutaneous organ(s)

Minor Criteria

1. More than 25% of mast cells in biopsy sections of bone marrow or other extracutaneous organs are spindle shaped or display atypical morphology, or, of all mast cells in bone marrow aspirate smears, > 25 % are immature or atypical
2. Detection of an activating point mutation at codon 816 of c-*KIT* in blood, bone marrow, or another extracutaneous organ
3. Mast cells in the bone marrow, blood, or other extracutaneous organs express CD2 and/or CD25 in addition to normal mast cell markers
4. Serum total tryptase persistently exceeds 20 ng/mL

[a] One major and one minor, or three minor, criteria are needed for the diagnosis of systemic mastocytosis.

Data from Horny HP, Metcalfe DD, Bennett JM, et al. Mastocytosis In: Swerdlow SH, Campo E, Harris NL, Jaffe ES, Pileri SA, Stein H, Thiele J, Vardiman JW, editors. WHO classification of tumors of hematopoietic and lymphoid tissues. Lyon, France: IARC Press; 2008: p. 54–63.

anaphylaxis or for MMAS.[54] Patients who have anaphylaxis of unknown etiology are more properly diagnosed with idiopathic anaphylaxis.

Diagnosis

For all patients with mastocytosis, the diagnostic process begins with a thorough history and physical examination. The diagnosis of CM is made based on the appearance of characteristic skin lesions and may be confirmed with skin biopsy showing mast cell infiltrates in the skin. In adults with cutaneous findings, systemic disease must be ruled out.

The diagnostic process for patients with suspected mast cell disease typically includes measurement of serum tryptase. If the tryptase is found to be elevated during an anaphylactic episode or other symptomatic event, the measurement should be repeated at least 24 hours after symptoms have resolved. If basal tryptase is elevated (>20 ng/mL), there is a high probability of SM.[52] Another blood test that may be useful, particularly in patients with borderline tryptase levels or an indeterminate clinical presentation, is a screening test for the D816V mutation in peripheral blood leukocytes. Both elevated tryptase and the D816V mutation are minor criteria for mastocytosis. Other blood tests that should be performed to evaluate for specific organ involvement include complete blood count and liver function tests.

Most of the diagnostic criteria for mastocytosis (see Table 23.3) depend on bone marrow histopathological findings. Bone marrow biopsy is therefore an essential part of the diagnostic process. It should be considered for adult patients with typical skin lesions and as a confirmatory test in patients with elevated tryptase levels. Typical findings on bone marrow biopsy in mastocytosis include multifocal dense mast cell aggregates (>15 mast cells per cluster) and abnormal morphological features of mast cells (Fig. 23.5D). Such features include spindle shapes, bi- or multilobed nuclei, and expression of CD2 or CD25.[52]

The differential diagnosis of SM includes a variety of disorders with similar clinical manifestations, such as MMAS and MCAS (described above), hereditary/acquired angioedema, carcinoid syndrome, and pheochromocytoma.

Treatment

THERAPEUTIC PRINCIPLES

Approach to Treatment of Mastocytosis

- Avoidance of symptomatic triggers
- Control of symptoms
  - Antihistamines (H_1 and H_2 receptor antagonists)
  - CysLT receptor antagonists
  - Aspirin
  - Disodium cromoglycate
  - Epinephrine
- Management of osteoporosis
- Cytoreductive therapy for aggressive disease
  - Imatinib
  - Other tyrosine kinase inhibitors
- Management of associated hematological disorders

There is no cure for mastocytosis, and treatment is focused on control of symptoms. All patients with mastocytosis should avoid exposures to symptomatic triggers. In addition, patients should have epinephrine autoinjectors available at all times to treat possible anaphylaxis. Pharmacological treatment of mastocytosis includes drugs that fall into three main categories: *(i)* those that block or interfere with mast cell mediators, *(ii)* those that prevent mast cell activation, and *(iii)* those that affect mast cell survival and proliferation.

Type 1 histamine (H_1)-receptor antagonists are not only particularly useful for the control of cutaneous symptoms, including pruritus and flushing, but may also aid in the management of headache and neuropsychiatric symptoms (headache, depression, cognitive impairment). Type 2 histamine (H_2)-receptor antagonists are used in the treatment of GI symptoms, such as abdominal pain, cramping, diarrhea, heartburn, nausea, and vomiting. CysLT receptor antagonists are used in conjunction with antihistamines as a second-line treatment for cutaneous symptoms. Aspirin may be helpful for some patients with flushing, particularly those with high levels of urinary prostaglandin D_2 or 11β-$PGF_{2\alpha}$ but must be used with caution, as NSAIDs may trigger anaphylaxis in some patients.[52]

Disodium cromoglycate (cromolyn) is a weak *in vivo* inhibitor of mast cell activation but is used orally to reduce GI and neuropsychiatric symptoms.[52] Certain H_1-receptor antagonists, including ketotifen and rupatadine, are reported to block mast cell activation.

Osteoporosis in patients with SM is typically treated in the same manner as in patients in the general population. Bisphosphonate drugs are the first line of treatment, along with appropriate intake of calcium and vitamin D. Interferon-α (IFN-α) has been reported to be helpful as a second-line agent in refractory disease.[52]

Cytoreductive therapy has a role in the treatment of aggressive systemic mastocytosis (ASM) and mast cell leukemia, although data from large clinical trials are lacking. Imatinib, a tyrosine kinase inhibitor, has been effective in reducing mast cell burden in only a small subset of patients who lack the D816V mutation in KIT. Although clinical responses to other tyrosine kinase inhibitors have been largely disappointing,[52] a very recent trial of midostaurin, which is active on D816V, showed substantial promise in patients with advanced SM.[55] More studies are needed to determine those patients most likely to benefit.

Additional treatment of patients with SM and associated clonal hematological non–mast cell lineage disease (AHNMD) depends on the associated hematological disorder. The general approach is to treat AHNMD as if SM were not present and to treat the SM as if the AHNMD were not present, while closely monitoring patients for mast cell activation symptoms during therapy.[56]

Prognosis

Each category of mastocytosis carries a different prognosis. Pediatric patients with CM have complete resolution by adulthood in the majority of cases. Patients with indolent systemic mastocytosis (ISM) generally have a normal life expectancy, and ISM rarely evolves into more aggressive categories. However, median survival in other forms of SM is diminished (24 months in SM-AHNMD, 48 months in ASM). Mast cell leukemia is the rarest and most severe form of mast cell disease, with median survival of only 2 months.[57]

SUMMARY AND FUTURE RESEARCH DIRECTIONS

ON THE HORIZON

- Newer mouse models of mast cell deficiency and of inducible basophil deficiency will help:
  - Clarify the role of mast cells in normal physiology
  - Improve understanding of the role of mast cells in atopic and nonatopic diseases
- Translational research will help:
  - Improve understanding of the mechanisms of mast cell- and basophil-mediated diseases
  - Identify additional mutations in *c-KIT* and other genes that are associated with mastocytosis
- Clinical research is needed:
  - To improve diagnostic tools for atopic diseases and mast cell disorders
  - To improve treatment of and eventually cure asthma and atopic diseases
  - To improve treatment of and eventually cure mastocytosis

The mouse models that have been classically used to study mast cell functions are based upon *Kit* mutations that result in other defects besides a lack of mast cells. Newer transgenic models of mast cell deficiency have been developed in recent years and will help clarify mast cell roles in both normal physiology and in disease processes by avoiding these confounding factors. Mouse models of inducible basophil deficiency have recently been developed and will be important in improving our understanding of basophil functions.[58]

There is a great need for better understanding of molecular mechanisms of mast cell–related human diseases for improved diagnosis and treatment in the long term. For most atopic diseases and for mastocytosis, the current standard of care consists of avoidance of triggers and symptomatic therapies aimed at mediator blockade.[59] Drugs that safely and selectively suppress the pathogenic functions of mast cells could substantially improve quality of life in patients suffering from common conditions, such as asthma, as well as rare ones, such as mastocytosis. At the same time, the critical functions for mast cells in innate immunity identified in mouse models require investigation in humans. Whether mast cell effector functions can be selectively exploited to enhance immune responses and vaccine efficacy is an area of active investigation.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Voehringer D. Protective and pathological roles of mast cells and basophils. Nat Rev Immunol 2013;13(5):362–75.
2. Arinobu Y, Iwasaki H, Gurish MF, et al. Developmental checkpoints of the basophil/mast cell lineages in adult murine hematopoiesis. Proc Natl Acad Sci USA 2005;102:18105–10.
3. Siracusa MC, Kim BS, Spergel JM, et al. Basophils and allergic inflammation. J Allergy Clin Immunol 2013;132:789–801.
4. Dahlin JS, Malinovschi A, Öhrvik H, et al. Lin- CD34hi CD117int/hi FcεRI + cells in human blood constitute a rare population of mast cell progenitors. Blood 2016;127:383–91.
5. Dwyer DF, Barrett NA, Austen KF, et al. Expression profiling of constitutive mast cells reveals a unique identity within the immune system. Nat Immunol 2016;17:878–87.
6. Akin C. Clonality and molecular pathogenesis of mastocytosis. Acta Haematol 2005;114:61–9.
7. Han NR, Oh HA, Nam SY, et al. TSLP induces mast cell development and aggravates allergic reactions through the activation of MDM2 and STAT6. J Invest Dermatol 2014;134:2521–30.
8. Lantz CS, Boesiger J, Song CH, et al. Role for interleukin-3 in mast-cell and basophil development and in immunity to parasites. Nature 1998;392:90–3.
9. Godfraind C, Louahed J, Faulkner H, et al. Intraepithelial infiltration by mast cells with both connective tissue-type and mucosal-type characteristics in gut, trachea, and kidneys of IL-9 transgenic mice. J Immunol 1998;160:3989–96.
10. Yanagida M, Fukamachi H, Ohgami K, et al. Effects of T-helper 2-type cytokines, interleukin-3 (IL-3), IL-4, IL-5, and IL-6 on the survival of cultured human mast cells. Blood 1995;86:3705–14.
11. Hsieh FH, Lam BK, Penrose JF, et al. T helper cell type 2 cytokines coordinately regulate immunoglobulin E-dependent cysteinyl leukotriene production by human cord blood-derived mast cells: profound induction of leukotriene C(4) synthase expression by interleukin 4. J Exp Med 2001;193:123–33.
12. Gurish MF, Boyce JA. Mast cells: ontogeny, homing, and recruitment of a unique innate effector cell. J Allergy Clin Immunol 2006;117:1285–91.
13. Ochi H, Hirani WM, Yuan Q, et al. T helper cell type 2 cytokine-mediated comitogenic responses and CCR3 expression during differentiation of human mast cells in vitro. J Exp Med 1999;190:267–80.
14. Brightling CE, Ammit AJ, Kaur D, et al. The CXCL10/CXCR3 axis mediates human lung mast cell migration to asthmatic airway smooth muscle. Am J Respir Crit Care Med 2005;171:1103–8.
15. Kanakura Y, Sonoda S, Nakano T, et al. Formation of mast-cell colonies in methylcellulose by mouse skin cells and development of mucosal-like

mast cells from the cloned cells in the gastric mucosa of W/Wv mice. Am J Pathol 1987;129:168–76.
16. Wernersson S, Pejler G. Mast cell secretory granules: armed for battle. Nat Rev Immunol 2014;14:478–94.
17. Balzar S, Fajt ML, Comhair SA, et al. Mast cell phenotype, location, and activation in severe asthma. Data from the Severe Asthma Research Program. Am J Respir Crit Care Med 2011;183:299–309.
18. Brightling CE, Bradding P, Symon FA, et al. Mast-cell infiltration of airway smooth muscle in asthma. N Engl J Med 2002;346:1699–705.
19. Bradding P, Arthur G. Mast cells in asthma—state of the art. Clin Exp Allergy 2016;46:194–263.
20. Kitamura Y, Shimada M, Go S, et al. Distribution of mast-cell precursors in hematopoietic and lymphopoietic tissues of mice. J Exp Med 1979;150:482–90.
21. Irani AM, Craig SS, DeBlois G, et al. Deficiency of the tryptase-positive, chymase-negative mast cell type in gastrointestinal mucosa of patients with defective T lymphocyte function. J Immunol 1987;138:4381–6.
22. Denburg JA, Telizyn S, Messner H, et al. Heterogeneity of human peripheral blood eosinophil-type colonies: evidence for a common basophil-eosinophil progenitor. Blood 1985;66:312–18.
23. Saito H, Hatake K, Dvorak AM, et al. Selective differentiation and proliferation of hematopoietic cells induced by recombinant human interleukins. Proc Natl Acad Sci USA 1988;85:2288–92.
24. Abu-Ghazaleh RI, Dunnette SL, Loegering DA, et al. Eosinophil granule proteins in peripheral blood granulocytes. J Leukoc Biol 1992;52:611–18.
25. Schneider LA, Schlenner SM, Feyerabend TB, et al. Molecular mechanism of mast cell mediated innate defense against endothelin and snake venom sarafotoxin. J Exp Med 2007;204:2629–39.
26. Douaiher J, Succar J, Lancerotto L, et al. Development of mast cells and importance of their tryptase and chymase serine proteases in inflammation and wound healing. Adv Immunol 2014;122:211–52.
27. Kanaoka Y, Maekawa A, Austen KF. Identification of GPR99 protein as a potential third cysteinyl leukotriene receptor with a preference for leukotriene E4 ligand. J Biol Chem 2013;288:10967–72.
28. Chen CY, Lee JB, Liu B, et al. Induction of interleukin-9-producing mucosal mast cells promotes susceptibility to IgE-mediated experimental food allergy. Immunity 2015;43:788–802.
29. Galli SJ, Tsai M. IgE and mast cells in allergic disease. Nat Med 2012;18:693–704.
30. Rivera J, Fierro NA, Olivera A, et al. New insights on mast cell activation via the high affinity receptor for IgE. Adv Immunol 2008;98:85–120.
31. Galli SJ, Metcalfe DD, Arver DA, et al. Basophils and mast cells and their disorders. In: Kaushansky KLM, Beutler E, Kipps TJ, et al, editors. Williams hematology. 8th ed. New York: McGraw-Hill Medical; 2010. p. 915–32.
32. McNeil BD, Pundir P, Meeker S, et al. Identification of a mast-cell-specific receptor crucial for pseudo-allergic drug reactions. Nature 2015;519:237–41.
33. Kajiwara N, Sasaki T, Bradding P, et al. Activation of human mast cells through the platelet-activating factor receptor. J Allergy Clin Immunol 2010;125:1137–45.
34. Modena BD, Dazy K, White AA. Emerging concepts: mast cell involvement in allergic diseases. Transl Res 2016;174:98–121.
35. Siddiqui S, Mistry V, Doe C, et al. Airway hyperresponsiveness is dissociated from airway wall structural remodeling. J Allergy Clin Immunol 2008;122:335–41, 341.e1-3.
36. Laidlaw TM, Boyce JA. Aspirin-exacerbated respiratory disease—new prime suspects. N Engl J Med 2016;374:484–8.
37. Abraham SN, St John AL. Mast cell-orchestrated immunity to pathogens. Nat Rev Immunol 2010;10:440–52.
38. Sun J, Sukhova GK, Wolters PJ, et al. Mast cells promote atherosclerosis by releasing proinflammatory cytokines. Nat Med 2007;13:719–24.
39. Lee DM, Friend DS, Gurish MF, et al. Mast cells: a cellular link between autoantibodies and inflammatory arthritis. Science 2002;297:1689–92.
40. Secor VH, Secor WE, Gutekunst CA, et al. Mast cells are essential for early onset and severe disease in a murine model of multiple sclerosis. J Exp Med 2000;191:813–22.
41. Younan G, Suber F, Xing W, et al. The inflammatory response after an epidermal burn depends on the activities of mouse mast cell proteases 4 and 5. J Immunol 2010;185:7681–90.
42. Tsai M, Starkl P, Marichal T, et al. Testing the "toxin hypothesis of allergy": mast cells, IgE, and innate and acquired immune responses to venoms. Curr Opin Immunol 2015;36:80–7.
43. Groschwitz KR, Ahrens R, Osterfeld H, et al. Mast cells regulate homeostatic intestinal epithelial migration and barrier function by a chymase/Mcpt4-dependent mechanism. Proc Natl Acad Sci USA 2009;106:22381–6.
44. Garfield RE, Irani AM, Schwartz LB, et al. Structural and functional comparison of mast cells in the pregnant versus nonpregnant human uterus. Am J Obstet Gynecol 2006;194:261–7.
45. Lange M, Niedoszytko M, Renke J, et al. Clinical aspects of paediatric mastocytosis: a review of 101 cases. J Eur Acad Dermatol Venereol 2013;27:97–102.
46. Cohen SS, Skovbo S, Vestergaard H, et al. Epidemiology of systemic mastocytosis in Denmark. Br J Haematol 2014;166:521–8.
47. Bibi S, Langenfeld F, Jeanningros S, et al. Molecular defects in mastocytosis: KIT and beyond KIT. Immunol Allergy Clin North Am 2014;34(2):239–62.
48. Sokol H, Georgin-Lavialle S, Canioni D, et al. Gastrointestinal manifestations in mastocytosis: a study of 83 patients. J Allergy Clin Immunol 2013;132:866–73.e1-3.
49. Moura DS, Georgin-Lavialle S, Gaillard R, et al. Neuropsychological features of adult mastocytosis. Immunol Allergy Clin North Am 2014;34:407–22.
50. van der Veer E, van der Goot W, de Monchy JG, et al. High prevalence of fractures and osteoporosis in patients with indolent systemic mastocytosis. Allergy 2012;67:431–8.
51. Bonadonna P, Lombardo C, Zanotti R. Mastocytosis and allergic diseases. J Investig Allergol Clin Immunol 2014;24:288–97.
52. Pardanani A. Systemic mastocytosis in adults: 2015 update on diagnosis, risk stratification, and management. Am J Hematol 2015;90:250–62.
53. Horny HP, Metcalfe DD, Bennett JM, et al. Mastocytosis. In: Swerdlow SH, Campo E, Harris NL, editors. WHO classification of tumors of hematopoietic and lymphoid tissues. Lyon, France: IARC Press; 2008. p. 54–63.
54. Valent P, Akin C, Arock M, et al. Definitions, criteria and global classification of mast cell disorders with special reference to mast cell activation syndromes: a consensus proposal. Int Arch Allergy Immunol 2012;157:215–25.
55. Gotlib J, Kluin-Nelemans HC, George TI, et al. Efficacy and safety of midostaurin in advanced systemic mastocytosis. N Engl J Med 2016;374:2530–41.
56. Valent P, Sperr WR, Schwartz LB, et al. Diagnosis and classification of mast cell proliferative disorders: delineation from immunologic diseases and non-mast cell hematopoietic neoplasms. J Allergy Clin Immunol 2004;114:3.
57. Lim KH, Tefferi A, Lasho TL, et al. Systemic mastocytosis in 342 consecutive adults: survival studies and prognostic factors. Blood 2009;113:5727–36.
58. Matsuoka K, Shitara H, Taya C, et al. Novel basophil- or eosinophil-depleted mouse models for functional analyses of allergic inflammation. PLoS ONE 2013;8:e60958.
59. Soter NA. Mastocytosis and the skin. Hematol Oncol Clin North Am 2000;14(3):537–55, vi.

MULTIPLE-CHOICE QUESTIONS

1. Which of the following is TRUE of the development of mast cells and basophils?
   A. Mast cells and basophils both develop and mature fully in bone marrow.
   B. Mast cells originate in bone marrow but mature in peripheral tissues; basophils originate and mature fully in bone marrow.
   C. Mast cells originate and mature fully in bone marrow; basophils originate in bone marrow but mature in peripheral tissues.
   D. Mast cells and basophils both originate in bone marrow and mature in peripheral tissues.

2. Which of the following mast cell mediators is preformed and stored in granules?
   A. Interleukin (IL)-6
   B. LTC_4
   C. Tryptase
   D. IL-4

3. Mast cells and basophils can be activated by which of the following mechanisms?
   A. immunoglobulin (Ig)E cross-linking
   B. Platelet-activating factor
   C. Binding of complement components C3a or C5a
   D. All of the above

4. Which of the following is TRUE of the diagnosis of potential mastocytosis?
   A. A child with an isolated mastocytoma requires a bone marrow biopsy to rule out systemic mastocytosis.
   B. An adult with urticaria pigmentosa and an elevated serum tryptase level requires a bone marrow biopsy to rule out systemic mastocytosis.
   C. Elevated serum tryptase level during an anaphylactic reaction is a diagnostic criterion for systemic mastocytosis.
   D. Detection of a D816V mutation in *c-KIT* is not useful in making a diagnosis of systemic mastocytosis.

24

Eosinophils and Eosinophilia

Anna Kovalszki, Peter F. Weller

Eosinophils are terminally differentiated, bone marrow–derived granulocytes that normally circulate in blood in low numbers and tend to localize in those tissues with mucosal epithelial surfaces. Increased blood or tissue eosinophils occur in helminth parasite infections, as well as in allergic diseases and a variety of other, often idiopathic, conditions. Conventionally, the major focus on eosinophils has been delineating the "effector" functions of these end-stage granulocytes, including what roles these cells play as helminthotoxic effector cells and the contribution they make to the immunopathogenesis of allergic diseases. More recent findings have indicated that eosinophil functions are considerably more extensive.[1-4] Eosinophils contain stores of multiple, preformed cytokines; engage in cognate cell–cell interactions with other cell types, including lymphocytes; and have roles in varied host immune and inflammatory responses not conventionally marked by quantitatively extensive eosinophil infiltration.[4]

PRODUCTION AND DISTRIBUTION OF EOSINOPHILS

Eosinophilopoiesis

The development of eosinophils in bone marrow can be elicited by three cytokines. Granulocyte macrophage–colony-stimulating factor (GM-CSF), interleukin (IL)-3, and IL-5 all promote eosinophilopoiesis. In contrast to IL-3 and GM-CSF, which also promote the development of other lineages, IL-5 uniquely promotes the development and terminal differentiation of eosinophils. IL-5 is produced by type 2 innate lymphoid cells,[5] CD8 T cells, natural killer (NK) cells, and other leukocytes, including eosinophils themselves. IL-5 is a defining cytokine product of T-helper cell-2 (Th2) CD4 T cells. The production of IL-5 by Th2 lymphocytes accounts for the eosinophilia accompanying T-helper 2 (Th2) cell–mediated immune responses characteristic of helminth infections and allergic diseases.

Eosinophilopoiesis develops over about a week. Retained in bone marrow is a pool of mature eosinophils. IL-5, alone and in concert with the chemokine eotaxin-1 (CCL11), rapidly releases this pool of mature eosinophils into the circulation to acutely increase blood eosinophilia and facilitate recruitment of eosinophils to sites of inflammation.[1] Blood eosinophils circulate with a half-life of about 8–18 hours. Eosinophils leave the circulation and localize in tissues, especially those with mucosal interfaces with the outside world, such as the gastrointestinal (GI) and lower genitourinary tracts. Although the mechanisms governing this homing of eosinophils to mucosal tissues are not fully known, the chemokine eotaxin-1 is involved in the homing of eosinophils to the GI tract, but not to the respiratory tract.[2] Eosinophils live longer compared with neutrophils and probably persist in tissues for several weeks. They are principally tissue-dwelling cells: as demonstrated in rodents, for every eosinophil present in the circulation, there are 300–500 in tissues.

Eosinophil Adherence Mechanisms

The transit of eosinophils from bone marrow, through the circulation, and into tissues is governed, in part, by multiple adherence molecules expressed on eosinophils (Fig. 24.1). As for other leukocytes, recruitment of eosinophils into tissue sites of inflammation utilizes combinatorial interactions involving specific adhesion molecules (via their expression and altered affinity states) that mediate cellular interactions with the vascular endothelium and actions of chemoattractant molecules. Eosinophils express several adhesion molecules that they broadly share with other leukocytes. These adhesion molecules mediate their initial rolling and subsequent adherence to endothelial cells. Similar to neutrophils, eosinophils can adhere via CD11/CD18 heterodimeric β_2 integrins to the intercellular adhesion molecule-1 and -2 (ICAM-1 and ICAM-2). Likewise, specific sialoglycoproteins mediate adherence between eosinophils and endothelial E- and P-selectins. Unlike neutrophils, but similar to lymphocytes, eosinophils are able to bind to vascular cell adhesion molecule-1 (VCAM-1). Eosinophils express two α_4 integrins, very late activation antigen-4 (VLA-4, $\alpha_4\beta_1$) and $\alpha_4\beta_7$, which bind to VCAM-1. Moreover, $\alpha_4\beta_7$ binds to the mucosal addressin cell adhesion molecule (MAdCAM). The β_2 integrin $\alpha d\beta_2$, which binds ICAM-3 and is expressed on other leukocytes, is an additional integrin that mediates eosinophil adhesion to VCAM-1. Enhanced expression of VCAM-1 on the vascular endothelium, as elicitable by IL-4 or IL-13 stimulation, may contribute to the localization of eosinophils in some tissue sites of inflammation.

In addition to mediating interactions with the endothelium, eosinophil adherence molecules, by their interactions with extracellular matrix components, modulate the activity of eosinophils that have exited the bloodstream. Eosinophil VLA-6, $\alpha_6\beta_1$, binds laminin. Both $\alpha_4\beta_1$ and $\alpha_4\beta_7$ interact with specific domains of tissue fibronectin, and these interactions can enhance eosinophil functional responses. Eosinophils express CD44 (PGP-1), which binds hyaluronic acid. Siglec-8, a sialic acid–binding immunoglobulin-like lectin, is expressed on eosinophils and binds sialoglycoconjugates.[6]

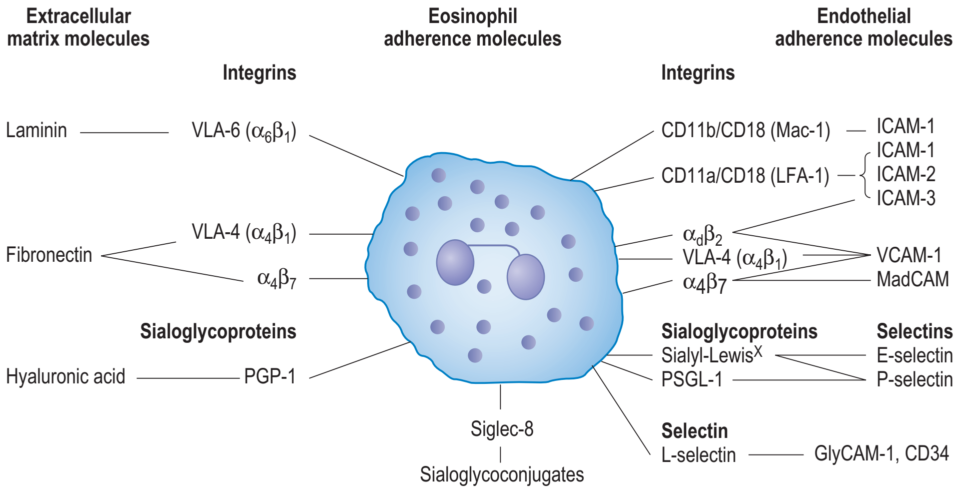

Extracellular matrix molecules
Eosinophil adherence molecules
Endothelial adherence molecules
Integrins
Integrins
Laminin
VLA-6 ($\alpha_6\beta_1$)
CD11b/CD18 (Mac-1)
ICAM-1
CD11a/CD18 (LFA-1)
ICAM-1
ICAM-2
ICAM-3
VLA-4 ($\alpha_4\beta_1$)
Fibronectin
$\alpha_4\beta_7$
$\alpha_d\beta_2$
VLA-4 ($\alpha_4\beta_1$)
$\alpha_4\beta_7$
VCAM-1
MadCAM
Sialoglycoproteins
Sialoglycoproteins
Selectins
Sialyl-Lewis^X
E-selectin
Hyaluronic acid
PGP-1
PSGL-1
P-selectin
Siglec-8
Sialoglycoconjugates
Selectin
L-selectin
GlyCAM-1, CD34

FIG 24.1 Adherence Mechanisms Utilized by Human Eosinophils to Bind to Vascular Endothelial Cells and the Extracellular Matrix Molecules. ICAM, intercellular adhesion molecule; VCAM, vascular cell adhesion molecule; MAdCAM, mucosal addressin cell adhesion molecule; VLA, very late activation antigen.

KEY CONCEPTS

Actions of Eosinophilopoietic Cytokines Interleukin (IL)-3, Granulocyte Macrophage–Colony-Stimulating Factor (GM-CSF), IL-5

Promote eosinophil development and maturation in bone marrow (IL-5).
Release a pool of mature eosinophils from bone marrow (IL-5).
Sustain the viability and antagonize apoptosis of mature eosinophils, enhance multiple effector responses of mature eosinophils.

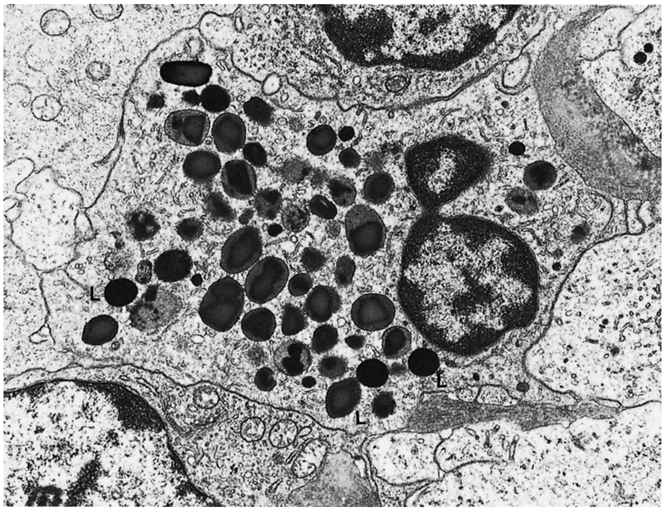


FIG 24.2 Transmission Electron Micrograph of a Human Eosinophil. The numerous cytoplasmic specific granules contain the electron-dense crystalline cores that are unique to eosinophils. In addition, lipid bodies are visible as globular, uniformly dark structures. (Original magnification × 11,180.) (Courtesy of Dr. Ann M. Dvorak, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston.)

Eosinophil Chemoattractants

Mobilization of eosinophils into tissues is governed by receptor-mediated chemoattractant stimuli. Chemoattractants promote the directed migration of eosinophils and may enhance the adhesion of eosinophils to vascular endothelium and their subsequent migration through the endothelium. Many compounds have been identified as eosinophil chemoattractants, including humoral immune mediators, such as platelet-activating factor (PAF) and the complement anaphylatoxins C5a and C3a; certain cytokines; and several chemokines, most notably the eotaxins. None of these is specific solely for eosinophils, but eotaxin-1, eotaxin-2, and eotaxin-3 exhibit the most restricted specificity.[1] Eotaxins signal through CCR3 chemokine receptors that are expressed on eosinophils as well as basophils, some Th2 cells, and some mast cells. Thus recruitment of eosinophils to sites of immunological reactions is governed by their response to chemoattractants that facilitate intravascular emigration and direct migration of extravascular eosinophils, as well as by the functional states of eosinophil adherence molecules and the differential expression of endothelial cell adherence ligands.

STRUCTURE OF EOSINOPHILS

Human eosinophils, unlike neutrophils, usually have a bilobed nucleus (Fig. 24.2). Defining attributes of eosinophils are their large, cytoplasmic "specific" granules that are morphologically distinct because of their unique content of crystalloid cores. Crystalloid cores are recognizable by transmission electron microscopy and usually appear electron dense (see Fig. 24.2). The cores and surrounding matrices of specific granules contain cationic proteins that account for the tinctorial staining of granules with eosin. Eosinophils at sites of inflammation can exhibit morphological changes in their specific granules, including loss of either matrix or core components from within intact granules, compatible with the extracellular release of granule constituents.

Lipid bodies, cytoplasmic structures distinct from granules (see Fig. 24.2), are roughly globular in shape and range in size from minute to the size of specific granules. Lipid bodies are dissolved by common alcohol-based hematological stains but are preserved and stain darkly with osmium fixation. Lipid bodies lack a delimiting membrane but contain internal membranes that are often obscured by overlying lipid. Lipid bodies are found in neutrophils and other cells, especially in association with inflammation; but eosinophils typically contain more lipid bodies compared with neutrophils. Lipid body formation in eosinophils is rapidly inducible within minutes. In eosinophils, key enzymes involved in eicosanoid formation, including prostaglandin H synthase, the 5- and 15-lipoxygenases, and leukotriene (LT) C_4 synthase, localize at lipid bodies; and lipid bodies are sites of eicosanoid synthesis.[7]

CELL-SURFACE RECEPTORS AND PROTEINS

Eosinophil receptors for immunoglobulins include those for immunoglobulin G (IgG), IgE, and IgA. The receptor for IgG on eosinophils is principally the low-affinity FcγγRII (CD32), whereas neutrophils have FcγRII and FcγRIII (CD16).

Eosinophil IgE receptors include the high-affinity IgE receptor FcεRI, typically found on basophils and mast cells, as well as FcεRII, the low-affinity IgE receptor, such as CD23, found on lymphocytes, monocytes, and antigen-presenting cells (APCs). Although FcεRI α-chain protein is present within eosinophils, its surface expression can be low or undetectable. Engagement of eosinophil FcεRI does not elicit exocytotic degranulation, as it does on basophils and mast cells. FcεRI may participate in IgE-mediated antigen uptake by antigen-presenting eosinophils. Eosinophil expression of IgE receptors is notable because IgE levels and eosinophil numbers frequently increase concomitantly in helminth infections as well as allergic diseases.

Eosinophils express FcαRI (CD89), which binds secretory IgA more potently than other forms of IgA. Engagement of FcαRI triggers eosinophil release of granule proteins. With the characteristic localization of eosinophils to mucosal surfaces of the respiratory, GI, and genitourinary tracts, this IgA receptor enables eosinophils to engage secretory IgA present at these mucosal sites.

Eosinophils have receptors for complement components, including C1q (CR1), C3b/C4b (CR1), iC3b (CR3), C3a, and C5a. Both C3a and C5a are eosinophil chemoattractants and stimulate the production of oxygen radicals by eosinophils. Eosinophils express several receptors for chemokines. CCR1 is a receptor for MIP-1α, MCP-3, and RANTES, whereas CCR3 is a receptor for eotaxin-1, eotaxin-2, eotaxin-3, MCP-3, and RANTES. Eosinophils express CXCR4 and respond to the ligand for this receptor, stromal cell–derived factor 1α.

Mature eosinophils, like their immature precursors, express receptors for the three cytokines, GM-CSF, IL-3, and IL-5, which promote eosinophilopoiesis and stimulate the functioning of mature eosinophils. In addition, eosinophils have receptors for a broad range of other cytokines, including IL-1α, IL-2, IL-4, interferon (IFN)-α and IFN-γ, tumor necrosis factor (TNF)-α, stem cell factor (c-KIT), and IL-16 (which signals via CD4 on eosinophils). Thus eosinophils are subject to stimulation by many cytokines, although little is understood about how many of them affect eosinophil functioning *in vivo*.

Of pertinence to interactions between eosinophils and B and T lymphocytes, eosinophils can express several relevant plasma membrane proteins. Class II major histocompatibility complex (MHC) proteins, generally absent on blood eosinophils, are induced for expression on eosinophils in sites of inflammation. In addition, eosinophils can express CD40, CD154 (CD40 ligand), CD153 (CD30 ligand), CD28 (B7–2), and CD86.[4]

Eosinophils express receptors for several lipid mediators, including PAF and leukotriene B_4 (LTB_4), which are chemoattractants for eosinophils and stimulate eosinophil degranulation and respiratory burst activity. Eosinophils also have receptors for prostaglandins D_2 and E_2 and for cysteinyl leukotrienes.

CONSTITUENTS OF EOSINOPHILS

Eosinophil-specific granules contain preformed proteins that include both specific cationic proteins and stores of diverse cytokines and chemokines.

Cationic Granule Proteins

Eosinophil granule cationic proteins have been extensively studied because of their abundance in the granules and their capacity to exert multiple effects on host cells and microbial targets. Major basic protein (MBP), named for its quantitative predominance within the granule and its markedly cationic (basic) isoelectric point of about 11, is a 13.8- to 14-kilodalton (kDa) protein. A homolog of MBP that is somewhat smaller (13.4 kDa) and less basic (pI 8.7) has been identified. MBP lacks enzymatic activity and probably exerts its varied effects via its markedly cationic nature. MBP was found to be toxic to both airway epithelium and helminths and to have antibacterial effects.[8]

A second granule protein is eosinophil peroxidase (EPO), an enzyme distinct from neutrophil myeloperoxidase. Cationic EPO (pI 10.8) uses hydrogen peroxide and halide ions to form hypohalous acids, which are toxic for parasites, bacteria, and tumor and host cells. EPO utilizes bromide in preference to chloride and is even more active with a pseudohalide, thiocyanate, to generate oxidant products, including hypobromous and hypothiocyanous acids.

Two additional granule proteins are eosinophil cationic protein (ECP) (18 kDa, pI 10.8) and eosinophil-derived neurotoxin (EDN) (18–19 kDa, pI 8.9). EDN, never demonstrated to be neurotoxic for humans, is so named only because, after it is injected intracerebrally into test rabbits, it elicits a characteristic neuropathological response. Both ECP and EDN are ribonucleases (RNases). EDN expresses 100 times more RNase activity compared with ECP, although their toxic effects on bacterial, parasitic, and mammalian target cells are not simply a result of their RNase catalytic activities.

Within the specific granule, MBP is localized to the crystalloid core and to tubulovesicular structures within and arising from the granules,[8] whereas ECP, EDN, and EPO are localized in the matrix of the granule around the core (see Fig. 24.2). MBP is also found in low amounts (~3% of eosinophil levels) in basophils, but whether this reflects endocytosis or endogenous synthesis is not known. Uptake of MBP and EPO into mast cells is known to occur via endocytosis. Small amounts of EDN and ECP are present in neutrophils, and as neutrophils contain messenger RNA (mRNA) transcripts for these, EDN and ECP are likely synthesized by neutrophils. Nevertheless, eosinophils are the dominant source of these four cationic proteins. The properties of these proteins and their numerous biological effects have been reviewed,[9] as these proteins have major effects not only in the potential role of eosinophils in host defense against helminth parasites but also in contributing to tissue dysfunction and damage in eosinophil-related allergic and other diseases.

Cytokines and Chemokines

Eosinophils are capable of elaborating at least four dozen diverse cytokines and chemokines, and studies continue to identify more cytokines released by eosinophils. The potential activities of eosinophil-derived cytokines are extensive. Eosinophil-derived cytokines include those with autocrine growth factor activities for eosinophils and those with potential roles in acute and chronic inflammatory responses. A notable feature of eosinophils as a source of cytokines is that they contain stores of these cytokines preformed within eosinophil granules and secretory vesicles.[10] Thus in contrast to most lymphocytes, which must be induced to synthesize *de novo* cytokines destined for release, eosinophils can immediately release preformed cytokine and chemokine proteins into the surrounding milieu. The local, rapid release of eosinophil-derived cytokines in tissues to effect adjacent cells could, and has been shown to, readily induce responses in varied cell types.

Eosinophils synthesize the three growth factor cytokines GM-CSF, IL-3 and IL-5, which promote eosinophil survival, antagonize eosinophil apoptosis, and enhance eosinophil effector responses. Other cytokines elaborated by human eosinophils that may have activities in acute and chronic inflammatory responses include IL-1α, IL-6, IL-8, IFN-γ, TNF-α, and MIP-1α. Human eosinophils can elaborate other various "growth" factors, including transforming growth factor (TGF)- α, TGF-β_1, vascular endothelial growth factor, platelet-derived growth factor (PDGF)-β, heparin-binding epidermal growth factor, and a proliferation-inducing ligand (APRIL). These cytokines have roles in contributing to epithelial hyperplasia and fibrosis, as well as other immune homeostatic activities. In addition, eosinophils are recognized as sources of specific cytokines and chemokines capable of stimulating or inhibiting lymphocyte responses, including IL-2, IL-4, IL-10, IL-12, IL-16, RANTES, and TGF-β_1. Of note, eosinophil cytokines include those associated with Th2 (IL-4, IL-5, IL-13), Th1 (IL-12, IFN-γ), and T-regulatory (IL-10, TGF-β) responses, emphasizing the diverse immunoregulatory potentials for eosinophil-secreted cytokines.[10]

ACTIVATED EOSINOPHILS

A well-recognized attribute of eosinophils is that, in conjunction with eosinophilic diseases, some blood and tissue eosinophils may exhibit various alterations, indicating that these cells have been "activated." These changes include increased metabolic activity, diminished density ("hypodense"), enhanced LTC_4 formation, and morphological alterations, including cytoplasmic vacuolization, alterations in granule numbers and size, and losses within specific granules of MBP-containing cores or matrices. Activated eosinophils may exhibit enhanced plasma membrane expression of some proteins, including CD69, human leukocyte antigen–D related (HLA-DR), and CD25.

Features associated with *in vivo* "activated" eosinophils can be elicited, in part, by exposing eosinophils to specific stimuli, including GM-CSF, IL-3, and IL-5. In addition, interactions with extracellular matrix components can further contribute to eosinophil activation. Eosinophil "activation," however, is not a singularly binary process, and some attributes of activation can be elicited without other attributes by mediators and mechanisms that remain to be delineated.

MECHANISMS OF EOSINOPHIL DEGRANULATION

As eosinophil granules contain four major cationic proteins and a multitude of preformed cytokines and chemokines, the processes by which eosinophils mobilize these granule constituents for their extracellular release are important in understanding the regulated functioning of eosinophils. Unlike mast cells or basophils that undergo acute exocytotic degranulation in response to cross-linking of their high-affinity Fcε receptors, an analogous mechanism to elicit comparable exocytotic degranulation of fluid-phase eosinophils has not been identified. Cross-linking of eosinophil IgG or IgA Fc receptors *in vitro* can stimulate release of eosinophil cationic proteins, but this rapid FcR-mediated acute "degranulation" process is cytolytic for eosinophils. In contrast, observations of eosinophils on the surfaces of large nonphagocytosable multicellular helminth parasites do provide evidence that eosinophils can degranulate by exocytosis to wholesale release granule contents on the surfaces of target parasites.

An alternative mechanism of mobilizing granule contents for secretion that eosinophils utilize is a process of vesicular transport-mediated "piecemeal" degranulation. Electron microscopic observations of lesional eosinophils provided evidence that eosinophil granule contents were mobilized *in vivo* by selective incorporation into small vesicles that traffic to the cell surface and release these granule contents. By this process, there may be agonist-elicited selective secretion of certain eosinophil-derived cytokines.[11] Ultrastructural studies have demonstrated that secretory vesicles arise from granules and transport cytokines, such as IL-4.[11] Insights into the selectivity and mechanisms of differential cytokine secretion have been gained by the finding, at least for IL-4, that a receptor for IL-4 mediates the transport of IL-4 from granules and within secretory vesicles.[12] How this process of vesicular transport is regulated and functions to selectively mobilize specific eosinophil granule–derived cytokines or cationic proteins remains under investigation.

In addition to regulated release of granule contents from viable eosinophils, a common, but often overlooked, occurrence is the lysis of eosinophils. Both cutaneous and pulmonary biopsy specimens of eosinophil-associated diseases contain free, extracellular, but still membrane-bound, eosinophil granules. These free extracellular granules express cytokine, chemokine, and cysteinyl leukotriene receptors and are secretion competent even outside of intact eosinophils.[13,14] Cytolytic stimuli elicit both the release of nuclear DNA to form extracellular DNA "traps" and the release of free secretion-competent eosinophil granules in humans.[15]

FUNCTIONS OF EOSINOPHILS

Conventional considerations of the roles that eosinophils may play have been guided by quantitative considerations so that those diseases characteristically marked by more prominent eosinophilia have occasioned the most interest. Thus studies have focused on the "effector" roles eosinophils play in host defense against helminth infections and in the immunopathogenesis of allergic and other eosinophilic diseases. Additional roles for eosinophils must be considered in immune or inflammatory responses not conventionally recognized to contain abundant eosinophils.[4]

ROLES IN HOST DEFENSE

Because the host response to infections with multicellular helminth parasites is characteristically associated with eosinophilia, it is often believed that eosinophils evolved to have a role in killing helminths, especially during their larval stages. Indeed, *in vitro* eosinophils can kill numerous helminths, organisms too large to be phagocytosed. Eosinophils adhere to the parasite and deposit eosinophil granule contents onto its surface. Cell products that can contribute to parasite death include MBP, ECP, EDN, and EPO.

As reviewed earlier,[16] the helminthotoxic roles of eosinophils *in vivo* are less certain in humans and rodents. In eosinophil-depleted mice, the intensities of primary and secondary infections with some helminths have not been greater than in eosinophilic mice, nor have IL-5 transgenic mice exhibited increased resistance to infection with some helminth species. Moreover, schistosome infections in two lines of eosinophil-ablated mice have shown no differences in measures of infection compared with normal mice.[17] Nevertheless, murine studies need to be interpreted with caution. Many experimental infections involve introducing helminth infections that are often host species–restricted into unnatural host mice, in which innate immune responses may be prominent. Natural human infections are usually a consequence of repeated exposures, during which acquired, rather than innate, immunity becomes prominent. Thus eosinophil functions as helminthotoxic cells *in vivo* remain unclear. Eosinophils might have alternative functions in host responses to helminths, including functioning as APCs and even favoring the survival of *Trichinella* larvae in muscles.[18,19]

ROLES IN DISEASE PATHOGENESIS

The abilities of eosinophils to release biologically active lipids as paracrine mediators of inflammation and to release preformed cationic and cytokine granule constituents enable eosinophils to contribute to the immunopathogenesis of various diseases, including asthma.[1] Eosinophils form several classes of biologically active lipids. Eosinophils may liberate PAF, whose diverse activities can be mediated either directly or by stimulating other cells to release leukotrienes, prostaglandins, and complement peptides. Stimulated eosinophils release LTC_4. LTD_4 and LTE_4 are formed from LTC_4 by the sequential enzymatic removal of glutamic acid and glycine from its tripeptide glutathione side chain. LTC_4 and especially LTD_4 have bronchoconstrictor activities, constrict terminal arterioles, dilate venules, and stimulate airway mucus secretion. Thus eosinophils are a potential source of two major types of mediator lipids, the sulfidopeptide leukotrienes and PAF.

Oxidants released by eosinophils, including superoxide anion, hydroxyl radical, and singlet oxygen, as well as EPO-catalyzed hypothiocyanous acid and other hypohalous acids, have the potential to damage host tissues.

Released eosinophil granule proteins are immunologically detectable in fluids, including blood, sputum, and synovial fluids, and in tissues, including the respiratory and GI tracts, skin, and heart, in association with various eosinophil-related diseases. The eosinophil cationic proteins, including MBP, ECP, and EPO, can damage various cell types. Thus extracellular release of eosinophil granule proteins, by degranulation or cytolysis of eosinophils, could contribute to local tissue damage by causing dysfunction and damage to adjacent cells.

OTHER EOSINOPHIL FUNCTIONS

Other potential functions for the eosinophil are not fully defined. In addition to the acute release of lipid, peptide, and cytokine mediators of inflammation, eosinophils probably contribute to chronic inflammation, including the development of fibrosis. Eosinophils can be a major source of the fibrosis-promoting cytokine TGF-β. Additional roles of eosinophils in modulating extracellular matrix deposition and remodeling are suggested by studies of normal wound healing. During dermal wound healing, eosinophils infiltrate into wound sites and sequentially express TGF-α early and TGF-β_1 later during wound healing. These findings suggest that eosinophils may contribute to the more chronic subepithelial airway fibrosis characteristic of chronic asthma.

Additional functions for eosinophils are indicated by the findings that they may be induced to express class II MHC proteins and can function as APCs.[20] Blood eosinophils lack HLA-DR expression, but eosinophils recovered from the airways 48 hours after segmental antigen challenge have been shown to express HLA-DR. Cytokines, including GM-CSF, IL-3, IL-4, and IFN-γ, induce eosinophil HLA-DR expression. Both murine and human eosinophils can function as HLA-DR–dependent MHC-restricted APCs in stimulating the proliferation of T cells. *In vivo,* murine eosinophils can process exogenous antigens in the airways, traffic to regional lymph nodes, and function as antigen-specific APCs to stimulate responses of CD4 T cells.[21]

Eosinophils, that normally become resident in submucosal and less prominently in other tissues, undoubtedly participate in ongoing homeostatic immune responses at these sites. Some of these responses are mediated by cytokines secreted by eosinophils, including IL-6 and APRIL to stimulate plasma cells development[22] and IL-4 to activate macrophages in fat tissue and effect glucose metabolism.[23] Further investigations will help delineate eosinophil's functional roles and interactions with other cells, so that the scope of eosinophil functions will probably extend beyond its currently more defined role as an effector cell contributing to allergic inflammation.

EOSINOPHILIA AND EOSINOPHILIC DISORDERS

Diverse infectious, allergic, neoplastic and idiopathic disease processes can be associated with increased blood and/or tissue eosinophil numbers. Blood eosinophilia, present when eosinophil numbers are in excess of their usual level of <450/μL of blood, may be intermittently, modestly, or (less frequently) markedly increased. Blood eosinophil numbers are not necessarily indicative of the extent of eosinophil involvement in affected tissues.

Some patients with sustained blood eosinophilia can develop organ damage, especially cardiac damage. This cardiac involvement can include the formation of intraventricular thrombi and endomyocardial fibrosis with secondary mitral or tricuspid regurgitation (Fig. 24.3). Such damage can complicate the sustained eosinophilia of hypereosinophilic syndromes and has been noted with eosinophilias accompanying other diseases, including eosinophilia with carcinomas, lymphomas, GM-CSF, or IL-2 administration, drug reactions, and parasitic infections. Most patients with eosinophilia, however, develop no evidence of endomyocardial damage. Conversely, cardiac disease can rarely present in patients without known eosinophilia. The pathogenesis of eosinophil-mediated cardiac damage involves both usually

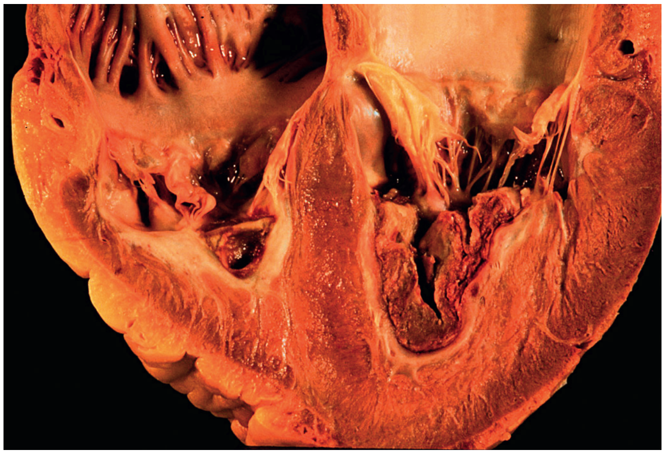


FIG 24.3 **Eosinophil Endomyocardial Disease.** A large thrombus is present in the apex of the left ventricle and the chordae tendineae are entrapped, leading to severe mitral valve regurgitation.

heightened numbers of eosinophils and some activating events, as yet ill-defined, that promote eosinophil-mediated tissue damage.

Cardiac damage progresses through three stages. In the first stage, there is damage to the endocardium and infiltration of the myocardium with eosinophils and lymphocytes, with eosinophil degranulation and myocardial necrosis. Elevated plasma levels of troponin can be a sensitive assay of early eosinophil-mediated cardiac damage. A similar acute eosinophilic myocarditis can develop with drug hypersensitivity reactions and may be more fulminant. The first stage is frequently clinically occult, although subungual splinter hemorrhages may be prominent. Elevations of serum troponins as a measure of myocardial injury should be evaluated. Echocardiography usually detects no abnormalities at this stage, although cardiac magnetic resonance imaging (MRI) is evolving as a technique to potentially detect cardiac involvement at an earlier stage. Uncommonly, death due to acute progressive cardiac disease can occur. Corticosteroid therapy during the acute stage may help control and prevent the evolution of myocardial fibrosis.

The second stage of heart disease, the formation of thrombi along the damaged endocardium, affects either or both ventricles and occasionally the atrium. Outflow tracts near the aortic and pulmonic valves are usually spared. These thrombi can embolize to the brain and elsewhere. Finally, in the fibrotic stage, progressive scarring leads to entrapment of chordae tendineae with the development of mitral and/or tricuspid valve regurgitation and to endomyocardial fibrosis, producing a restrictive cardiomyopathy. Echocardiography and MRI are valuable in detecting intracardiac thrombi and the manifestations of fibrosis. Patients with sustained eosinophilia should be monitored by using echocardiography and serum troponin assays for evidence of cardiac disease. In an older series of patients referred to the National Institutes of Health (NIH), much of the mortality among these patients with hypereosinophilia was attributable to end-stage congestive heart failure. In contemporary times, earlier recognition of cardiac involvement, mitral valve replacement with bioprostheses and additional therapeutic options for hypereosinophilic syndromes (see below) have largely minimized the morbidity and mortality attributable to end-stage eosinophilic endomyocardial disease.

INFECTIOUS DISEASES ASSOCIATED WITH EOSINOPHILIA

Eosinophilia is encountered only with specific infectious diseases. With active bacterial or viral infections, eosinopenia is characteristic. This suppression of blood eosinophils is, in part, caused by heightened endogenous corticosteroid production as well as by inflammatory mediators released during these infections. This suppression of eosinophilia, with either serious bacterial infections or marked inflammation, accounts for the absence of otherwise expected eosinophilia in some patients with helminth infections, including those with hyperinfection strongyloidiasis.[24] As a general clinical guideline, patients with a febrile illness and an increased or even normal blood eosinophilia are not likely to have common bacterial or viral infections, unless they have adrenal insufficiency or a confounding medication-elicited eosinophilia.

Helminth Parasites

Helminth parasites are multicellular metazoan organisms—the "worm" parasites. Infections with diverse helminths elicit eosinophilia (Chapter 31).[24] Although eosinophilia may provide a hematological clue to the presence of helminth infection, the absence of blood eosinophilia does not exclude such infections. The eosinophilic response to helminths is determined both by the host's immune response and by the parasite, including its distribution, migration, and development within the infected host. The level of eosinophilia tends to parallel the magnitude and extent of tissue invasion by helminth larvae or adults. In several helminth infections, the migration of infecting larvae or subsequent developmental stages through tissues is greatest early in infection, and hence the magnitude of the elicited eosinophilia will be the greatest in these early phases. In established infections, local eosinophil infiltration will often be present around helminths within tissues, without significant blood eosinophilia. Eosinophilia may be absent in those helminth infections that are well contained within tissues (*e.g.,* intact echinococcal cysts) or are solely intraluminal within the intestinal tract (*e.g., Ascaris,* tapeworms). In some established infections, increases in blood eosinophilia may be episodic. Intermittent leakage of cyst fluids from echinococcal cysts can transiently stimulate increases in blood eosinophilia and also cause symptoms attributable to allergic or anaphylactic reactions (urticaria, bronchospasm). For tissue-dwelling helminths, increases in eosinophilia may occur principally in association with migration of adult parasites, as in loiasis and gnathostomiasis.

Helminth infections more likely to elicit prolonged hypereosinophilia in adults include filarial and hookworm infections and strongyloidiasis (Table 24.1).[24] Trichinellosis can elicit an acute hypereosinophilia. *Strongyloides stercoralis* infection, difficult to diagnose solely by stool examinations, is especially important to exclude, not only because it elicits modest to even marked eosinophilia but also because, unlike other helminths, it can develop into a disseminated, often fatal, disease (hyperinfection syndrome) in patients given immunosuppressive corticosteroids.[24] Enzyme-linked immunosorbent assay (ELISA) serology has proved valuable in detecting strongyloidiasis and should be obtained for patients with eosinophilia likely to receive corticosteroids. Some tissue- or blood-dwelling helminths that are not diagnosable by stool examinations but that can cause marked eosinophilia require diagnostic examination of blood or biopsied tissues or specific serological tests.[24] Infections with these organisms include filarial infections, trichinellosis, and visceral larva migrans. In

TABLE 24.1 Parasitical Diseases Capable of Causing Marked (>3000/µL) or Long-Standing Eosinophilia

Helminth Infection	Hypereosinophilia	Chronic Eosinophilia
Angiostrongyliasis costaricensis		
Ascariasis	+ During early lung migration	
Hookworm infection	+ During early lung migration	+ Common cause of low-grade eosinophilia
Strongyloidiasis	Uncommonly	+ Self-perpetuating, may last > 50 years
Trichinellosis	+ With heavy infections	
Visceral larva migrans	+ Principally in children	
Gnathostomiasis		
Cysticercosis		
Echinococcosis		+ May be episodic with cyst fluid leakage
Filariases:		
Tropical pulmonary eosinophilia	+	+
Loiasis	+ Especially in expatriates	+
Onchocerciasis	+	+
Flukes:		
Schistosomiasis	+ During early infection in nonimmune patients	+
Fascioliasis	+ During early infection	+
Clonorchiasis	+ During early infection	+
Paragonimiasis	+ During early infection	+
Fasciolopsiasis	+ During early infection	+

Adapted from Wilson ME, Weller PF. Eosinophilia. In: Guerrant RL, Walker DH, Weller PF, eds. Tropical Infectious Diseases: Principles, Pathogens and Practice, 3rd ed. Philadelphia, Penn.: Churchill Livingstone; 2011:943.

children, owing to their propensity for geophagous pica and ingestion of dirt contaminated by dog ascarid eggs, visceral larva migrans caused by *Toxocara canis* is a potential etiology for sustained eosinophilia. ELISA serological testing can evaluate this possibility.

Other Infections: Protozoa and Fungi

Infections with single-celled protozoan parasites do not characteristically elicit eosinophilia. This is true of all intestinal, blood-, and tissue-infecting protozoa, with three exceptions. Two intestinal protozoans, *Dientamoeba fragilis* and *Isospora belli,* can at times be associated with low-grade eosinophilia. Hence, in patients with symptoms of enteric infection and eosinophilia, diagnostic trophozoites of *D. fragilis* or oocysts of *I. belli* should be sought in stool examinations. Fecal examinations for *I. belli* oocysts must be specifically requested, as they are not usually detected in routine stool ova and parasite examinations. Other enteric protozoa do not elicit eosinophilia and, if detected in stool examinations, should not be accepted as causes of eosinophilia. *Sarcocystis,* a myositis producing protozoan, can elicit modest eosinophilia.

Three fungal diseases can be associated with eosinophilia. (i) Aspergillosis is accompanied by eosinophilia only in the form of allergic bronchopulmonary aspergillosis, not when it is an invasive pathogen. (ii) Coccidioidomycosis, following primary infection, especially in conjunction with erythema nodosum, and at times with progressive disseminated disease, can elicit blood eosinophilia and may cause an eosinophilic meningitis. (iii) Basidiobolomycosis infection can also be associated with eosinophilia.[24]

Human Immunodeficiency Virus and Retroviral Infections

Eosinophilia can uncommonly accompany human immunodeficiency virus (HIV) infections for several reasons. First, leukopenia may increase eosinophil percentages without reflecting true eosinophilia. Second, adverse reactions to medications may elicit eosinophilia. Third, patients with acquired immunodeficiency syndrome (AIDS) who develop adrenal insufficiency as a result of cytomegalovirus and other infections may exhibit eosinophilia as a consequence. In addition, modest, and uncommonly marked, eosinophilia and eosinophilic pustular folliculitis can be observed in some patients with HIV infection.[24] Eosinophilia more commonly develops with HTLV-1 infections.[24]

ALLERGIC DISEASES ASSOCIATED WITH EOSINOPHILIA

Among the noninfectious diseases associated with eosinophilia (Table 24.2) are allergic diseases, notably those mediated by IgE-dependent mechanisms. In these diseases, including allergic rhinitis, conjunctivitis, and asthma, eosinophils are present in involved tissues as well as often being increased in blood.

MYELOPROLIFERATIVE AND NEOPLASTIC DISEASE

Eosinophilia can occur with specific neoplastic diseases, as well as in some disorders of uncertain etiology, including some hypereosinophilic syndromes.

Hypereosinophilic Syndromes

A syndrome previously termed *idiopathic hypereosinophilic syndrome* is not a single entity but rather a constellation of leukoproliferative disorders characterized by sustained overproduction of eosinophils. The three original diagnostic criteria for this syndrome were eosinophilia in excess of 1500/µL of blood persisting for longer than 6 months; lack of an identifiable parasitic, allergic, or other etiology for eosinophilia; and signs and symptoms of organ involvement. In contemporary practice, if eosinophilia is sustained over a month and the other criteria are reliably met before the 6-month time frame, a diagnosis can be made and treatment promptly initiated. Moreover, in recent years there has been increasing recognition that hypereosinophilic syndromes (HES), even at times without evident organ damage, encompass a spectrum of disorders, and progress has been made in identifying the underlying defects in some of these (Fig. 24.4).[25]

KEY CONCEPTS

Hypereosinophilic Syndromes

Eosinophilia sustained in excess of 1500/µL.
Absence of allergic, parasitic, or other etiologies for eosinophilia.
Usually evidence of organ involvement.

TABLE 24.2 Eosinophil-Associated Diseases and Disorders

Allergic or Atopic Diseases
Asthma
Allergic rhinitis
Eosinophilic esophagitis
Atopic dermatitis
Allergic urticaria
Nasal polyps

Myeloproliferative and Neoplastic Disorders
Hypereosinophilic syndromes: myeloproliferative, lymphoproliferative, and others
Leukemia
Lymphoma- and tumor-associated
Mastocytosis

Pulmonary Syndromes
Parasite-induced eosinophilic lung diseases:
Loeffler syndrome: patchy migratory infiltrates, resolving over weeks, seen with transpulmonary migration of helminth parasites, especially *Ascaris*
Tropical pulmonary eosinophilia: miliary lesions and fibrosis; heightened immune response causing one form of lymphatic filariasis; increased immunoglobulin E (IgE) and antifilarial antibodies
Pulmonary parenchymal invasion: paragonimiasis
Heavy hematogenous seeding with helminths: trichinellosis, schistosomiasis, larva migrans
Allergic bronchopulmonary aspergillosis
Chronic eosinophilic pneumonia: dense peripheral infiltrates, fever; progressive, blood eosinophilia may be absent; steroid responsive
Acute eosinophilic pneumonia: acute presentation diagnosed by bronchoalveolar lavage or biopsy
Eosinophilic granulomatosis with polyangiitis (Churg–Strauss syndrome) vasculitis: small- and medium-sized arteries; granulomas, necrosis; asthma often antecedent; extrapulmonary (*e.g.,* neurological, cutaneous, cardiac, or gastrointestinal) involvement likely
Drug- and toxin-induced eosinophilic lung diseases
Other: hypereosinophilic syndromes, neoplasia, bronchocentric granulomatosis

Skin and Subcutaneous Diseases
Skin diseases: atopic dermatitis, blistering diseases, including bullous pemphigoid, urticarias, drug reactions
Diseases of pregnancy: pruritic urticarial papules and plaques syndrome, herpes gestationis
Eosinophilic pustular folliculitis
Eosinophilic cellulitis (Wells syndrome)
Kimura disease and angiolymphoid hyperplasia with eosinophilia
Shulman syndrome (eosinophilic fasciitis)
Episodic angioedema with eosinophilia: recurrent periodic episodes with fever, angioedema, and secondary weight gain; may be long-standing without untoward cardiac dysfunction

Gastrointestinal Disorders
Eosinophilic gastroenteritides
Inflammatory bowel disease: eosinophils in lesions; occasionally blood eosinophilia with ulcerative colitis

Rheumatologic Diseases
Vasculitis: Eosinophilic granulomatosis with polyangiitis (Churg–Strauss) and cutaneous necrotizing eosinophilic vasculitis

Immunological Reactions
Medication-related eosinophilias
Immunodeficiency diseases: Job's syndrome and Omenn syndrome
Transplant rejections

Endocrine
Hypoadrenalism: Addison disease, adrenal hemorrhage, hypopituitarism

Other Causes of Eosinophilia
Atheromatous cholesterol embolization
Hereditary
Serosal surface irritation, including peritoneal dialysis and pleural eosinophilia

Adapted from Weller PF. Eosinophilia and eosinophil-related disorders. In: Adkinson NF, Jr., Yunginger JW, Busse WW, et al., eds. Allergy: Principles and Practice, 6th ed. Philadelphia, Penn.: Mosby; 2003:1105.

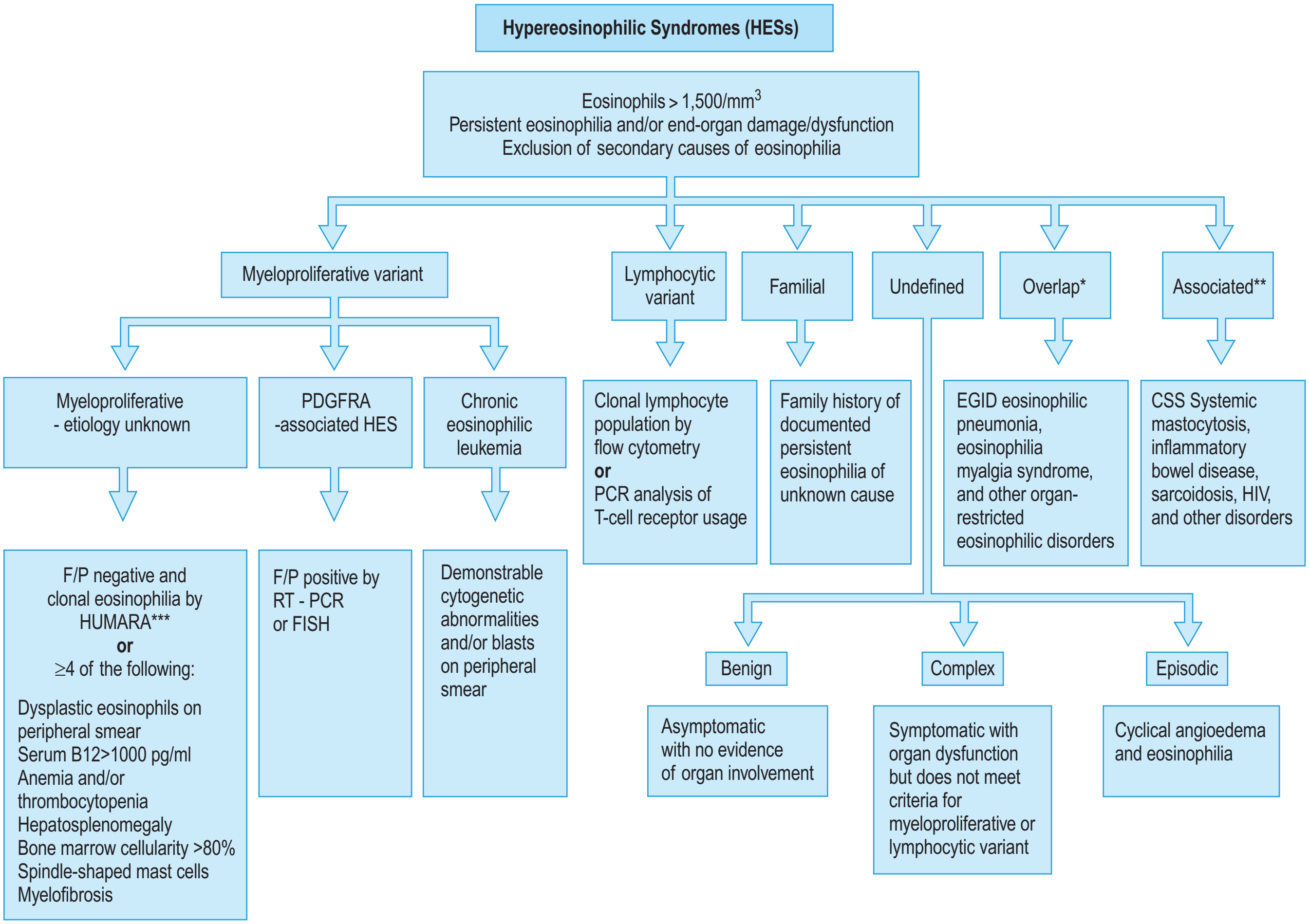

Hypereosinophilic Syndromes (HESs)
Eosinophils > 1,500/mm3
Persistent eosinophilia and/or end-organ damage/dysfunction
Exclusion of secondary causes of eosinophilia
Myeloproliferative variant
Lymphocytic variant
Familial
Undefined
Overlap*
Associated**
Myeloproliferative - etiology unknown
PDGFRA -associated HES
Chronic eosinophilic leukemia
Clonal lymphocyte population by flow cytometry or PCR analysis of T-cell receptor usage
Family history of documented persistent eosinophilia of unknown cause
EGID eosinophilic pneumonia, eosinophilia myalgia syndrome, and other organ-restricted eosinophilic disorders
CSS Systemic mastocytosis, inflammatory bowel disease, sarcoidosis, HIV, and other disorders
F/P negative and clonal eosinophilia by HUMARA*** or ≥4 of the following:
Dysplastic eosinophils on peripheral smear
Serum B12>1000 pg/ml
Anemia and/or thrombocytopenia
Hepatosplenomegaly
Bone marrow cellularity >80%
Spindle-shaped mast cells
Myelofibrosis
F/P positive by RT - PCR or FISH
Demonstrable cytogenetic abnormalities and/or blasts on peripheral smear
Benign
Complex
Episodic
Asymptomatic with no evidence of organ involvement
Symptomatic with organ dysfunction but does not meet criteria for myeloproliferative or lymphocytic variant
Cyclical angioedema and eosinophilia

FIG 24.4 **Classification of Hypereosinophilic Syndromes Based on a Workshop Summary Report.** Specific syndromes discussed at the workshop are indicated in bold. *Incomplete criteria, apparent restriction to specific tissues/organs. †Peripheral eosinophilia, >1500/mm^3 in association with a defined diagnosis. ‡ Presence of the FLPL1/PDGFRA (F/P) mutation. § Clonality analysis based on the digestion of genomic DNA with methylation-sensitive restriction enzymes followed by polymerase chain reaction (PCR) amplification of the CAG repeat at the human androgen receptor gene (HUMARA) locus at the X chromosome. CSS, Churg-Strauss syndrome (now called eosinophilic granulomatosis with polyangiitis); EGID, eosinophil gastrointestinal diseases; FISH, fluorescence *in situ* hybridization. (From Klion AD, Bochner BS, Gleich GJ, et al. Approaches to the treatment of hypereosinophilic syndromes: a workshop summary report. J Allergy Clin Immunol 2006; 117:1294, with permission from the American Academy of Allergy, Asthma and Immunology.)

Some patients with HES, termed *myeloproliferative variants of HES,* exhibit features common to myeloproliferative disorders, including elevated vitamin B_{12} and lactate dehydrogenase (LDH) levels, splenomegaly, cytogenetic abnormalities, myelofibrosis, anemia, myeloid dysplasia, and often elevated serum level of mast cell tryptase. In many patients with myeloproliferative HES, the molecular defect has been identified as a chromosome 4 deletion that yields a fusion gene encoding a FIP1LI/PDGFRA (PDGF-α receptor) (F/P) protein that constitutively expresses receptor kinase activity. This fusion gene can be diagnostically evaluated by reverse transcription–polymerase chain reaction (RT-PCR) or fluorescence *in situ* hybridization (FISH) (Chapter 96). Importantly, the majority of patients with this fusion mutation, which constitutes a form of chronic eosinophilic leukemia (CEL), respond to therapy with imatinib, which is considered the first line of therapy for FIP1LI/PDGFRA-positive HES.[26] For patients with any evidence of cardiac involvement, including elevated troponin levels, it is recommended that glucocorticoids be administered along with initiation of imatinib therapy. Other patients with eosinophilia without F/P mutations have also responded to imatinib, indicating that other receptor tyrosine kinase mutations can underlie some of these CEL/myeloproliferative forms of HES.[27] The presence of more than four myeloproliferative features commonly seen in mutation-positive disease, predicted response in those without known mutations. Some of these features included dysplastic eosinophils, vitamin B_{12} level >1000 picograms per milliliter (pg/mL), tryptase level ≥12 nanograms per milliliter (ng/mL), anemia/thrombocytopenia, hypercellular marrow, and spindled mast cells, reticulin fibrosis, and dysplastic megakaryocytes on bone marrow biopsy.[28] In addition, clonal abnormalities in the eosinophil lineage have been reported in a few patients (see Fig. 24.4).

Another variant form of HES is a lymphoproliferative form resulting from clonal expansions of lymphocytes, often $CD4^{+}CD3^{-}$ Th2-like lymphocytes, which elaborate IL-5.[29] These aberrant T cells can be sought by flow cytometry or T-cell receptor (TCR) analysis. These patients, who may have elevated IgE levels, usually do not develop eosinophilic endomyocardial disease but are at risk for developing T-cell lymphomas.[29]

In addition to these recognized variants, there are a substantial number of patients with HES for whom the etiologies of the eosinophilia remain unknown.[26] Some such patients develop no signs or symptoms of disease and can be monitored without therapy.[30] For those who require therapy, including those with lymphoproliferative variants,[29] glucocorticosteroids are the mainstay of treatment.[26] With glucocorticoid therapy, partial or complete remission of eosinophilia within 1 month has been reported to occur in 85% of patients.[26] Second-line agents include hydroxyurea and IFN-α.[26] The neutralizing anti-IL-5 monoclonal antibody (mAb), mepolizumab, has beneficial and steroid-sparing effects in those with FIP1L1-PDGFRA negative hypereosinophilic syndromes,[31] but it is approved by the US Food and Drug Administration (FDA) for treating severe eosinophilic phenotype asthma but not yet approved for hypereosinophilic syndromes. Anti-CD52 mAb (alemtuzumab) and allogeneic hematopoietic cell transplantation have been used for particularly severe and refractory HES.

In contrast to older reports, with earlier diagnosis and therapy and with more varied and targeted therapeutic options, morbidity and, particularly, mortality in HES syndromes have been reduced.

Eosinophilia With Tumors or Leukemias

The F/P-positive and related chromosomal fusion gene mediated myeloproliferative variants of HES are forms of CEL.[27] Eosinophilia is a characteristic of the M4Eo subtype of acute myeloid leukemia, having the common M4 characteristic of chromosomal 16 abnormalities. Other forms of eosinophilic leukemia, often with specific cytogenetic and molecular genetic abnormalities, have been recognized.[27] Eosinophilia may accompany chronic myelogenous leukemia (often with basophilia) but is uncommon with acute lymphoblastic leukemia. Eosinophilia may be observed in some patients with lymphoma, including Hodgkin disease, especially the nodular sclerosing form, T-cell lymphoblastic lymphoma, and adult T-cell leukemia/lymphoma. A small proportion of patients with carcinomas, especially of mucin-producing epithelial cell origin, have associated blood and tissue eosinophilia. Eosinophilia may accompany angioimmunoblastic lymphadenopathy, mycosis fungoides, Sézary syndrome, and lymphomatoid papulosis. Eosinophilia occurs in about 20% of patients with systemic mastocytosis and may be the presenting finding in the absence of cutaneous manifestations.

ORGAN SYSTEM INVOLVEMENT AND EOSINOPHILIA

Eosinophilic syndromes limited to specific organs, such as eosinophilic pneumonias or eosinophilic GI disorders (EGIDs; Chapter 46), characteristically do not extend beyond their own target organ, and hence lack the multiplicity of organ involvement often found with non–organ-specific hypereosinophilic syndromes. They also do not have the predilection to develop secondary eosinophil-mediated cardiac damage, for reasons that are not known.

Pulmonary Eosinophilias

Blood eosinophilia can infrequently accompany pleural fluid eosinophilia, a nonspecific response seen with various disorders, including trauma and repeated thoracenteses. In addition, several pulmonary parenchymal disorders can be associated with eosinophilia (see Table 24.2).[32]

Helminth parasites are responsible for four forms of eosinophilic lung disease.[24,32] The first form, Loeffler syndrome, is marked by blood eosinophilia, eosinophilic patchy pulmonary infiltrates that appear and resolve over weeks, and, at times, bronchospasm, and is typically caused by those helminth parasites (*Ascaris lumbricoides/Ascaris suum),* and less commonly hookworm and *Strongyloides* that migrate through the lungs early in their developmental lifecycle.[24] Stool examinations are not helpful, as the pulmonary response is elicited by infecting larval forms months before productive egg-laying from later adult stages begins in the intestines. Diagnosis is made on epidemiological grounds.[24]

The second form of helminth-induced lung disease is the syndrome of tropical pulmonary eosinophilia, which develops in a minority of patients infected with lymphatic-dwelling filarial species.[23] This syndrome is characterized by marked blood eosinophilia, a paroxysmal nonproductive cough, wheezing, occasional weight loss, lymphadenopathy, and low-grade fevers. On chest X-rays, increased bronchovesicular markings, diffuse interstitial lesions 1–3 mm in diameter or mottled opacities, usually more prominent in lower lung fields, are common. Patients have markedly increased numbers of blood and alveolar eosinophils, and elevations in both total serum IgE and antifilarial antibodies.

A third form of helminth-induced lung disease is caused by helminths that invade the pulmonary parenchyma, notably lung flukes that cause paragonimiasis.

The fourth form of lung disease is caused by larger than usual numbers of helminth organisms that are carried hematogenously into the lungs. Examples include schistosomiasis, trichinellosis, and larva migrans.

Bronchopulmonary aspergillosis constitutes another type of eosinophil-associated pulmonary disease. Two forms of idiopathic eosinophilic pneumonia are recognized.[32] In *chronic eosinophilic pneumonia,* patients may exhibit peripheral pulmonary infiltrates that can extend across lobar fissures. Chronic eosinophilic pneumonia is of unknown etiology and is responsive to corticosteroids but is prone to relapse. An *acute form of eosinophilic pneumonia,* which manifests as fever, pulmonary infiltrates, and respiratory insufficiency, is diagnosable by finding eosinophils in bronchoalveolar lavage (BAL) fluids or on lung biopsy. Acute eosinophilic pneumonia, which often follows new exposures to cigarette or other smoke or dusts, responds to corticosteroid treatment and does not relapse.

The major vasculitis associated with eosinophilia is eosinophilic granulomatosis with polyangiitis (EGPA, formerly called *Churg-Strauss syndrome*) (Chapter 58).[33] Late-onset asthma, eosinophilia, and at times transient pulmonary infiltrates antedate the development of systemic vasculitis in about half the cases. Pulmonary involvement is seen in almost all patients, and pulmonary infiltrates occur in three-quarters of them. Nasal and sinus involvement is common. Corticosteroids, anti-IgE mAb, or anticysteinyl leukotriene agent therapies for asthma may mask the evolution of EGPA. Neurological, cutaneous, cardiac, and GI organ involvement is common.[33] Cardiac involvement includes pericarditis and small vessel cardiac vasculitis, and, much less commonly, endomyocardial thrombosis and fibrosis.

Diverse drugs are capable of eliciting pulmonary eosinophilia. More commonly implicated medications include nonsteroidal antiinflammatory drugs (NSAIDs) and antimicrobial medications. Likewise, toxic agents, including those from occupational exposure, can be responsible for pulmonary eosinophilia. Each of these reactions has a defined etiological stimulus and hence differs from idiopathic and other eosinophilic diseases, but the clinical presentation of drug- and toxin-elicited pulmonary eosinophilias can resemble other forms of pulmonary eosinophilia, including acute or chronic eosinophilic pneumonia.

Skin and Subcutaneous Diseases

A number of cutaneous diseases can be associated with heightened blood eosinophils,[34] including atopic dermatitis, blistering disorders including bullous pemphigoid, drug reactions, and two diseases associated with pregnancy: (i) herpes gestationis and (ii) the syndrome of pruritic urticarial papules and plaques of pregnancy. Eosinophilic pustular folliculitis is seen mostly in patients with HIV infections and in those treated for hematological malignancies or after bone marrow transplantation. In patients with cutaneous involvement and eosinophilia, angiolymphoid hyperplasia with eosinophilia and Kimura disease, eosinophilic cellulitis (Wells syndrome), eosinophilic fasciitis, and eosinophilic pustular folliculitis can be differentiated on the basis of histopathology of biopsied lesions. Another syndrome, episodic angioedema with eosinophilia, is characterized by recurring episodes of angioedema, urticaria, fever, and marked blood eosinophilia. This syndrome responds to glucocorticosteroid therapy.

Gastrointestinal Diseases

EGIDs (Chapter 46), including eosinophilic esophagitis, eosinophilic gastroenteritis, and eosinophilic colitis, represent a heterogeneous collection of disorders in which there may be eosinophilic infiltration of the mucosa, the muscle layer or the serosa, the last of which can lead to eosinophilic ascites. Peripheral blood eosinophilia may occur in EGIDs, although with eosinophilic esophagitis, peripheral blood eosinophil counts are often normal. Eosinophils are present in the lesions of collagenous colitis and ulcerative colitis, but blood eosinophilia is usually absent. GI eosinophilia elicited by intestinal helminths and eosinophilic enterocolitis as a result of hypersensitivity reactions to medications must be excluded in patients with these diseases who have tissue eosinophilia.

Rheumatological Disorders

Of the various forms of vasculitis, only two are commonly associated with eosinophilia. The principal eosinophil-related vasculitis is EGPA, formerly called Churg-Strauss syndrome (as discussed above; and in Chapter 58). Cutaneous necrotizing eosinophilic vasculitis with hypocomplementemia and eosinophilia is a distinct vasculitis of small dermal vessels that are extensively infiltrated with eosinophils. This form of vasculitis may occur in patients with connective tissue diseases. In addition, eosinophilia may uncommonly accompany rheumatoid arthritis itself but more commonly results from adverse reactions to treatment medications (including NSAIDs, gold, and tetracyclines) or concomitant vasculitis.

Immunological Disorders

CLINICAL PEARLS

Eosinophilia and Drug Reactions

Drug Reactions	Examples
Interstitial nephritis	Semisynthetic penicillins, cephalosporins
Pulmonary infiltrates	Nitrofurantoin, sulfas, nonsteroidal antiinflammatory drugs (NSAIDs)
Pleuropulmonary	Dantrolene
Hepatitis	Semisynthetic penicillins, tetracyclines
Hypersensitivity vasculitis	Allopurinol, phenytoin
Asthma, nasal polyps	Aspirin
Eosinophilia–myalgia	L-Tryptophan contaminant syndrome
Asymptomatic	Ampicillin, penicillins, cephalosporins
Cytokine-mediated	Granulocyte macrophage–colony-stimulating factor (GM-CSF), interleukin (IL)-2
DRESS (drug reaction with eosinophilia and systemic symptoms)	Minocycline, allopurinol, anticonvulsants

Adapted from Weller PF. Eosinophilia and eosinophil-related disorders. In: Adkinson NF, Jr., Yunginger JW, Busse WW, et al., eds. Allergy: Principles and Practice, 6th ed. Philadelphia, Penn.: Mosby; 2003:1105.

Adverse reactions to medications are a common cause of eosinophilia. Although often considered as hypersensitivity reactions, in most instances of drug-associated eosinophilia, the mechanism leading to eosinophilia is not understood. Eosinophilia may develop without other manifestations of adverse drug reactions, such as rashes or drug fevers. In addition, drug-induced eosinophilia may be associated with distinct clinicopathological patterns in which eosinophilia accompanies drug-induced diseases that are characteristically limited to specific organs with or without associated blood eosinophilia. When organ dysfunction develops, cessation of drug administration is necessary. Drug-induced interstitial nephritis may be accompanied by blood eosinophilia, and eosinophils may be detectable in urine. Unlike granulocyte–colony-stimulating factor (G-CSF) therapy, therapy with GM-CSF can lead to prominent blood and tissue eosinophilia. Administration of IL-2 or of IL-2-stimulated lymphocytes can be followed by the development of eosinophilia, most likely as a result of stimulated production of IL-5. Reactions to medications, often anticonvulsants, minocycline, and allopurinol, can elicit DRESS (*d*rug *r*eaction with *e*osinophilia and *s*ystemic *s*ymptoms).[34] In addition to cutaneous eruptions, fever, lymphadenopathy, hepatitis, nephritis, atypical lymphocytosis, GI tract involvement, and eosinophilia are common but variable elements of this drug-induced syndrome, which can be fatal. The triggering medication must be stopped, and corticosteroids are often administered.

Some primary immunodeficiency syndromes are associated with eosinophilia.[35] Hyper-IgE syndrome is characterized by recurrent staphylococcal abscesses of the skin, lungs, and other sites; pruritic dermatitis; hyperimmunoglobulinemia E; and eosinophilia of blood, sputum, and tissues. Eosinophilia is characteristic of Omenn syndrome, combined immunodeficiency with hypereosinophilia (Chapter 35).

Infiltration of eosinophils accompanies rejection of lung, kidney, and liver allografts. Tissue and blood eosinophilia occur early in the rejection process, and eosinophil counts and eosinophil granule protein levels (in urine, BAL fluids, and involved allograft tissues) have correlated with prognosis, severity, and response to rejection therapy.

Endocrine Diseases

The loss of endogenous adrenoglucocorticosteroid production in Addison disease, adrenal hemorrhage, or hypopituitarism can cause increased blood eosinophilia, although usually not more than mild to moderate.

Other Causes of Eosinophilia

The syndrome of atheromatous cholesterol embolization is at times associated with hypocomplementemia, eosinophilia, and eosinophiluria. Rarely, cases of hereditary eosinophilia among family members have been recognized. Irritation of serosal surfaces can be associated with eosinophilia, and related diseases can include Dressler syndrome; eosinophilic pleural effusions; peritoneal and, at times, blood eosinophilia developing during chronic peritoneal dialysis; and perhaps the eosinophilia that follows abdominal irradiation.

THERAPEUTIC PRINCIPLES

Therapy of Specific Eosinophilic Diseases

Eosinophil-Associated Diseases With Identifiable Etiologies	
Parasitic infections	Treat causative parasite
Drug-reaction related eosinophilias	Terminate eliciting medication
Adrenal insufficiency	Corticosteroid replacement therapy
Allergic/atopic diseases	Varied, may include topical or inhaled corticosteroids
Distinct Eosinophilic Syndromes Involving Specific Organs	
Eosinophilic pulmonary diseases:	
Acute eosinophilic pneumonia	Corticosteroids
Chronic eosinophilic pneumonia	Corticosteroids, interferon-α
Eosinophilic granulomatosis with polyangiitis	Corticosteroids, interferon-α
Hypereosinophilic Syndromes	
F/P-positive myeloproliferative	Imatinib
Lymphoproliferative and other	Corticosteroids, interferon-α, hydroxyurea, anti–interleukin (IL)-5 monoclonal antibody (mAb), other

EVALUATION OF EOSINOPHILIA

Because a diversity of disorders can be accompanied by eosinophilia, evaluation of the patient with eosinophilia requires a consideration of features based on the patient's history, physical examination, and other laboratory, radiographic, or diagnostic tests.[26,36,37] An initial approach can focus on identifying eosinophilic diseases that have a defined treatable etiology. These include infections with helminth parasites, and for these, the approach should be guided by information obtained from patient history about potential exposures; results of the history and physical examinations with regard to signs and symptoms of any clinically apparent associated illness; results of standard biochemical and radiographic tests for evidence of organ involvement; and results of specific parasitological tests, including potentially stool, urine, blood, sputum, or tissue examinations, as well as results of serological tests.[24] The duration and magnitude of the eosinophilia may suggest some entities, especially if it is prolonged or markedly elevated (see Table 24.1). Other causes of eosinophilia that are amenable to treatment include eosinophilia secondary to medications, for which cessation of the offending drug may be indicated if the eosinophilia is accompanied by organ damage. Likewise, if eosinophilia is secondary to glucocorticosteroid deficiency, diagnostic testing can corroborate this deficiency and lead to the administration of replacement corticosteroids and consequent resolution of the eosinophilia.

Because allergic diseases usually are associated with at least some degree of eosinophilia, clinical and laboratory evidence of such disease should be sought. If the eosinophilia is not attributable to allergic diseases, parasitic infections, medications, or steroid deficiency, further evaluation will be guided by whether the patient has evidence of organ disease and, if so, which organs are involved (see Table 24.2). This is germane, for instance, in defining whether the patient has a distinct eosinophilic pulmonary, GI, or cutaneous syndrome. Bone marrow examinations in most patients with eosinophilia are not usually informative, revealing only evidence of enhanced eosinophilopoiesis; but bone marrow should be examined if there is suspicion of a hematological malignancy or myeloproliferative disorder. For patients with sustained eosinophilia who meet the criteria for HES, diagnostic testing should aim to identify which variant form of HES the patient may have, in which case, bone marrow examination is often needed (see Fig. 24.4).

ON THE HORIZON

- Identify the causes of those hypereosinophilic syndromes for which etiologies currently remain unknown.
- Delineation of the signaling mechanisms that govern the agonist-specific, differential secretion of cytokines that are preformed and stored within eosinophil granules and secretory vesicles.
- Continue to evaluate the therapeutic efficacies of anticytokine therapeutics, including anti–interleukin (IL)-5 neutralizing antibodies, in the treatment of the varied forms of eosinophilic diseases.
- Identify biomarkers that are predictive of eosinophil-mediated tissue damage and that can be used for clinical diagnostic and therapeutic monitoring tools.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Blanchard C, Rothenberg ME. Biology of the eosinophil. Adv Immunol 2009;101:81–121.
2. Rosenberg HF, Dyer KD, Foster PS. Eosinophils: changing perspectives in health and disease. Nat Rev Immunol 2013;13:9–22.
3. Spencer LA, Weller PF. Eosinophils and Th2 immunity: contemporary insights. Immunol Cell Biol 2010;88:250–6.
4. Shamri R, Xenakis JJ, Spencer LA. Eosinophils in innate immunity: an evolving story. Cell Tissue Res 2011;343:57–83.
5. Nussbaum JC, Van Dyken SJ, von Moltke J, et al. Type 2 innate lymphoid cells control eosinophil homeostasis. Nature 2013;502:245–8.
6. Bochner BS. Siglec-8 on human eosinophils and mast cells, and Siglec-F on murine eosinophils, are functionally related inhibitory receptors. Clin Exp Allergy 2009;39:317–24.
7. Bozza PT, Magalhaes KG, Weller PF. Leukocyte lipid bodies—biogenesis and functions in inflammation. Biochim Biophys Acta 2009;1791:540–51.

8. Melo RCN, Perez SAC, Spencer LA, et al. Intragranular vesiculotubular compartments are involved in piecemeal degranulation by activated human eosinophils. Traffic 2005;6:866–79.
9. Hogan SP, Rosenberg HF, Moqbel R, et al. Eosinophils: biological properties and role in health and disease. Clin Exp Allergy 2008;38:709–50.
10. Spencer LA, Szela CT, Perez SA, et al. Human eosinophils constitutively express multiple Th1, Th2, and immunoregulatory cytokines that are secreted rapidly and differentially. J Leukoc Biol 2009;85:117–23.
11. Melo RC, Weller PF. Piecemeal degranulation in human eosinophils: a distinct secretion mechanism underlying inflammatory responses. Histol Histopathol 2010;25:1341–54.
12. Spencer LA, Melo RCN, Perez SAC, et al. Cytokine receptor-mediated trafficking of preformed IL-4 in eosinophils identifies an innate immune mechanism of cytokine secretion. Proc Natl Acad Sci USA 2006;103:3333–8.
13. Neves JS, Perez SA, Spencer LA, et al. Eosinophil granules function extracellularly as receptor-mediated secretory organelles. Proc Natl Acad Sci USA 2008;105:18478–83.
14. Neves JS, Weller PF. Functional extracellular eosinophil granules: novel implications in eosinophil immunobiology. Curr Opin Immunol 2009;21:694–9.
15. Ueki S, Melo RC, Ghiran I, et al. Eosinophil extracellular DNA trap cell death mediates lytic release of free secretion-competent eosinophil granules in humans. Blood 2013;121:2074–83.
16. Klion AD, Nutman TB. The role of eosinophils in host defense against helminth parasites. J Allergy Clin Immunol 2004;113:30–7.
17. Swartz JM, Dyer KD, Cheever AW, et al. *Schistosoma mansoni* infection in eosinophil lineage-ablated mice. Blood 2006;108:2420–7.
18. Padigel UM, Hess JA, Lee JJ, et al. Eosinophils act as antigen presenting cells to induce immunity to Strongyloides stercoralis in mice. J Infect Dis 2007;196:1844–51.
19. Fabre V, Beiting DP, Bliss SK, et al. Eosinophil deficiency compromises parasite survival in chronic nematode infection. J Immunol 2009;182:1577–83.
20. Akuthota P, Wang HB, Spencer LA, et al. Immunoregulatory roles of eosinophils: a new look at a familiar cell. Clin Exp Allergy 2008;38:1254–63.
21. Wang HB, Ghiran I, Matthaei K, et al. Airway eosinophils: allergic inflammation recruited professional antigen-presenting cells. J Immunol 2007;179:7585–92.
22. Chu VT, Frohlich A, Steinhauser G, et al. Eosinophils are required for the maintenance of plasma cells in the bone marrow. Nat Immunol 2011;12:151–9.
23. Wu D, Molofsky AB, Liang HE, et al. Eosinophils sustain adipose alternatively activated macrophages associated with glucose homeostasis. Science 2011;332:243–7.
24. Wilson ME, Weller PF. Eosinophilia. In: Guerrant RL, Walker DH, Weller PF, editors. Tropical infectious diseases: principles, pathogens and practice. 3rd ed. Philadelphia, Penn.: Elsevier; 2011. p. 939–49.
25. Klion AD, Bochner BS, Gleich GJ, et al. Approaches to the treatment of hypereosinophilic syndromes: a workshop summary report. J Allergy Clin Immunol 2006;117:1292–302.
26. Ogbogu PU, Bochner BS, Butterfield JH, et al. Hypereosinophilic syndrome: a multicenter, retrospective analysis of clinical characteristics and response to therapy. J Allergy Clin Immunol 2009;124:1319–25.
27. Valent P, Gleich GJ, Reiter A, et al. Pathogenesis and classification of eosinophil disorders: a review of recent developments in the field. Expert Rev Hematol 2012;5:157–76.
28. Khoury P, Desmond R, Pabon A, et al. Clinical features predict responsiveness to imatinib in platelet derived growth factor receptor alpha-negative hypereosinophilic syndrome. Allergy 2016;71:803–10.
29. Lefevre G, Copin MC, Staumont-Salle D, et al. The lymphoid variant of hypereosinophilic syndrome: study of 21 patients with CD3-CD4+ aberrant T-cell phenotype. Medicine (Baltimore) 2014;93:255–66.
30. Chen YY, Khoury P, Ware JM, et al. Marked and persistent eosinophilia in the absence of clinical manifestations. J Allergy Clin Immunol 2014;133:1195–202.
31. Rothenberg ME, Klion AD, Roufosse FE, et al. Treatment of patients with the hypereosinophilic syndrome with mepolizumab. N Engl J Med 2008;358:1215–28.
32. Akuthota P, Weller PF. Eosinophilic pneumonias. Clin Microbiol Rev 2012;25:649–60.
33. Comarmond C, Pagnoux C, Khellaf M, et al. Eosinophilic granulomatosis with polyangiitis (Churg-Strauss): clinical characteristics and long-term follow-up of the 383 patients enrolled in the French Vasculitis Study Group cohort. Arthritis Rheum 2013;65:270–81.
34. Long H, Zhang G, Wang L, et al. Eosinophilic skin diseases: a comprehensive review. Clin Rev Allergy Immunol 2016;50:189–213.
35. Ben m'rad M, Leclerc-Mercier S, Blanche P, et al. Drug-induced hypersensitivity syndrome: clinical and biologic disease patterns in 24 patients. Medicine (Baltimore) 2009;88:131–40.
36. Williams KW, Milner JD, Freeman AF. Eosinophilia associated with disorders of immune deficiency or immune dysregulation. Immunol Allergy Clin North Am 2015;35:523–44.
37. Klion A. Hypereosinophilic syndrome: current approach to diagnosis and treatment. Annu Rev Med 2009;60:293–306.

MULTIPLE-CHOICE QUESTIONS

1. Which of the following may NOT contribute to or elicit eosinophilia?
   A. $CD3^-$, $CD4^+$ aberrant lymphocyte subsets
   B. FIP1L1-PDGFRα chromosomal gene rearrangement
   C. *Giardia lamblia* infection
   D. *Strongyloides stercoralis* infection

2. A 45-year-old male who has had eosinophilia of between 2500 and 5000/ mm^3 documented for the past 6 months has been referred to you. Your evaluation could include:
   A. Flow cytometry for lymphocyte phenotyping
   B. Serum immunoglobulin E (IgE)
   C. Serum troponin
   D. *Strongyloides* enzyme-linked immunosorbent assay (ELISA) serology
   E. All of the above

3. Which three cytokines promote eosinophil development in bone marrow and sustain eosinophil viability?
   A. Transforming growth factor (TGF)-β, interleukin (IL)-2, IL-4
   B. IL-7, IL-2, IL-13
   C. Granulocyte macrophage–colony-stimulating factor (GM-CSF), IL-3, IL-5
   D. Interferon (INF)-γ, IL-8, IL-10

25

Host Defenses to Viruses

Barry T. Rouse, Scott N. Mueller

Viruses as obligate intracellular parasites require their host to replicate them and to facilitate their spread to others. In humans, most clinically relevant infections were derived from other animals, and this process continues. Recent examples include human immunodeficiency virus (HIV), Ebola virus, severe acute respiratory syndrome (SARS) virus, and Zika virus. Viral infections are rarely lethal, even if they are highly cytolytic to individual cells. Mortality commonly occurs when viruses jump species, when the virus undergoes a major antigenic change (*i.e.*, influenza viruses), or when host immunity is compromised. HIV (Chapter 39) represents one of the most dramatic human examples of an exotic virus that kills its host. However, HIV kills slowly, providing ample time to spread to new hosts and an effective strategy for persistence in the species. Death or dire consequences following virus infection in mammals with inadequate immunity are well illustrated by observations that fetuses or neonates, especially if deprived of passive immunity, succumb to many agents well tolerated by healthy adults. The science of viral immunology seeks to understand mechanisms of virus–host interactions with a view to applying this knowledge to the design of effective vaccines and immunomodulators that control virus infections. These objectives are facilitated by an increasing wealth of immunological techniques, an expanding array of genetically manipulated animal models, and an abundance of high throughput technologies, which generate data that can be subjected to complex computational analysis. Such analyses can yield signatures indicative of optimal immunogenicity and vaccine efficacy or failure and can explain the variable outcome of infections in individual hosts. In most situations, defense against viruses involves multiple immune components, and the impact of a single mechanism varies greatly according to the method by which individual viruses enter, replicate, and spread within the host. In this chapter, we highlight the principal means by which the host achieves immunity after infection by viruses. Table 25.1 presents an overview.

VIRAL ENTRY AND INFECTION

Access to target tissues presents numerous obstacles for entry and infection by most human viruses. Most effective of these are the mechanical barriers provided by skin and the mucosal surfaces, as well as the chemically hostile environment of the gut (Fig. 25.1). A number of common human viral pathogens enter through the gastrointestinal tract, including rotavirus, enteric adenoviruses, and hepatitis A virus (HAV). These are usually spread via person-to-person contact or contaminated food and water. Respiratory infections caused by influenza viruses, rhinoviruses, coronaviruses, measles virus, varicella-zoster virus (VZV), and respiratory syncytial virus (RSV) are often spread by aerosol transmission, as well as person-to-person contact. Many of the herpes viruses target the skin or the mucosae, such as herpes simplex virus (HSV) and VZV. HSV, in particular, can infect the oral and genital mucosae, the eye, and skin through small cuts and abrasions. Other herpes viruses, such as Epstein-Barr virus (EBV) and cytomegalovirus (CMV), target mucosae. CMV can also spread vertically from mother to baby or rarely via blood transfusions. Human papillomavirus (HPV) targets skin and mucosae and causes warts and may transform cells, inducing cancers, such as cervical cancer. Some viruses, such as West Nile virus, Dengue virus, Semliki forest virus, and Zika virus, can enter through skin via insect vectors. HIV and hepatitis B virus (HBV) are commonly spread via sexual contact. HIV, HBV, and hepatitis C virus (HCV) can also infect humans by direct entry into the bloodstream via transfusions or contaminated needles.

Most human viruses replicate only in certain target tissues, this being mainly the consequence of viral receptor distribution. Many viruses use two receptors, such as the use of the CD4 coreceptor and the chemokine receptor CCR5 on T cells by HIV. After attachment to a cellular receptor, viruses may fuse with the cell membrane or be endocytosed and then gain entry into the cytoplasm or nucleus by fusing with the vesicular membrane (enveloped viruses, such as HSV and HIV), or translocate across the cell membrane or induce lysis of the endocytic vesicle once in the cytoplasm (nonenveloped viruses, such as Norwalk virus and poliovirus).[1] Viruses then utilize host cell machinery and specialized virally encoded proteins to replicate rapidly within the cell. Once they have multiplied within the cell, many viruses induce cytolysis to facilitate release of new infectious virions (*e.g.*, poxviruses, poliovirus, and herpes viruses). Other viruses are released from infected cells by budding through the cell membrane in the absence of cell death (*e.g.*, HIV and influenza virus). Having entered the body, however, viruses encounter numerous innate defenses and activate the components of adaptive immunity. The latter usually assures that clinical disease, if not infection, will not become evident. Successful exploitation of these defenses through the use of vaccines (Chapter 90) remains a central challenge for many human viruses, particularly those that cause chronic infections, such as HIV and HCV.[2]

INNATE IMMUNITY TO VIRUSES

Viral infection induces an extensive array of defense mechanisms in the host. Innate defenses come into play to block or inhibit

TABLE 25.1 Viral Infections and Immunity

Viral Event	Obstacles	Time Course
Transmission	Mechanical and chemical barriers	0
Infection and replication	Innate immunity	0 →
Infection stopped or spreads	Viral antigens transported to lymphoid tissues	Within 24 hours
Infection controlled	Specific antibodies and cell-mediated immunity	4–10 days
Sterile immunity	Immune memory	14 days to years
Viral persistence if infection not controlled	Immune disruption or evasion	Weeks to years

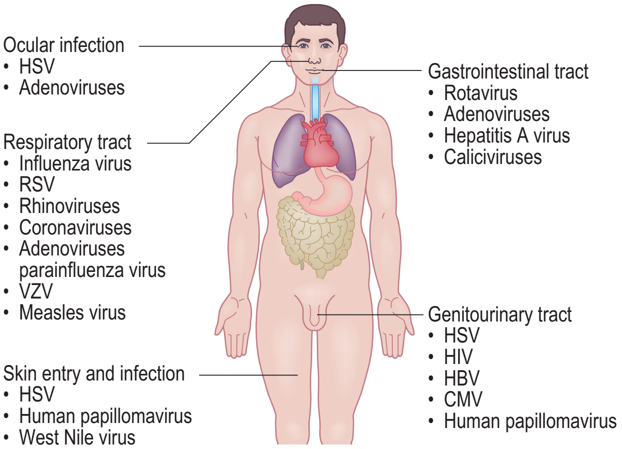


FIG 25.1 Common Routes of Entry and Infection for Human Viral Pathogens. CMV, cytomegalovirus; HBV, hepatitis B virus; HIV, human immunodeficiency virus; HSV, herpes simplex virus; RSV, respiratory syncytial virus; VZV, varicella-zoster virus.

TABLE 25.2 Sensors of Viral Infection

Toll-Like Receptors (TLRs)	
TLR3	dsRNA, MCMV, VSV, LCMV, HSV, EBV
TLR7 and TLR8	ssRNA, Influenza virus, HIV, VSV
TLR9	dsDNA, HSV, MCMV
TLR2	MV hemagglutinin protein, HSV, HCMV
TLR4	MMTV envelope protein, RSV
RIG-I-Like Helicases (RLHs)	
RIG-I	Influenza virus, VSV, HCV, JEV, MV, RSV, Sendai virus, EBV
MDA-5	Poly(I:C), MV, Sendai virus, VSV, MCMV, Picornaviruses
NOD-Like Receptors (NLRs)	
NLRP3	Influenza virus, Sendai virus, Adenovirus, Vaccinia virus
NOD2	Influenza virus, VSV, RSV
Other sensors	
AIM2	Vaccinia virus, MCMV
ZBP1 (DAI)	Cytosolic dsDNA, HSV
IFI16	Cytosolic dsDNA, HSV
cGAS	Cytosolic dsDNA, HSV

AIM2, absent in melanoma-2; IFI16, Gamma-interferon-inducible protein Ifi-16; cGAS, cyclic GMP-AMP Synthase; ZBP1, Z-DNA-binding protein 1; DAI, DNA-dependent activator of IFN; dsRNA, double-strand RNA; EBV, Epstein-Barr virus; HCMV, human cytomegalovirus; HCV, hepatitis C virus; HIV, human immunodeficiency virus; HSV, herpes simplex virus 1/2; JEV, Japanese encephalitis virus; LCMV, lymphocytic choriomeningitis virus; MCMV, murine cytomegalovirus; MDA-5, melanoma differentiation-associated gene; MMTV, mouse mammary tumor virus; MV, measles virus; NLR, NOD-like receptor; RLH, RIG-I-like helicase; RSV, respiratory syncytial virus; ssRNA, single-strand RNA; TLR, Toll-like receptor; VSV, vesicular stomatitis virus.

initial infection, to protect cells from infection, or to eliminate virus-infected cells. Innate mechanisms occur well before the effectors of adaptive immunity become active, but they are critical for the initiation of adaptive immunity via the elicitation of inflammation that promotes immune cell activation. The innate immune defenses are initiated via pattern recognition receptors (PRRs), which recognize pathogen-associated molecular patterns (PAMPs)[3] (Chapter 3). These include transmembrane receptors of the Toll-like receptor (TLR) family, two families of intracellular receptors including the NOD-like receptors (NLRs) and the RIG-I–like helicases (RLHs), as well as the sensor molecule absent in melanoma-2 (AIM2). Additionally, the molecules cyclic guanosine monophosphate–adenosine monophosphate (GMP-AMP) synthase (cGAS), DDX41, IFI16, and Z-DNA–binding protein 1 (ZBP1) can sense cytosolic DNA (Table 25.2). These cellular sensors promote the expression of interleukin-1 (IL-1) and IL-18, type I (α/β) interferon (IFN-I), and a variety of IFN-stimulated genes and inflammatory cytokines, and chemokines. TLRs are cell surface or endosomal membrane–bound proteins expressed by numerous cells, including dendritic cells (DCs), macrophages, lymphocytes, and parenchymal cells. Expression of TLRs is largely inducible in most cell types, although some (TLR7/8/9) are constitutively expressed at high levels by specialized plasmacytoid DCs for rapid IFN production. Different TLR molecules recognize specific viral products, such as single- and double-stranded RNA (TLR 3 and TLR7/8, respectively) or double-stranded DNA (TLR9).

The RLHs retinoic acid-inducible gene I (*RIG-I*) and melanoma differentiation-associated gene (*MDA-5*) mediate cytoplasmic recognition of viral nucleic acids. These activate mitochondrial antiviral signaling (MAVS) proteins to stimulate IFN-I production and activate inflammasomes, which are molecular complexes that facilitate the activation of caspases and induce the production of proinflammatory IL-1β and IL-18.[4] NLRs are a second class of cytosolic sensors of PAMPs that activate inflammasomes via the adapter protein ASC. These include the NLRP (or NALP), NOD, and IPAF/NAIP receptors. Three major inflammasomes have been shown to be involved in antiviral immunity: the NLRP3 inflammasome, the RIG-I inflammasome, and the AIM2 inflammasome.[3]

The innate defense system consists of multiple cellular components and many specialized proteins. The longest known and best-studied antiviral proteins are the α/β IFNs, which act by binding to the type I IFN receptor and result in the transcription of more than 100 IFN-stimulated genes. One consequence of this "antiviral state" is the inhibition of cell protein synthesis and the prevention of viral replication.[5] Multiple leukocyte subsets are involved in innate defense, including macrophages, DCs, neutrophils, natural killer (NK) cells, natural killer T cells (NKT cells), and γδT cells. Furthermore, tissue cells, including fibroblasts, epithelial cells, and endothelial cells, express PRRs and respond to viral infection via the production of innate cytokines, including IFN-I and IL-1. IFN-I is a critical link between the innate and adaptive immune system, via activation of DCs and T cells, as well as protecting T cells from NK cell-mediated attack.[6] IFN-Is can also activate NK cells and induce other cytokines that promote

KEY CONCEPTS

Major Antiviral Innate Defense Mechanisms

- Acting to block infection:
  - Natural antibodies
  - Complement components
  - Some cytokines and chemokines
- Acting to protect cells from infection:
  - Interferon-α/β
  - Interferon-γ (IL-γ)
  - IL-1, IL-18
- Acting to destroy or inhibit virus-infected cells:
  - Natural killer (NK) cells
  - Natural killer T cells (NKT cells)
  - Macrophages
  - Neutrophils
  - γδ T cells
  - Nitric oxide
- Involved in regulating antiviral inflammatory response:
  - ILs-1, 6, 10, 12, 18, 23, 33
  - Transforming growth factor (TGF)-β
  - Chemokines (CCL2, 3, 4, 5)

NK responses, such as IFN-γ and IL-12. NK cells produce proinflammatory cytokines; they can kill infected cells and interact with DCs, and are an important component of innate defense against viruses. NK cells can protect against some herpes viruses, which downregulate major histocompatibility complex (MHC) expression in the cells they infect. NK cells are also important in resistance to mouse and human CMV and possibly to HIV, influenza virus, and Ebola virus.[7] NK cells have also recently been shown to possess traits of adaptive immunity and, like T and B cells, can form populations of memory cells.[8] NK cells are regulated by an array of activating and inhibitory receptors, whose expression and function are just beginning to be understood. Uninfected cells are usually protected from NK cell cytolysis as they deliver negative signals, such as high expression of MHC molecules. In contrast, virus-infected cells are killed either because they deliver positive signals or because they lack adequate MHC-negative signals. NK cells may also control excessive immune responses to viruses by killing CD4$^+$ T cells and indirectly regulating cytotoxic T lymphocyte (CTL) responses. NKT cells may provide some antigen-specific innate immune protection against certain viruses, such as influenza virus.[9]

Several classes of innate host proteins function in antiviral defense. These include natural antibodies, which may play a role in defense against some viral infections, as well as pentraxins and complement proteins.[10] Some viruses may be directly inactivated by complement activation or be destroyed by phagocytic cells that bind and ingest complement-bound virions. Several proinflammatory cytokines and chemokines induced by virus infection also play key roles in defense. Foremost among these is IL-1 and other members of the IL-1 family, including IL-18 and IL-33.[11] These cytokines influence both innate and adaptive immune cells and play critical roles in antiviral defense. Other antiviral cytokines are produced early following infection, such as TNF-α, IFN-γ, IL-12, IL-6, and chemokines, such as MIP-1α. In particular, IL-12 is a potent inducer of IFN-γ from NK cells. Inflammatory chemokines may also play an important role in innate antiviral defense by orchestrating macrophage, neutrophil, DC, and NK cell responses at the site of infection. Not only are these components of innate immunity involved in mediating initial protection against viruses; several components (*e.g.,* the PRRs; the cytokines IFN-I, IL-I, IL-33, and IL-12; and phagocytes, including macrophages, monocytes, and DCs) serve to shape the nature and effectiveness of the subsequent adaptive response to viral pathogens. For instance, DCs require innate signals, such as IFN-I and IL-12, for maturation and optimal T-cell activation. Furthermore, CD8$^+$ T cells responding to viruses need IFN-I and IL-33 signals for expansion and memory formation. Thus both the magnitude and the type of innate response induced by virus infection have a marked influence on the generation of adaptive immune responses.

ADAPTIVE IMMUNITY TO VIRUSES

Innate immunity generally only slows, rather than stops, viral infection, allowing time for the adaptive immune response to begin. The two major divisions of adaptive immunity, antibody-mediated and T-cell–mediated, are mainly directed at different targets. Antibodies usually function by binding to free viral particles and, in so doing, block infection of the host cell (Chapter 15). In contrast, T cells act principally by recognizing and destroying virus-infected cells or by orchestrating an inflammatory response that includes several antiviral components (Chapters 16, 17). As all viruses replicate within cells and many can spread directly between cells without reentering the extracellular environment, resolution of infection is reliant more on T-cell function than on antibody function. However, broadly neutralizing antiviral antibodies have the potential to be effective therapies against many different human infections, including HIV, influenza viruses, and Ebola virus. Recent advances have allowed researchers to isolate and identify human monoclonal antibodies (mAbs) against these and other pathogens,[12] offering promise of new therapies as well as significant insight for vaccine design. Antiviral antibodies are also very important as an immunoprotective barrier against reinfection. It is the presence of antibodies at portals of entry—most often mucosal surfaces—that is of particular relevance to influenza, HSV, and HIV infections.[13] Yet, how to generate vaccines that induce optimal antibody responses, including broadly neutralizing antibodies, remains an important unsolved problem.

Initiation of adaptive immunity is closely dependent on early innate mechanisms that activate antigen-presenting cells (APCs), principally subsets of DCs. APCs and lymphocytes are drawn into lymphoid tissues by chemokine and cytokine signals and are retained there for a few days to facilitate effective intercellular interactions. The architecture of the secondary lymphoid tissues supports the coordinated interactions among the cells of the adaptive immune system[14] through a network of supportive stromal cells and local chemokine gradients[15](Chapter 2). The induction events occur in lymph nodes draining an infection site or in the spleen if virus enters the bloodstream. The passage of viral antigens to lymph nodes usually occurs in DCs. Some viruses are able to compromise the function of APCs, such as HSV and measles virus, which can inhibit DC maturation.

B-cell activation occurs following antigen encounter in the B-cell follicles, and possibly the T-cell zones, in the spleen or lymph nodes. Some activated B cells become short-lived plasma cells, whereas others move to the edges of the B-cell follicles and interact with antigen-specific helper CD4$^+$ T cells via presentation of antigenic peptides on B-cell MHC class II molecules. These Bcl6-dependent CD4 T follicular helper (Tfh) cells are specialized for providing help for B-cell responses and are needed to promote and regulate B-cell responses.[16] Activated B cells initiate germinal center (GC) reactions with the help of CD4 Tfh cells, ensuring somatic hypermutation and affinity maturation for the selection of high-affinity, antibody-producing, long-lived plasma cells, as well as memory B cells.[17] At the molecular level, upregulation of the transcription factors Blimp-1, XBP-1, and IRF-4 dictates plasma cell formation, whereas Pax-5 expression delineates B cells destined for GC reactions and the memory B-cell lineage.

Antibody binding to epitopes expressed by native proteins at the surface of free virions usually blocks viral attachment or penetration of target cells. Sometimes the consequence is viral lysis (with complement proteins also involved), opsonization, or sensitization for destruction by Fc receptor–bearing cells that mediate antibody-dependent cellular cytotoxicity (ADCC). Occasionally, however, Fc receptor binding of antibody-bound virus may facilitate infection and result in more severe tissue damage. This occurs in Dengue fever and may happen in some instances in HIV infection. The antibody involved in the protection of mucosal surfaces in humans is predominantly secretory immunoglobulin A (IgA), but serum-derived IgG may also be protective, particularly in such sites as the vaginal mucosa.[13] Both antibody isotypes act mainly to block infection of epithelial cells, although in some instances, the antibody may transport antigen from within the body across epithelial cells to the outside. Mucosal antibody persists for a much shorter period compared with serum antibody, which explains, in part, why immunity to mucosal pathogens is usually of much shorter duration compared with immunity to systemic viral infections.

KEY CONCEPTS

Antiviral T- and B-Cell Immunity

Effector Systems	Recognized Molecules	Control Mechanisms
Antibody	Surface proteins or virions	Neutralization of virus, opsonization, or destruction of infected cells by ADCC
Antibody + complement	Surface proteins expressed on infected cells	Infected cell destruction by ADCC or complement-mediated lysis
Mucosal antibody (IgA)	Surface proteins or virions	Viral neutralization, opsonization, and transcytosis
$CD4^+$ T cells	Viral peptides (10–20 mers) presented on MHC class II surface, internal or nonstructural proteins presented by APCs	Antiviral cytokine and chemokine production; help for $CD8^+$ T-cell and B-cell responses; killing infected cells; regulatory functions to reduce immunopathology
$CD8^+$ T cells	Viral peptides (8–10 mers) presented on MHC class I surface, internal, or nonstructural proteins presented on infected cells or by cross-presentation	Killing infected cells or purging virus without cell death; antiviral cytokine and chemokine production

ADCC, antibody-dependent cellular cytotoxicity; APC, antigen-presenting cell; IgA, immunoglobulin A; MHC, major histocompatibility complex.

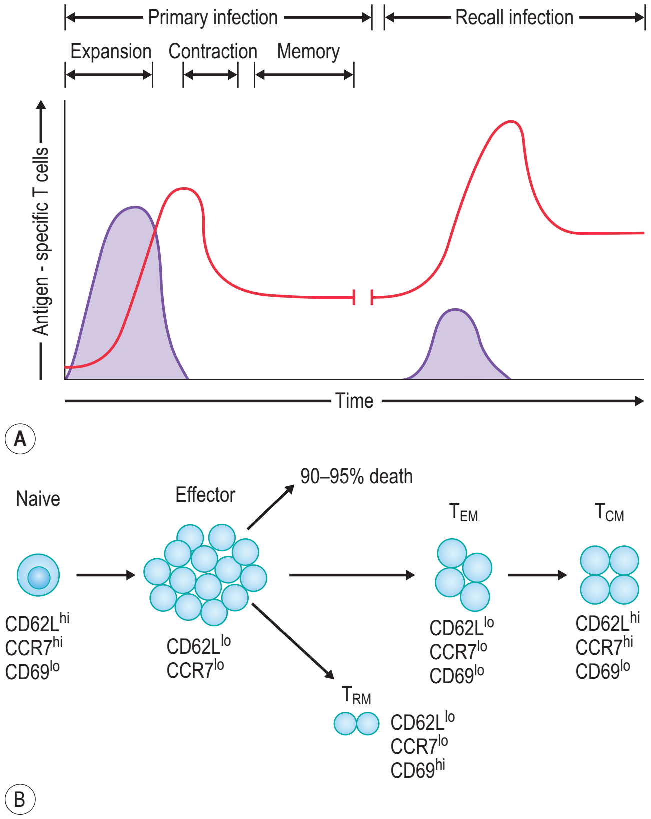

Primary infection
Recall infection
Expansion
Contraction
Memory
Antigen - specific T cells
Time
A
Naive
Effector
90–95% death
T_{EM}
T_{CM}
$CD62L^{hi}$ $CCR7^{hi}$ $CD69^{lo}$
$CD62L^{lo}$ $CCR7^{lo}$
$CD62L^{lo}$ $CCR7^{lo}$ $CD69^{lo}$
$CD62L^{hi}$ $CCR7^{hi}$ $CD69^{lo}$
T_{RM}
$CD62L^{lo}$ $CCR7^{lo}$ $CD69^{hi}$
B

FIG 25.2 Expansion/Contraction/Memory Phases of Adaptive Immunity and Memory Cell Subsets. (A) Dynamics of primary and secondary (recall) T-cell responses to viral infection. Both primary and recall T-cell responses undergo expansion and contraction phases, followed by stable immune memory. Recall responses induce a larger effector pool and reduced contraction further boosting the memory pool. (B) Effector and memory T-cell differentiation. Antigen stimulation expands effector cells, most of which die during the contraction phase. Effector memory T (T_{EM}) cells that are formed gradually convert to central memory T (T_{CM}) cells over time, with corresponding changes in surface marker expression. Some effector T cells develop into resident memory T (T_{RM}) cells that persist in the tissues and do not reenter the circulation.

Like B-cell responses, T-cell responses to viral infections also begin within lymphoid tissues. Specific $CD8^+$ CTL precursors recognize antigen in the context of MHC class I–peptide antigen complexes on DCs. The $CD8^+$ T cells become activated, proliferate, and differentiate into effectors. Expansion of these naïve antigen-specific precursors is considerable, often exceeding 10 000-fold, and results in an effector population that can account for 40% or more of a host's total $CD8^+$ T-cell population (Fig. 25.2). Various factors, including antigen and APCs, costimulatory molecules (*e.g.*, CD28 and 4–1BB), and inflammatory cytokines (*e.g.*, IFN-I and IL-12) are required to program the development of functional effector lymphocytes. In some infections, $CD4^+$ T-cell help is also important to prime robust CTL responses via signals, including CD40 that are delivered to DCs.[18] Activated CTL effectors then exit lymphoid organs and access almost all body locations via the bloodstream. However, effectors do not stay activated for long once the virus is cleared, and approximately 95% die by a process termed *activation-induced cell death.*

Following this contraction phase, the remaining cells differentiate into memory cells, which remain as a more or less stable population in the host for many years. They represent an expanded pool of CTL precursors that can be activated upon secondary encounter with antigen and provide enhanced protection upon reinfection with the same virus (see next section). Although much of our knowledge of T-cell responses to viruses has been obtained from murine studies, it is increasingly clear that the fundamental principles are the same or similar in humans.[19]

T-cell immunity against a particular virus involves both CD4$^+$ and CD8$^+$ T-cell subsets that recognize peptides derived from viral antigens bound to surface MHC proteins (class II and class I, respectively) (Chapters 5, 6). Complexes of viral peptides bound to MHC class II proteins are generated by APCs from scavenged and processed virus-infected cells or viral particles. Antigen–MHC class I complexes are expressed on the surface of infected cells, and antigen can also be transferred to APCs from infected cells by a process known as *cross-presentation.* Recent experiments in mice have also demonstrated a role for transfer of antigen between DCs as they migrate from infected tissues to the lymphoid tissues. Multiple subsets of DCs exist and specialize somewhat in antigen presentation on MHC-I or MHC-II.[20] During the process of activation, T cells can receive signals from multiple DC types in a temporally controlled sequence that coordinates CD4$^+$ and CD8$^+$ T-cell interactions.[18] Use of MHC class I and class II tetramers to directly visualize antigen-specific CD8$^+$ and CD4$^+$ T-cell responses, respectively, has demonstrated the significant size of T-cell responses to viruses, such that the majority of the activated T cells seen at the peak of the response are virus-specific.

CTLs function by recognizing virus-infected cells and killing them; this often involves perforins and cytotoxic granules containing granzymes. Effector CTLs can also induce death in target cells following engagement of the Fas ligand on the CTL with Fas on target cells. Both pathways lead to apoptosis of the target cell, involving the degradation of nucleic acids, including those of the virus. Alternatively, CD8$^+$ T cells also mediate defense through the release of various cytokines after antigen recognition. Some of the cytokines and chemokines most highly produced by CTLs include IFN-γ, TNF-α, lymphotoxin-α, and RANTES (CCL5) (Chapters 9, 10). These cytokines can have multiple antiviral effects on infected cells and on the cells around them, including purging of virus from infected cells without killing the cells. This is particularly important for such viruses as HSV, which infects nonrejuvenating cells, such as nerve cells.

CD4$^+$ T cells are involved in antiviral defense as well as being modulators of inflammatory reactions to viruses. Multiple functional subsets of CD4$^+$ T cells are recognized based largely on the types of cytokines produced when they recognize antigen. CD4$^+$ T cells are more broadly reactive than CD8$^+$ T cells; they recognize larger peptides processed from viral proteins and are restricted by MHC class II. These CD4$^+$ T cells participate in antiviral immunity in several ways. They can act as helper cells for the development of high-affinity antibody responses and for more functional CD8$^+$ T-cell responses.[16,21] Additionally, CD4$^+$ T cells act as effectors and orchestrate inflammatory reactions, which either serve a protective function or, in some cases, become prolonged causing chronic tissue damage (Chapter 16). The latter can happen in HCV-mediated hepatitis and HSV-mediated stromal keratitis. Occasionally, CD4$^+$ T cells can mediate direct cytotoxicity, but they are less effective than CD8$^+$ T cells. The principal subsets of CD4$^+$ T cells involved in inflammatory reactions are T helper-1 (Th1) cells (producing mainly IFN-γ, tumor necrosis factor [TNF]-α, IL-2) and Th17-producing cells (IL-17a and IL-22). A third effector subset, Th2 cells producing (IL-4, IL-5, and IL-13), also participates in inflammatory reactions, although in the case of viruses, these are usually more tissue damaging than protective. This situation can occur in responses to RSV infection. Regulatory T cells (Tregs) are a further subset of CD4$^+$ T cells of particular importance, since these cells largely act to regulate the function of effector subsets and, in so doing, influence the severity and duration of inflammatory reactions[22] (Chapter 18). Tregs produce antiinflammatory cytokines, such as IL-10 and TGF-β, and can be distinguished from other CD4$^+$ subsets by their expression of a unique transcription factor FoxP3. The balance of CD4$^+$ T-cell subset representation in response to a virus infection is critical. In situations where responses become overtly tissue damaging and chronic, the balance favors effector subsets. In such situations, changing the balance to favor Tregs can result in diminished lesions.

IMMUNOLOGICAL MEMORY

Immunological memory is a cardinal feature of adaptive immunity. The goal of vaccinology is to induce long-lived immunological memory to protect against reinfection (Chapter 90). Following infection with certain viruses, memory can be exceptionally long-lived, potentially for the life of the host (*e.g.,* yellow fever and smallpox viruses).[19,23] Memory is defined by the persistence of specific lymphocytes and antibody-producing plasma cells rather than that of antigen to induce continuous lymphocyte activation. Humoral memory to viruses involves long-lived plasma cells in bone marrow, which provide a continuous low-level source of serum antibody. This maintenance of humoral immunity also involves a population of homeostatically maintained memory B cells, which may be required to maintain stable numbers of long-lived plasma cells over time. The pool of memory T cells is regulated by low-level homeostatic division controlled by the cytokines IL-7 and IL-15. For memory CD8$^+$ T cells, IL-7 is primarily important for survival, whereas IL-15 is crucial for low-level proliferation to maintain the size of the memory T-cell pool.

KEY CONCEPTS

Principles of Antiviral Immunity

Many human viral infections are successfully controlled by the immune system.

Certain emerging viruses may overwhelm the immune system and cause severe morbidity and mortality.

Other viruses have developed mechanisms to overwhelm or evade the immune system and persist.

Individuals with defects in innate or adaptive immunity demonstrate more severe viral infections.

T-cell immunity is more important for control than are antibodies in many viral infections.

Antibodies are important to minimize reinfection, particularly at mucosal sites.

Immune memory is often sufficient to prevent secondary disease, although not in all viral infections.

Tissue-specific immune memory may be important to rapidly protect against reinfection at peripheral sites (*e.g.,* skin and mucosae).

Immunological memory is defined by a pool of antigen-specific cells whose increased frequency enables rapid control of viral reinfection (see Fig. 25.2). IL-7Rα-expressing effector T cells are the precursors of this memory pool. This population of cells, which constitutes about 5–10% of the effector pool, preferentially survives the contraction phase and gradually differentiates into a stable memory population.[24] Upon reinfection, these memory cells can be rapidly activated and, by virtue of their increased frequency, mediate more rapid clearance of the viral pathogen. Moreover, repeated stimulation of memory cells via multiple infections with the same virus, or prime-boost vaccine regimes, further increases the size of the antigen-specific memory T-cell pool.[25] Restimulation also affects the activation status and tissue distribution of memory T cells, which may enhance protection from viral infection in mucosal and other tissues.

Experiments in humans and mice have demonstrated that memory T cells are heterogeneous. Memory T cells were divided into effector memory (T_{EM}) and central memory (T_{CM}) subsets, defined by expression of two surface molecules involved in T-cell migration: CD62L and CCR7.[24] The $CD62L^{lo}CCR7^{lo}$ T_{EM} subset is found primarily in nonlymphoid tissues and the spleen, whereas the $CD62L^{hi}CCR7^{hi}$ T_{CM} subset is largely present in lymph nodes and the spleen. The current model predicts that effector T cells form the T_{EM} subset and that these cells gradually convert to a T_{CM} phenotype over time (Fig. 25.2B). Although the conditions that control the rate of this conversion are unknown, it is likely that the amounts of antigen and inflammatory signals received during the effector phase greatly influence this. It has also been shown that $CD4^+$ T-cell help is required for the generation of long-lived memory $CD8^+$ T cells, via interactions with DCs.[21]

Studies suggest that T_{CM} are capable of mounting stronger proliferative responses following reinfection. Tissue-specific homing of T_{EM} cells permits them to enter sites of potential viral infection, such as skin and mucosae. However, we now know that many memory T cells found at sites of previous viral infections take up long-term residence in tissues.[26] This includes skin, intestines, lungs, the liver, and the brain. These resident memory T cells (T_{RM} cells) are sequestered from the circulation and provide rapid protection against viruses, such as HSV, in skin, where they localize with a unique dendritic morphology and undergo slow surveillance of the tissue (Fig. 25.3). Notably, activation of T_{RM} cells can trigger enhanced early inflammation to drive local immunity. This is in contrast to T_{EM} cells, which continue to migrate through nonlymphoid tissues, rather than being sequestered in peripheral tissues, and also differs from the $CD8^+$ and $CD4^+$ T_{CM}, which migrate largely through lymphoid organs (spleen and lymph nodes). These differences may define the physiological *raison d'être* for these memory T-cell subsets, highlighting that measurement of memory T cells in human peripheral blood is a poor representation of the total-body memory T cell pool.

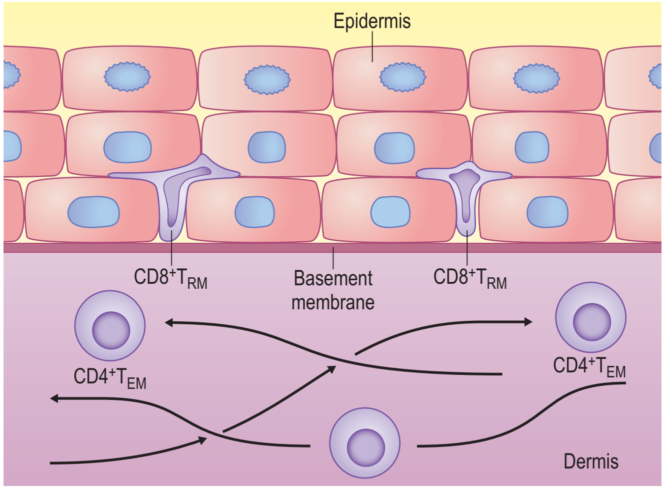


FIG 25.3 Unique Subsets of Memory $CD8^+$ and $CD4^+$ T Cells Reside Within Peripheral Tissues, at Sites of Previous Viral Infection, and Provide Rapid Protection Against Reinfection. Resident memory $CD8^+$ T cells (T_{RM}) remain localized in the epidermis in skin after herpes simplex virus (HSV) infection. Resident memory $CD4^+$ T (T_{EM}) cells continue to migrate through the dermal layers of skin, with access to blood and lymphoid tissues.

T_{RM} cells can be detected in tissues by using markers, such as CD69 and CD103, although these are imperfect identifiers, including in human tissues. T_{RM} cells in different anatomical locations share a common genetic signature and require common transcription factors for their formation. Yet, these cells also adopt unique gene expression that is imprinted by the tissue environment, and presumably imparts specialized functions on T_{RM} cells in each location. However, memory in certain peripheral tissues, such as lungs, appears to wane over time, suggesting that memory T cells may not persist in sufficient numbers in this site. This rationalizes a need for vaccines that induce optimal numbers of memory T cells in tissues as well as blood.

IMMUNE EVASION AND IMMUNITY TO CHRONIC VIRAL INFECTIONS

Many, if not all, viruses employ immune blunting or delay tactics to circumvent aspects of the immune system, allowing them time to replicate further or escape detection (Table 25.3).[27] One such mechanism may involve killing or infecting APCs. Viruses may also delay or prevent apoptosis induced by CTLs within infected cells. Other viral evasion measures aimed at the $CD8^+$ T cell–mediated antiviral defense system inhibit antigen processing, thereby minimizing effector CTL induction. To escape CTL killing, many viruses also downregulate the MHC molecules on the surface of infected cells. In addition, viruses may produce various mimics or modulators/inhibitors of cytokines, chemokines, or

TABLE 25.3 Mechanisms and Examples of Viral Immune Evasion

Mechanism	Example
Interference with viral antigen processing and presentation	HSV (ICP47), EBV (EBNA-1), HIV (Nef, Tat), HPV (E5), CMV (UL6)
Evasion of NK cell function	HIV (Nef), EBV (EBNA-1), CMV (UL40, UL18)
Inhibition of cell apoptosis	Adenovirus (RID complex and E1B), HIV (Nef), EBV (BHRF-1)
Destruction of T cells	HIV
Interference with antiviral cytokines and chemokines	EBV (IL-10 homologue), CMV(US28 chemokine receptor homologue), vaccinia virus (IL-18-binding protein), HIV (Tat chemokine activity)
Inhibition of complement action	HSV, pox viruses
Inhibition of DC maturation	HSV, vaccinia virus
Frequent antigenic variation	Influenza virus, HIV
Infection of immune privileged site	Measles virus, VZV and HSV (neurons)
Immune exhaustion	HIV, HCV, HBV

CMV, cytomegalovirus; DC, dendritic cell; EBV, Epstein–Barr virus; HBV, hepatitis B virus; HCV, hepatitis C virus; HIV, human immunodeficiency virus; HPV, human papillomavirus; HSV, herpes simplex virus; IL-18, interleukin-18; NK, natural killer; RID, receptor internalization and degradation; VZV, varicella-zoster virus.

other components of the immune system or their receptors. Viruses also resort to antigenic hypervariability to escape antibody or T-cell recognition. This can occur during transmission from host to host (*e.g.*, influenza virus), or within hosts during chronic infection through the generation of viral escape mutants. The latter is particularly important for HIV and HCV infections.

The success of many viral pathogens rests in their ability to subvert the host immune response. The most successful human viruses can escape the immune system and persist for the life of the host.[28] Two well-studied examples of this are CMV and EBV. T-cell responses to these viruses are prominent and readily detectable in humans, and yet the immune system is unable to clear either pathogen completely. However, these viruses generally remain undetectable in immunocompetent individuals. Other viral infections, such as those caused by the herpes viruses HSV and VZV, are marked by periods of latency when no virus can be detected. Yet, periods of viral reactivation, often triggered by stress, can lead to episodes of disease. These are controlled by the immune response, which plays a central role in controlling herpes virus latency.[29]

Many of the most medically important human viruses are associated with persistent viremia. These include those causing chronic infections, such as HIV, HCV, HBV, and human T-lymphotropic virus (HTLV), among others. Such chronic viral infections are marked by high levels of persisting antigen and can result in skewed T-cell immunodominance hierarchies, altered tissue localization of immune cells, and severely impaired T-cell function.[30] This altered T-cell function is hierarchical and results in functional T-cell defects ranging from reduced cytokine production and altered proliferative capacity (exhaustion) to death (deletion) of the responding T cells (Fig. 25.4).

Sustained viral antigen levels and inflammation are responsible for this immune dysfunction. This is in stark contrast to normal memory T-cell development, which occurs in the absence of persisting antigen (see previous section). Studies have demonstrated that signaling through multiple inhibitory receptors expressed on the cell surface contributes to exhaustion during chronic infections.[30] This includes the receptor programmed death (PD)-1, expression of which may be essential for preventing excessive immunopathology by effector T cells and yet appears to contribute directly to failed immunity to HIV infection and other chronic human viral infections. Although the molecular mechanisms of exhaustion remain unclear, differential involvement of transcription factors and altered gene expression define exhausted T cells. These studies implicated multiple inhibitory receptors as a potential therapeutic targets, and although combinations of these checkpoint inhibitor blockade therapies are proving highly beneficial to the treatment of certain cancers,[31] similarly efficacious responses have yet to be demonstrated during chronic virus infection.

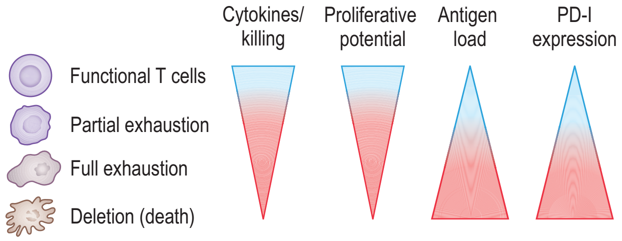


FIG 25.4 Hierarchical Model of T-Cell Exhaustion During Persistent Viral Infection. T-cell function (cytokine production, killing, and proliferative potential) is negatively influenced by increasing levels of antigen. Low levels of persistent antigen may lead to partial loss of function and intermediate levels of programmed death (PD)-1 expression. High, sustained levels of antigen over time can lead to full loss of function, high levels of PD-1, and eventually cell death (deletion).

OUTCOMES OF VIRUS INFECTION: IMMUNITY OR IMMUNOPATHOLOGY

Typically, individual humans respond to a virus infection in different ways. When the common cold or even pandemic influenza infection occurs, only a small percentage of exposed persons may develop overt clinical disease. In the prevaccine days, poliomyelitis was a much-feared consequence of poliovirus infection, but only a very small percentage of infected persons developed the paralyzing complications. Similarly, only an unfortunate few develop life-threatening meningoencephalitis following infection with the insect-transmitted West Nile virus. It is particularly characteristic of chronic viral infections that clinical expression is highly variable. With HCV, for example, in 70–80% of patients, some form of chronic liver disease develops, and the virus is not cleared. However, in up to 30%, the infection is controlled, the virus is cleared, and immunity to reinfection develops. The latter group of individuals make a type of immune response that includes protective antibodies along with an appropriate pattern of T-cell responsiveness.[32]

We do not fully understand the reasons for the varying outcomes of virus infections in different persons, and almost certainly multiple factors are involved. Many of these factors impact the response pattern made by the innate immune system, which, in turn, affects the magnitude and type of adaptive immune response that occurs. Some of the circumstances that do influence the outcome of infection include genetic susceptibility of the host, the age of the host when infected, the dose and route of infection, the variable induction in the host of antiinflammatory cells and proteins, and the presence of concurrent infections and past exposure to cross-reactive antigens.[32]

IMMUNOPATHOLOGY AND AUTOIMMUNITY

Immune responses against virus-infected cells often result in tissue damage, especially if cell killing is involved or if there is extensive recruitment and activation of inflammatory cell types, such as macrophages and sometimes neutrophils. If the response is brief and is quickly repaired, it is usually deemed an immunoprotective event. A prolonged tissue-damaging effect resulting from an immune reaction against viruses is considered immunopathology. Such situations most commonly involve persistent viruses, which are themselves often mildly cytodestructive in the absence of an immune reaction. Chronic tissue damage initiated by viruses can also result in development of an autoreactive and an occasionally oncogenic response. For example, some autoimmune diseases may be initiated or exacerbated by viral infections, but no named virus has been regularly incriminated as a cause of human autoimmune disease.[33] Circumstantial evidence exists for a virus link in multiple sclerosis (MS), insulin-dependent diabetes, and possibly systemic lupus erythematosus (SLE). In MS, many viruses have been isolated from patients, although no specific one has been tied to the disease etiology. The current hypothesis is that viral infections set up an

TABLE 25.4 Lesions Resulting From Immunopathology

Primarily involving $CD8^+$ T cells acting as cytotoxic T lymphocytes or sources of proinflammatory cytokines	Murine lymphocytic choriomeningitis virus; Hepatitis B virus (HBV)–induced chronic hepatitis Coxsackie B virus–induced diabetes Coxsackie B virus–induced myocarditis Demyelination caused by some strains of mouse coronavirus and Theiler virus
Primarily involving $CD4^+$ T cells that produce Th1 cytokines	Demyelination caused by some strains of mouse coronavirus and Theiler's virus; Herpes simplex virus (HSV)–induced stromal keratitis
Involvement of $CD4^+$ T cells that produce Th2 cytokines	Respiratory syncytial virus (RSV)–induced pulmonary lesions
Involvement of antibody	Glomerulonephritis in chronic hepatitis B Dengue hemorrhagic fever

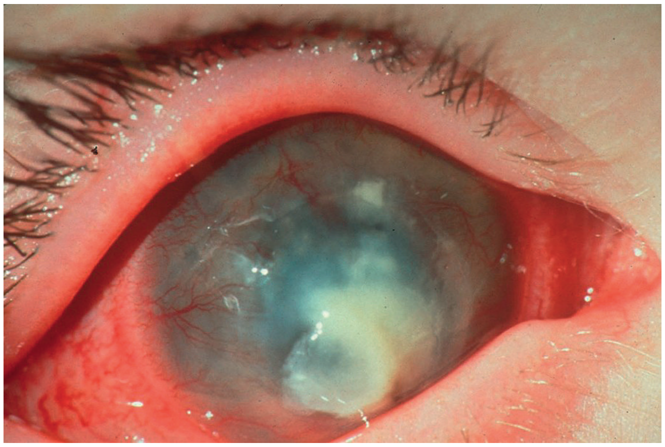


FIG 25.5 **Example of Herpetic Stromal Keratitis (Hsk) in the Human Eye After Herpes Simplex Virus-1 (HSV-1) Infection.** Inflammation of the eye and eyelid can be observed, as well as neovascularization and substantial necrosis, ulceration, and opacity of the cornea.

inflammatory environment that may exacerbate or tip the balance toward disease in genetically susceptible individuals.

Immunopathological reactions involving viruses have several mechanisms, but T cells are usually involved as orchestrators of inflammatory events (Table 25.4). The clearest example of immunopathology involving a virus is lymphocytic choriomeningitis virus (LCMV) in the mouse. This model has dominated ideas and has set several paradigms in viral immunology in general. The first virus-induced immunopathological lesions recognized were glomerulonephritis and arteritis, noted in mice persistently infected with LCMV. The lesions were assumed to represent inflammatory reactions to tissue-entrapped immune complexes that activate complement. Similar immune complex–mediated lesions occur in other infections, including lung lesions found in severe influenza, respiratory syncytial virus infection, viral hepatitis, and arthritis. However, only rarely have viral antigens been shown to contribute to the antigen component of the complex. An example where the inclusion of viral antigen in immune complexes has been demonstrated is chronic HBV infection of humans. Autoimmune diseases, such as SLE, also result from immune complex–mediated tissue damage. However, evidence linking viruses to the etiology or pathogenesis of SLE is scarce, since the immune complexes in SLE do not appear to include viral antigens at any stage.

Thanks largely to the LCMV model, it is clear that $CD8^+$ T-cell recognition of viral antigens can result in tissue damage. In LCMV infection, damage occurs in the leptomeninges of immunocompetent mice infected intracerebrally. Hepatitis can also occur in mice infected intravenously. Neither lesion becomes evident if the $CD8^+$ T-cell response is suppressed. $CD8^+$ T cell–mediated immunopathology can be a causative mechanism of chronic hepatitis associated with HCV and HBV infection, although the tissue damage also involves inflammatory $CD4^+$ T cells. Additional viral immunopathology models where lesions result primarily from $CD8^+$ T-cell involvement include myocarditis and insulin-dependent diabetes associated with coxsackie B virus infection. In both instances, $CD8^+$ T cells mainly orchestrate events, but tissue damage may result from the bystander effects of cytokines and other molecules, such as lipid mediators, metalloproteinases, and components of the oxygen burst. Although coxsackie virus can be a cause of diabetes in the mouse, attempts to relate viral infection directly to the etiology of human diabetes have so far failed.

CLINICAL RELEVANCE

Hypothesized Role of Viruses in Autoimmunity

Molecular mimicry: similar epitopes shared by virus and host
Bystander activation: chronic release of cytokines and host antigens activates local autoreactive lymphocytes
Viral persistence: chronic viral antigen presentation on host cells leads to prolonged immunopathology

Immunopathological reactions against viruses can also involve subsets of $CD4^+$ T cells, which can be either Th1 or Th17 or both. One well-studied example involves persistent infection with Theiler virus in mice.[34] This infection causes a demyelinating syndrome that resembles the autoimmune disease experimental allergic encephalomyelitis. In both situations, $CD4^+$ T cells that produce Th1 cytokines appear to serve as pathological mediators. Furthermore, in both models an increase in the involvement of myelin-derived autoantigens occurs as the disease progresses. Once again, such observations indicate the possible role of a virus in an autoimmune disease. With the Theiler virus model, the virus persists in the nervous system and chronically stimulates $CD4^+$ T cells to secrete an array of cytokines. The demyelinating events appear to result from cytokine action on oligodendrocytes. Myelin components, such as myelin basic protein, proteolipid protein, and myelin oligodendroglial glycoprotein, may be released and can participate as additional antigen in immunoinflammatory events. This scenario is referred to as *epitope spreading.*

Another model of virus-induced immunopathology that mainly involves the Th1 subset of $CD4^+$ T cells is stromal keratitis caused by HSV infection (Fig. 25.5).[35] The pathogenesis of this immunopathological lesion is unusual in that it occurs and progresses when viral antigens can no longer be demonstrated. The chronic immunoinflammatory lesions are mainly orchestrated by $CD4^+$ T cells, but multiple early events induce the subsequent pathology. Viral replication, the production of certain cytokines and chemokines (IL-1, IL-6, IL-12, and CXCL8), recruitment of inflammatory cells (*e.g.,* neutrophils), and neovascularization of the avascular cornea all precede immunopathology. Recently, it has become evident that Th17 T cells participate in stromal keratitis lesions. The role of Th17 T cells as orchestrators of inflammatory reactions has been a major research focus, especially in lesions of autoimmune diseases.[36] When Th17 T cells are the principal mediators of tissue damage, abundant neutrophils are recruited to inflammatory sites, with such cells being mainly responsible for tissue damage.

A further mechanism of viral-induced immunopathology and autoimmunity is molecular mimicry.[33] Molecular mimicry represents shared antigenic epitopes, either B- or T-cell antigen, between the host and virus. This concept originated with streptococci and their association with rheumatic fever. With human autoimmune disease, there is little direct support for viral molecular mimicry; however, some animal models have been used to prove the theoretical case, where a viral antigen is expressed as a self-protein in the islet cells of the pancreas. In this model, subsequent infection with the virus induces diabetes. However, this is not true mimicry and may be more closely related to viral antigen persistence in a model such as Theiler disease.

As discussed previously, the outcome of a T-cell response to a virus may be critically dependent on the balance of the T-cell-type response thus, tissue damage is likely to be more severe and prolonged if $CD8^+$ or Th1 and Th17 $CD4^+$ T cells are predominant. Lesions become milder and may resolve when the balance favors Tregs. Accordingly, therapeutic approaches that can shift the balance of T cells are under trial.

KEY CONCEPTS

Phases of Immunity Affected by Regulatory T Cells (Tregs)

Interference with antigen presentation by dendritic cells
Inhibition of T-cell proliferation
Inhibition of molecules involved in tissue-specific migration of effector cells
Inhibition of T-cell effector functions in lymphoid and nonlymphoid tissues

TRANSLATIONAL RESEARCH OPPORTUNITIES

Reversing T-cell exhaustion in patients suffering from chronic infections or cancer will be a key clinical target in the near future. The discovery of multiple inhibitory receptors on exhausted T cells (*e.g.,* PD-1, LAG-3, 2B4, TIM-3) has provided the opportunity to selectively improve T-cell function through blockade of these inhibitory receptors. This may be combined with blockade of immunosuppressive cytokines (*e.g.,* IL-10) or enhancement of signals stimulatory to the response (*e.g.,* IL-7 therapy), as well as with more traditional antiviral therapies and vaccination. The challenge that lies ahead will be in determining which combination of inhibitory and stimulatory signals will need to be manipulated in different diseases and in different groups of patients.

The design of a new generation of vaccines to target diseases, such as HIV and influenza, may require tailor-made solutions for patients who respond poorly to vaccination or respond improperly, as with adverse effects, such as autoimmune reactions. High-throughput approaches now allow for the generation of a molecular signature of vaccination or infection.[37] Such systems' biology approaches are expected to result in novel screening for immune protection parameters after vaccination. In the near future, this should also assist in the formulation of new vaccines containing key immune activators, such as those that stimulate certain subsets of T cells or induce appropriate homing molecule expression on these cells to direct them to tissues where they are required to mediate protection (*e.g.,* mucosal sites, or skin).

ON THE HORIZON

Pressing Issues in Need of Solutions

Design of new vaccines that induce broadly neutralizing antibodies
Design of new vaccines that induce tissue-resident and circulating memory T-cell subsets
Overcoming immune dysfunction during chronic viral infections for successful viral clearance
Improving the efficacy of vaccines to viruses using systems biology approaches
Therapies for reducing immunopathology during viral infections

In some individuals, viral infections cause mild, or sometimes debilitating, tissue damage. Factors that influence whether a viral infection results in immunopathology varies from individual to individual. These factors include age, the route of infection, preexisting immunity, host genetics, and the host's viral burden or virome. Our knowledge of the influence of these factors on the outcome of viral infection is expected to improve rapidly in the coming decades. Recent advances have shed considerable light on the various proinflammatory and antiinflammatory mediators produced during viral infections. These represent key targets for novel therapies in the near future via the use of small-molecule inhibitors.

CONCLUSIONS

Humans are infected by many pathogenic viruses. In most cases, these infections are controlled by the immune system with limited damage to the host. However, certain viruses, particularly in cases where the host's immune system is impaired, can cause significant damage to the host's tissues. As our understanding of the mechanisms underlying innate immune defenses, antigen presentation, T- and B-cell responses, and Tregs continues to improve, so too does the ability to design better vaccines and therapies to boost the immune control of viral infections. Although this remains a challenging goal, particularly for many human viruses, such as HIV, HCV, and HSV, these rapid advances continue to provide many avenues for further investigation.

ACKNOWLEDGMENTS

Barry T. Rouse is supported by grants from the National Institutes of Health and Scott N. Mueller by the Australian Research Council and the Australian National Health and Medical Research Council.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Marsh M, Helenius A. Virus entry: open sesame. Cell 2006;124:729–40.
2. Bachmann MF, Jennings GT. Vaccine delivery: a matter of size, geometry, kinetics and molecular patterns. Nat Rev Immunol 2010;10:787–96.

3. Iwasaki A. A virological view of innate immune recognition. Annu Rev Microbiol 2012;66:177–96.
4. Kanneganti TD. Central roles of NLRs and inflammasomes in viral infection. Nat Rev Immunol 2010;10:688–98.
5. Garcia-Sastre A, Biron CA. Type 1 interferons and the virus-host relationship: a lesson in detente. Science 2006;312:879–82.
6. Crouse J, Kalinke U, Oxenius A. Regulation of antiviral T cell responses by type I interferons. Nat Rev Immunol 2015;15:231–42.
7. Jost S, Altfeld M. Control of human viral infections by natural killer cells. Annu Rev Immunol 2013;31:163–94.
8. O'Sullivan TE, Sun JC, Lanier LL. Natural killer cell memory. Immunity 2015;43:634–45.
9. Godfrey DI, Uldrich AP, McCluskey J, et al. The burgeoning family of unconventional T cells. Nat Immunol 2015;16:1114–23.
10. Bottazzi B, Doni A, Garlanda C, et al. An integrated view of humoral innate immunity: pentraxins as a paradigm. Annu Rev Immunol 2010;28: 157–83.
11. Garlanda C, Dinarello CA, Mantovani A. The interleukin-1 family: back to the future. Immunity 2013;39:1003–18.
12. Corti D, Lanzavecchia A. Broadly neutralizing antiviral antibodies. Annu Rev Immunol 2013;31:705–42.
13. Iwasaki A. Antiviral immune responses in the genital tract: clues for vaccines. Nat Rev Immunol 2010;10:699–711.
14. Qi H, Kastenmuller W, Germain RN. Spatiotemporal basis of innate and adaptive immunity in secondary lymphoid tissue. Annu Rev Cell Dev Biol 2014;30:141–67.
15. Mueller SN, Germain RN. Stromal cell contributions to the homeostasis and functionality of the immune system. Nat Rev Immunol 2009;9: 618–29.
16. Crotty S. Follicular helper CD4 T cells (T_{FH}). Annu Rev Immunol 2011;29:621–63.
17. Corcoran LM, Tarlinton DM. Regulation of germinal center responses, memory B cells and plasma cell formation-an update. Curr Opin Immunol 2016;39:59–67.
18. Bedoui S, Heath WR, Mueller SN. $CD4^+$ T cell help amplifies innate signals for primary $CD8^+$ T cell immunity. Immunol Rev 2016;272: 52–64.
19. Ahmed R, Akondy RS. Insights into human $CD8^+$ T-cell memory using the yellow fever and smallpox vaccines. Immunol Cell Biol 2011;89: 340–5.
20. Merad M, Sathe P, Helft J, et al. The dendritic cell lineage: ontogeny and function of dendritic cells and their subsets in the steady state and the inflamed setting. Annu Rev Immunol 2013;31:563–604.
21. Laidlaw BJ, Craft JE, Kaech SM. The multifaceted role of $CD4^+$ T cells in $CD8^+$ T cell memory. Nat Rev Immunol 2016;16:102–11.
22. Veiga-Parga T, Sehrawat S, Rouse BT. Role of regulatory T cells during virus infection. Immunol Rev 2013;255:182–96.
23. Amanna IJ, Slifka MK, Crotty S. Immunity and immunological memory following smallpox vaccination. Immunol Rev 2006;211:320–37.
24. Mueller SN, Gebhardt T, Carbone FR, et al. Memory T cell subsets, migration patterns, and tissue residence. Annu Rev Immunol 2013;31: 137–61.
25. Jameson SC, Masopust D. Diversity in T cell memory: an embarrassment of riches. Immunity 2009;31:859–71.
26. Mueller SN, Mackay LK. Tissue-resident memory T cells: local specialists in immune defence. Nat Rev Immunol 2016;16:79–89.
27. Finlay BB, McFadden G. Anti-immunology: evasion of the host immune system by bacterial and viral pathogens. Cell 2006;124:767–82.
28. Virgin HW, Wherry EJ, Ahmed R. Redefining chronic viral infection. Cell 2009;138:30–50.
29. Rouse BT, Kaistha SD. A tale of 2 alpha-herpesviruses: lessons for vaccinologists. Clin Infect Dis 2006;42:810–17.
30. Wherry EJ, Kurachi M. Molecular and cellular insights into T cell exhaustion. Nat Rev Immunol 2015;15:486–99.
31. Baumeister SH, Freeman GJ, Dranoff G, et al. Coinhibitory pathways in immunotherapy for cancer. Annu Rev Immunol 2016;34:539–73.
32. Rouse BT, Sehrawat S. Immunity and immunopathology to viruses: what decides the outcome? Nat Rev Immunol 2010;10:514–26.
33. Fujinami RS, von Herrath MG, Christen U, et al. Molecular mimicry, bystander activation, or viral persistence: infections and autoimmune disease. Clin Microbiol Rev 2006;19:80–94.
34. Olson JK, Ercolini AM, Miller SD. A virus-induced molecular mimicry model of multiple sclerosis. Curr Top Microbiol Immunol 2005;296: 39–53.
35. Biswas PS, Rouse BT. Early events in HSV keratitis—setting the stage for a blinding disease. Microbes Infect 2005;7:799–810.
36. Korn T, Bettelli E, Oukka M, et al. IL-17 and Th17 Cells. Annu Rev Immunol 2009;27:485–517.
37. Pulendran B. Systems vaccinology: probing humanity's diverse immune systems with vaccines. Proc Natl Acad Sci USA 2014;111:12300–6.

MULTIPLE-CHOICE QUESTIONS

1. Many of the most medically important human viruses are associated with persistent viremia, including human immunodeficiency virus (HIV) and hepatitis C virus (HCV). What are the effects of high levels of persisting antigen on the immune response?
   - **A.** Increased killing capacity by cytotoxic T cells
   - **B.** Altered tissue distribution and impaired functions of T cells
   - **C.** Upregulation of T-cell proliferation
   - **D.** Enhanced cytokine production by virus-specific T cells
   - **E.** Reduced expression of coinhibitory molecules by the responding virus-specific T cells

2. Many viral infections can cause immunopathological reactions, such as herpetic stromal keratitis induced by herpes simplex virus. What is a major parameter in the pathogenesis of this disease?
   - **A.** Inhibition of inflammatory cell recruitment into the infected tissues
   - **B.** Recruitment of cytotoxic CD4 T cells that cause ocular pathology
   - **C.** Recruitment and retention of pathogenic CD8 T cells in the late stage of the response
   - **D.** Recruitment and retention of pathogenic Th1 CD4 T cells into the ocular tissues
   - **E.** Reduction in Th17 CD4 T cells in the ocular tissues

3. Immunological memory is a cardinal feature of adaptive immunity to virus infection. Memory T cells can be divided into multiple subsets. Which of the following statements accurately describe the major subsets of memory T cells?
   - **A.** Effector memory T cells migrate through lymphoid tissues, central memory T cells migrate through nonlymphoid tissues, and tissue-resident memory T cells circulate in blood.
   - **B.** Effector memory T cells migrate through nonlymphoid tissues, central memory T cells persist in nonlymphoid tissues, and tissue-resident memory T cells persist predominantly in lymphoid tissues.
   - **C.** Effector memory T cells persist in nonlymphoid tissues and do not enter blood, central memory T cells persist in lymphoid tissues, and tissue-resident memory T cells migrate predominantly through nonlymphoid tissues.
   - **D.** Effector memory T cells migrate exclusively in blood, central memory T cells migrate through lymphoid tissues, and tissue-resident memory T cells are retained in nonlymphoid tissues.
   - **E.** Effector memory T cells migrate through nonlymphoid tissues, central memory T cells migrate through lymphoid tissues, and tissue-resident memory T cells are retained in tissues.

26

Host Defenses to Intracellular Bacteria

Stephen T. Reece, Stefan H.E. Kaufmann

The evolutionary relationship between humans and bacteria is so intimate that it is impossible to imagine the development of one without the other.[1] Although this coexistence is generally mutually beneficial, clear boundaries do exist between the two and are intensely defended. We tend to think of the human host as the defender and bacteria as transgressors of these boundaries. Evolution of human immunity has been accompanied by evolution of ingenious bacterial mechanisms to not only survive its onslaught but also to manipulate it to enhance survival. This idea is instructively mirrored in the lifestyle of intracellular bacteria. These bacteria actively seek out an environment inside human cells where they can flourish; this is not an easy environment in which to survive. Human cells have developed an ability to differentiate bacterial from host components and direct host cells to clear the invader. The most successful intracellular pathogens have adapted to the intracellular environment of a particular cell target, proliferate only slowly, and can live for long periods completely undetected by the immune system, as we see in the case of the tuberculosis (TB) bacterium *Mycobacterium tuberculosis.* In other instances, for example, in listeriosis, intracellular infection is more explosive, with the rich intracellular environment harnessed to rapidly amplify bacterial growth. In some cases, intracellular bacteria live for a very long time in the human body, sometimes for a person's entire lifetime. A wide spectrum of pathologies ensues from intracellular infection, making most intracellular bacteria highly clinically relevant. Moreover, new concepts on the influence of intracellular bacteria on host cell differentiation point to their ability to change infected cell phenotype to enhance survival.

This chapter evaluates the current interpretation of this fascinating interplay between human and microbe, sheds light on how the human immune system functions, and how cellular phenotype can be molded in cells whose fates were previously believed to be strictly predetermined. Finally, such insights can inform new therapeutic and prophylactic approaches to keep intracellular bacterial infections under control.

BALANCE OF PROTECTION AND PATHOLOGY DEFINES THE CHRONIC NATURE OF INTRACELLULAR BACTERIAL INFECTION

Some bacteria, such as *Listeria monocytogenes,* are fully eradicated once the host immune response has reached its peak activity. Most often, the intracellular habitat provides a protective niche that promotes persistent infection in the face of an ongoing immune response. Here, the bacteria can persist for long periods without causing clinical signs of illness, but bacterial growth can be reactivated to cause disease if the immune response becomes compromised. This occurs in *M. tuberculosis* infection, resulting in disease years or decades after primary infection. In fact, disease need not arise from infection at all. In many regions, for example, the majority of adults harbor *M. tuberculosis* without suffering from clinical disease. However, disease can develop directly after primary infection, during maturation of the immune response, or with regression once the immune response is sufficiently strong. Yet, sterile eradication of the pathogen is rarely achieved: bacteria persist latently, and illness may reemerge at a later time. For example, *Rickettsia prowazekii* may persist for decades after convalescence from typhus to cause Brill-Zinsser disease later.

CLINICAL PEARLS

Distinguishing Clinical Characteristics of Infections With Intracellular Bacteria

Nonsterilizing immunity
Persistent bacteria, sometimes latent infection
Formation of long-lasting tissue granulomas containing low numbers of viable bacteria
Critical role of T cells in protection, role of antibodies less well established but likely to play an as-yet unappreciated role
Critical role of immune response in pathology
Lack of effective vaccines
Host-directed therapies toward enhancing antimicrobial mechanisms while limiting host pathology

Several intracellular bacteria possess components that can profoundly influence the course of disease, for example, the lipopolysaccharides (LPSs) of brucellae and salmonellae. Chronic persistence inside host cells, however, depends on the target cell remaining intact and physiologically active. Accordingly, many intracellular bacteria are of low toxicity and do not have dramatic direct effects on their host. Instead, pathogenesis is largely determined by the immune response. Classic examples of this concept include granuloma liquefaction in acute TB, which severely affects lung function, and eye scarring as a consequence of chronic or recurring *Chlamydia trachomatis* infection that ultimately leads to trachoma.

The survival of intracellular bacteria has major consequences for pathology. Although many intracellular bacteria show some organ tropism, dissemination to other organs frequently occurs, resulting in different disease forms. For example, TB is generally manifested in the lung in 80% of cases, yet many other organs can be affected. In contrast to other *Salmonella enterica* serovars,

TABLE 26.1 Major Infectious Diseases Caused by Intracellular Bacteria

Disease	Pathogen	Prevalence	Incubation Time	Route of Infection	Target Cell
Granulomatous Intracellular Bacteria					
Tuberculosis	*Mycobacterium tuberculosis*	Worldwide	Years (latency after primary infection and disease reactivation) Weeks (miliary TB)	Inhalation of bacteria-containing microdroplets	Macrophage
Leprosy	*Mycobacterium leprae*	South America Africa India Southeast Asia	Years	Smear infection through mucosa/ inhalation	Macrophage Schwann cell
Typhoid fever	*Salmonella enterica* serovars Typhi and Paratyphi	Worldwide	7–10 days	Fecal–oral	Macrophage
Brucellosis	*Brucella* spp.	Worldwide	Weeks to months	Zoonosis; cows, goats, pigs; inhalation, gut, skin abrasion	Macrophage
Listeriosis	*Listeria monocytogenes*	Worldwide	Days to months	Fecal–oral	Macrophage Hepatocyte
Nongranulomatous Intracellular Bacteria					
Legionnaires' disease	*Legionella pneumophila*	Worldwide	2–10 days	Inhalation	Macrophage
Rocky Mountain spotted fever	*Rickettsia rickettsiae*	Western hemisphere	1 week	Tick bite	Vascular endothelial cell Smooth muscle cell
Urogenital infection	*Chlamydia trachomatis* serovars D-K	Worldwide	1–3 weeks	Sexual intercourse	Epithelial cell
Conjunctivitis, trachoma	*Chlamydia trachomatis* serovars A-C	Africa	Conjunctivitis: 1–3 weeks Trachoma: years	Eye	Epithelial cell
Cat scratch disease	*Bartonella henselae* *B. quintana* *B. bacilliformis*	Worldwide	Bacillary angiomatosis Peliosis hepatitis Endocarditis Bacteremia with fever Neuroretinitis: 1–3 weeks	Flea, sandfly, or mosquito bite; animal scratch or bite	Erythrocyte Endothelial cell

the serovars Typhi and Paratyphi are not restricted to the gastrointestinal (GI) tract but are disseminated to internal organs, primarily the liver and spleen. In these cases, the type of clinical disease depends markedly on the infected tissue type.

KEY CONCEPTS

Characteristic Features of Intracellular Bacterial Infections

Persistence of bacteria inside mononuclear phagocytes (*i.e.,* macrophages)
Low to absent bacterial-mediated toxicity to the host
Protection requires cytokine-mediated activation of infected phagocytes
Interferon-γ (IFN-γ) and tumor necrosis factor-α (TNF-α) produced by antigen-specific T cells are key cytokines for protection

INTRACELLULAR BACTERIAL INFECTIONS OF CLINICAL RELEVANCE (Table 26.1)

Granulomatous Infections

Tuberculosis

The major entry of tubercle bacilli into the human body is via inhalation into the lung. These inhaled bacteria are then engulfed by alveolar macrophages, which transport the pathogens to the lung interstitia. Their exact fate after these events is enigmatic. Moreover, most infections in humans result in an asymptomatic carrier state or the so-called latent TB infection (LTBI). Infection with *M. tuberculosis* starts with the so-called Ghon complex, characterized by a caseous lesion in the midlung as well as in the draining lymph nodes. These primary lesions can progress, but their development rarely causes disease directly.[2] Moreover, bacteria from these sites can disseminate to other regions of the lung and systemically, causing disease of the kidneys, liver, and central nervous system (CNS). Containment of the primary lesions, which leads to LTBI, is a function of an effective, predominately cellular antitubercular immune response. Infection of immunocompromised patients, notably those with acquired immunodeficiency syndrome (AIDS), or newborns frequently results in systemic disease (miliary TB). TB represents a major health problem worldwide, including an increasing incidence in many industrialized countries. In 2015, the World Health Organization (WHO) estimated that 10.4 million active TB cases were diagnosed worldwide and close to 1.8 million people died of the disease.[3] The much larger estimated number of 2 billion individuals infected with *M. tuberculosis* well illustrates the dissociation of infection from disease. The emergence of multidrug-resistant strains and extremely drug-resistant strains has complicated treatment with currently available antibiotic therapy, and even when treatment is successful, recurrence of disease can occur. The currently available live whole cell vaccine BCG (Bacille Calmette–Guérin), an attenuated strain derived from the etiological agent of bovine TB *Mycobacterium bovis,* shows only low and variable protection against pulmonary TB.

Leprosy

Mycobacterium leprae is most likely transmitted by contact with patients who shed microorganisms in nasal secretions and lesion exudates. It primarily affects the nerves and the skin, frequently leading to stigmatizing deformities. In skin, bacilli target keratinocytes, histiocytes, and macrophages, whereas in peripheral nerves, Schwann cells are the major target for entry. Leprosy is a spectral disease. The tuberculoid pole is characterized by rigorous T-cell responses, which succeed in restricting microbial growth in well-defined lesions containing few bacilli. In contrast, at the lepromatous pole, bacterial growth is unrestricted and lesions contain abundant bacilli within macrophages lacking signs of activation. Several types of immunosuppression have been implicated in this latter type of disease. Infection of Schwann cells promotes nerve damage and anesthesia. This results in injuries and secondary infections that significantly exaggerate the disease. Despite the success of multidrug therapy in reducing the number of registered leprosy cases worldwide, some 216 000 new cases were reported in 2013. This suggests that active transmission of *M. leprae* is still occurring and that more effective interventions are required to prevent it.[4]

Atypical Mycobacterial Infections

Mycobacterial species present in the environment are typically unable to persist within activated macrophages and thus rarely cause disease in individuals with competent immune status.[5]

As a consequence of human immunodeficiency virus (HIV) infection, however, nontuberculous mycobacteria (NTM), primarily *Mycobacterium avium/Mycobacterium intracellulare,* have gained clinical importance, and these infections are recognized as one of the most common complications of AIDS in industrialized nations.

Mycobacterium scrofulaceum occasionally causes lymphadenitis in children, and *Mycobacterium kansasii* primarily causes infections in older men with preexisting lung disease.

Mycobacterium ulcerans causes a severe subcutaneous infection characterized by chronic skin ulcerations, known as *Buruli ulcer.* This pathology is caused—at least in part—by elaboration of a mycolactone toxin by the bacillus that exhibits highly cytopathic effects.[6] Buruli ulcer is most predominant in West African countries that accounted for most of the 2251 cases reported globally in 2014.[7]

Typhoid or Enteric Fever

Salmonella enterica serovars Typhi and Paratyphi A, Paratyphi B, and Paratyphi C are leading causes of community-acquired bloodstream infections in low- and middle-income countries. The route of transmission is fecal–oral and largely occurs via contaminated water sources. Bacteria are disseminated within mononuclear phagocytes (MPs) from the GI tract to macrophage-rich organs, particularly the liver, spleen, and lymph nodes. Accordingly, typhoid is characterized by systemic symptoms, such as prolonged fever and malaise, with sustained bacteremia, although diarrhea or constipation may also be present. In some cases, an asymptomatic carrier state can persist as a result of chronic infection of the gallbladder, which maintains the environmental reservoir of infection in endemic areas. Typhoid fever remains a major cause of morbidity and mortality, with approximately 21 million new cases and over 190 000 deaths per year worldwide.[8]

Gastroenteritis

S. enterica serovars Typhimurium and Enteritidis, often referred to as nontyphoidal salmonella (NTS), are the major causes of salmonella gastroenteritis in humans, which occurs mainly as a result of the ingestion of contaminated food or water. The bacteria rapidly cross the intestinal epithelia and replicate in the lamina propria, inducing an influx of polymorphonuclear neutrophils (PMNs), which is generally sufficient to resolve the infection within 7 days. In rare cases, bacteria enter the bloodstream and cause systemic bacteremia, most notably in patients with AIDS, where death can occur as a result of septic shock.[8]

Listeriosis

Listeriosis monocytogenes is increasingly recognized to cause foodborne gastroenteritis. Clinical listeriosis affects mainly pregnant women, older adults, fetuses, and neonates. Disease manifestations are most severe in patients with a compromised immune system, in whom the CNS becomes involved and fatal bacteremia can result. Additionally, as these bacteria are able to cross the placenta, listeriosis is a major cause of perinatal and neonatal disease, typically resulting in abortion. Listeria outbreaks are sporadic with low incidence but high fatality and affect high-income countries, such as the United States.

Brucellosis

Brucellosis is the most common global zoonosis of humans with approximately 500 000 cases per year.[9] It is caused by *Brucella abortus, B. melitensis,* or *B. suis,* which primarily infect cows, goats, and pigs, respectively. The bacteria are transmitted to humans via inhalation, through abraded skin or the GI tract. Lesions are primarily found within macrophage-rich tissues, especially the spleen and bone marrow. Human brucellosis is characterized by systemic symptoms, particularly undulant fever. Although the disease often remains subclinical, it becomes chronic in some patients, and relapses and remissions may occur. Interest in brucellosis has increased in the last 5 years because of elevated levels of detection resulting from better surveillance.

Lymphogranuloma Venereum

Lymphogranuloma venereum (LGV), a sexually transmitted disease, is highly prevalent in Africa, Southeast Asia, and Latin America. LGV has recently emerged as infection of sexually active homosexual men in Europe and the United States. It is caused by the L1, L2, and L3 serotypes of *Chlamydia trachomatis,* which are disseminated from the urogenital tract to local lymph nodes and then to skin. Accordingly, LGV is characterized by lymph node swelling and skin lesions, which are accompanied by systemic complications.[10]

Melioidosis

Burkholderia pseudomallei is a gram-negative bacillus and the causative agent of melioidosis, endemic in Southeast Asia and Northern Australia. The disease can be acquired through inhalation and ingestion or through cuts in the skin. Susceptible hosts can suffer abscess formation in multiple organs and, in some cases, disseminated infection, resulting in septic shock accompanied by pneumonia. There are an estimated 165 000 cases of melioidosis per year globally, resulting in approximately 89 000 deaths.[11]

Tularemia

This rare zoonosis in humans caused by *Francisella tularensis* is mainly found in rabbits and has recently gained wider recognition because of its potential for dual use.[12] Infection can be spread to humans via contaminated animals or tick bites. This gram-negative bacterium survives in macrophages and primarily causes acute pneumonia as well as skin sores, with subsequent involvement of lymph nodes.

Nongranulomatous Infections

Legionnaires' Disease or Legionellosis

Legionnaires' disease is caused by *Legionella pneumophila,* an environmental bacterium that persists within amoeba living in water reservoirs (*e.g.,* air-cooling systems), from where it is spread aerogenically. Infection is exacerbated by a compromised immune status. Characteristically, Legionnaires' disease presents as atypical pneumonia associated with general symptoms and is complicated by extrapulmonary infection, renal failure, and lung abscesses. Cases of Legionnaires' disease in the United States increased from 0.39 to 1.36 per 100 000 people from 2000 to 2011.[13]

Chlamydial Urethritis, Cervicitis, and Conjunctivitis

C. trachomatis serovars D–K enter and persist in epithelial cells of the urogenital tract, causing cervicitis and urethritis. In women, infertility can develop as a result of chronic or recurrent infection. In neonates, congenital infection during birth may result in conjunctivitis and pneumonia. Urogenital infections by chlamydiae occur worldwide and are now considered the most common sexually transmitted bacterial disease, with an estimated 100 million new infections occurring annually.[14]

Trachoma

Smear infections of the eye with *C. trachomatis* serovars A, B, and C cause inclusion conjunctivitis. As a consequence of multiple chronic infections and of the resulting immune response, scars develop that eventually injure the cornea, leading to trachoma. Approximately 84 million people are infected with *C. trachomatis* worldwide, 7.6 million of whom suffer from visual impairment.[15]

Chlamydia Pneumoniae

C. pneumoniae (formerly known as *C. trachomatis* TWAR strain) is the cause of mild respiratory disease in young adults and may cause serious infections in older, more debilitated patients. Atypical pneumonia may also be caused by *Chlamydia psittaci,* although this zoonosis, transmitted by birds, is relatively rare.

Typhus

Rickettsia prowazekii, R. typhi, and *R. tsutsugamushi* cause diseases of varying severity.[16] They are transmitted by arthropods and infect vascular endothelial cells at the site of an insect bite or scratch, causing skin reactions. Subsequently, pathogens are disseminated to the central organs, and more general symptoms develop. Globally, typhus is of minor importance.

Rocky Mountain Spotted Fever, Ehrlichiosis

Rocky Mountain spotted fever is caused by *Rickettsia rickettsii.* Infection of the vascular endothelium leads to systemic symptoms and skin manifestations that may be followed by shock and neurological complications.[16] Worldwide, this disease, as well as Mediterranean spotted fever caused by *Rickettsia conorii,* is of minor importance; as is probably Ehrlichiosis, a newly emerging zoonosis transmitted by ticks and caused by various *Ehrlichia* spp., mainly *E. chaffeensis.*[17] Disease manifestations include generalized symptoms, such as fever and muscle pain.

Bartonella

Bartonella spp. represent gram-negative facultative intracellular pathogens transmitted by insect vectors, such as fleas, sandflies, and mosquitoes.[18] The most clinically relevant species are *B. henselae, B. quintana,* and *B. bacilliformis. B. henselae* causes cat scratch disease (CSD) resulting in local lymphadenopathy in the lymph node draining the scratch site accompanied by fever, headache, and splenomegaly. Oculoglandular involvement (Parinaud syndrome), encephalopathy, neuroretinitis, or osteomyelitis can occur, albeit in rare cases. In immunosuppressed patients, bacillary angiomatosis and peliosis can occur, characterized by pseudotumoral proliferation of endothelial cells. Bacteria persist within erythrocytes with the intracellular location providing a protective niche.

GRANULOMA PATHOLOGY AS HALLMARK OF INTRACELLULAR BACTERIAL INFECTION

KEY CONCEPTS

Balance of Protection and Host Pathology in Granulomas

- Macrophage activation results in bacterial death (protective)
- Intracellular bacterial killing by "killer molecules" from T cells (protective)
- Lysis of infected macrophages by T cells results in release of bacteria and killing by more effective effector cells (protective) or bacterial dissemination (pathogenic)
- Development of central necrosis in granulomas results in death of tissue and bacteria (protective/pathogenic)
- Fibrotic encapsulation of granuloma results in containment of infection (protective)
- Overexuberant tissue fibrosis and necrosis (pathogenic)
- Liquefaction of central necrotic tissue in granulomas results in bacterial replication, cavity formation, and transmission of bacteria (pathogenic and contagious)

A characteristic feature of many infections caused by intracellular bacteria is the eventual need for tissue remodeling by the host at the site of infection. Granulomas are the result of an inability to rapidly clear host tissue of intracellular bacteria and represent a fascinating site of the host–pathogen interface (Fig. 26.1). The longevity of the granuloma depends directly on the continuous presence of the microbial pathogen, and the lesion generally disappears after its sterile eradication. Granulomas form the focus of the coordinated cross-talk between different types of T cells, B cells, and infected and uninfected mononuclear phagocytes (MPs) and dendritic cells (DCs). Even if the immune system fails to completely eliminate bacteria inside the granuloma, the latter performs a protective function by containing microbes within distinct foci and preventing their dissemination. At the same time, the granuloma can be detrimental to the host because it can interfere with physiological organ functions.[19] More detailed study of cellular phenotype within granulomas is starting to establish how cellular differentiation is orchestrated and how the granuloma develops.

Granulomatous lesions are generally initiated by nonspecific inflammatory signals mediated by bacterial products, chemokines, and proinflammatory cytokines that are produced by endothelial cells and MPs at the site of infection. Inflammatory phagocytes (of both monocytic and granulocytic origin) are attracted to the site of microbial replication, and an infiltrative, sometimes exudative, lesion develops. Following the accumulation and

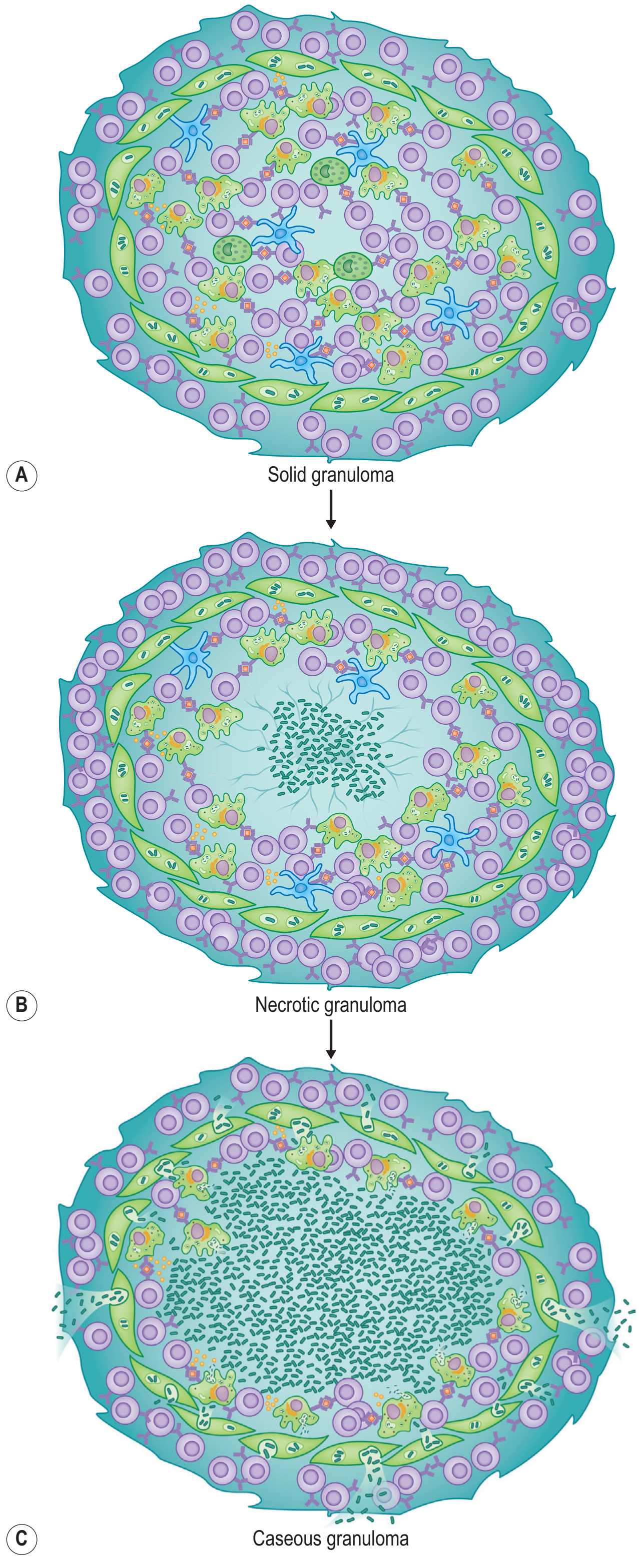


FIG 26.1 **Development of Granuloma Pathology and Implications for Tuberculosis (TB).** This figure depicts three distinct yet continuous stages of granuloma pathology in the lung due to *Mycobacterium tuberculosis* infection. (A) Solid granuloma: Composed largely of T cells and infected and uninfected MPs. These granulomas are defined by a lack of central necrosis and likely are representative of an ability to control *M. tuberculosis* replication. (B) Caseous/necrotic granulomas: These structures contain a central region of demarcated necrotic cell death. Bacteria are often detected within the caseous necrotic region and in proximal cells, notably mononuclear phagocytes (MPs). Since calcified caseous granulomas containing few bacteria have been observed, development of central necrosis may be a consequence of antibacterial mechanisms resulting in sacrifice of host cells to contain infection. (C) Cavity formation: These structures result from inability of caseous granulomas to contain bacterial replication. The acellular necrotic region, containing a large number of extracellular bacteria, increases in size and can liquefy and empty into the lung airways, resulting in transmission of viable bacteria via cough. Therefore granuloma formation is central to human-to-human spread of TB. Dissemination of bacteria through the bloodstream results in disease manifestation in other organs, such as the meninges and the urinary bladder.

activation of increasing numbers of MPs and DCs, this lesion takes an increasingly structured granulomatous form. A significant number of B cells is also found, which seem to influence granuloma morphology. Once specific T cells have been attracted to the lesion, it transforms into a productive granuloma that provides the most appropriate tissue site for antibacterial protection. Here, activation of MPs by interferon-γ (IFN-γ) and tumor necrosis factor-α (TNF-α) inhibits microbial growth. However, unbridled macrophage activation can have tissue-damaging effects, and mechanisms within the granuloma tightly regulate these effects. Eventually, the granuloma is encapsulated by a fibrotic wall, and its center becomes necrotic. Both tissue reactions are primarily protective, the former by promoting bacterial containment and the latter by reducing the nutrient and oxygen supply to the pathogen. The combined effects of chronic macrophage activation, persistence of intracellular bacteria, and hypoxia likely lead to enhanced cell death in the center of granulomas, resulting in the formation of a caseum. Caseation may favor the local replication of normally facultative intracellular bacteria in the cellular detritus, as well as microbial dissemination to distant tissue sites and to the environment to transmit infection. Hypoxia also has pronounced effects on enzyme functions that can dictate macrophage phenotype.

THE INTERDEPENDENCE OF INNATE AND ADAPTIVE IMMUNITY IN PROTECTION AGAINST INTRACELLULAR BACTERIA

Innate Immune Mechanisms as First-Line Defense

The interaction between host cell and pathogen that defines the intracellular lifestyle consists of a number of different layers. The first layer that differentiates intracellular bacteria from other bacteria, notably commensal bacteria that colonize the host but do not cause infection, is that of host cell entry. Extracellular bacteria are typically engulfed by professional phagocytes, which include tissue macrophages, DCs, and PMNs. This uptake is enhanced by host components of the complement system and antibodies, which bind to complement receptors (CRs) and Fc receptors, respectively, on professional phagocytes.[20] *M. tuberculosis* actively targets macrophages, where it must counteract numerous antimicrobial mechanisms operative in these cells (see below). Intracellular bacteria also use elaborate mechanisms to enter nonprofessional phagocytes, by which they must subvert host endocytic processes that are normally engaged in traffic of cellular cargoes. In some cases, this provides a less hostile environment because of their inability to efficiently mobilize antibacterial effector mechanisms. *Bartonella* spp., unique among intracellular bacteria, can enter red blood cells, thus allowing transmission via blood-sucking insect vectors. This represents a particularly advantageous niche as red blood cells lack the machinery to drive the adaptive immune responses required for protection. Entry into nonphagocytic host cells requires bacteria to induce their own internalization. Bacteria that colonize the GI tract (*i.e., L. monocytogenes* or salmonellae) or mucosal membranes of the urogenital tract (*i.e., C. trachomatis*) must mediate tight adhesion to the host cell membrane and be capable of mediating the uptake process. Broadly, two processes are utilized by bacteria to induce uptake into a nonphagocytic cell. The "zipper" mechanism is mediated by binding of a bacterial cell surface protein binding a cognate receptor on the host cell membrane. *L. monocytogenes* entry into intestinal epithelial cells depends on engagement of InlA to E-cadherin to mediate uptake.

Salmonellae and *C. trachomatis* use a "trigger" mechanism to induce internalization and inject multiple factors into the host cell cytoplasm to mediate uptake. These proteins are delivered by the needle-like structures that form part of bacterial type III secretion systems (T3SSs). These injected proteins then target host proteins involved in host cell signaling and actin remodeling to induce bacterial entry. The *C. trachomatis*–secreted proteins Tarp, CT166, and CT694 reversibly stimulate the Rho-family guanosine triphosphatase (GTPase) Rac1 to trigger internalization. Similarly, salmonellae inject T3SS factors to stimulate the Rho-family GTPases Cdc42 and Rac1. The success of these mechanisms of induced uptake enables intracellular bacteria to persist inside diverse cell types. *Rickettsia* spp., *C. trachomatis, M. leprae,* and *L. monocytogenes* ultimately target vascular endothelial cells, epithelial cells, Schwann cells, and hepatocytes, respectively, as their preferred intracellular habitats.

To prevent intracellular infection, the host depends on its ability to discriminate between host and bacterial molecules. As already mentioned, bacteria targeting the intracellular environment often do so via mucosal surfaces already populated by commensal organisms (the microbiome) that do not alert host defenses. The host must, therefore, discriminate between commensal and pathogenic bacteria via recognition of conserved molecular motifs of bacteria, named *pathogen-associated molecular patterns* (PAMPs). This occurs via host receptors broadly defined as pattern recognition receptors (PRRs, Table 26.2). The best-characterized group of PRRs is that of the so-called Toll-like receptors (TLRs). The TLR system constitutes an innate scanning mechanism of microbial pattern recognition to distinguish between a wide spectrum of bacteria and viruses. TLRs are present as homo- or heterodimers on the plasma membrane or within intracellular endosome/phagosome compartments.[21] PAMPs of bacterial origin comprise di- and triacyl lipoproteins, LPSs, and flagellin, which are recognized by TLR-2/6, TLR-2/1, TLR-4/4, or TLR-5/5, respectively. The vast array of mycobacterial cell wall lipids, such as lipoarabinomannan (LAM), trehalose dimycolate (TDM), and phosphatidylinositol mannosides (PIMs) bind either TLR-2 or TLR-4. Lipoteichoic acid (LTA) of gram-positive bacteria is recognized by TLR-2. TLR-9 binds low-methylated bacterial DNA containing CpG motifs within endosomes.

TABLE 26.2 Major Pattern Recognition Receptors Involved in Sensing of Intracellular Bacteria

Pattern Recognition Receptor	Location	Ligand
Toll-Like Receptors		
TLR1	Plasma membrane	Triacyl lipoprotein
TLR2	Plasma membrane	PGA, porins, LAM
TLR4	Plasma membrane	LPS
TLR5	Plasma membrane	Flagellin
TLR6	Plasma membrane	Diacyl lipoprotein
TLR7 (human TLR8)	Endosome	ssRNA
TLR9	Endosome	CpG DNA
Scavenger Receptors		
SR-A	Plasma membrane	LPS, LTA, CpG DNA, proteins
MARCO	Plasma membrane	LPS, proteins
CD36	Plasma membrane	Diacyl lipoprotein
LOX-1	Plasma membrane	Protein
SREC	Plasma membrane	Protein
C-Type Lectins		
DC-SIGN	Plasma membrane	LPS, ManLAM, capsular polysaccharide
MINCLE	Plasma membrane	Mycobacterial cord factor: TDM
NOD-Like Receptors		
NOD1	Cytoplasm	D-glutamyl-meso-diaminopimelic acid
NOD2	Cytoplasm	MDP
NLRP1	Cytoplasm	MDP
NLRP3	Cytoplasm	RNA, LPS, LTA, MDP
NLRC4	Cytoplasm	Flagellin
Naip5	Cytoplasm	Flagellin
AIM2-Like Receptors		
AIM2	Cytoplasm	dsDNA
IFI16	Cytoplasm	dsDNA
STING/cGAS Pathway		
cGAS	Cytoplasm	dsDNA

Note: We omit PRRs (*e.g.,* TLR3, which binds viral produced double-stranded RNA) not classically associated with intracellular bacteria. AIM2, absent in melanoma-2; CD36, cluster of differentiation 36; CpGDNA, cytosine-phosphatidyl-guanine DNA; DC-SIGN, dendritic cell–specific intercellular adhesion molecule-3-grabbing nonintegrin; dsDNA, double-stranded DNA; LAM, lipoarabinomannan; LOX-1, lipoxygenase-1; LPS, lipopolysaccharide; LTA, lipoteichoic acid; ManLAM, mannose lipoarabinomannan; MARCO, macrophage receptor with collagenous structure; MDP, muramyl dipeptide; MINCLE, macrophage-inducible C-type lectin; NLR, NOD-like receptor; NOD, nucleotide-binding domain; PGA, peptidoglycans; PRR, pattern recognition receptor; SR, scavenger receptor; ssRNA, single-stranded RNA; SREC, scavenger receptor expressed by endothelial cell-I; TDM, trehalose dimycolate; TLR, Toll-like receptor.

Scavenger receptors and C-type lectins are also PRRs and function at the cell membrane.[22] Scavenger receptors were first defined by their ability to transport modified forms of low-density lipoproteins inside cells, indicating their ability to also interact with host molecules. However, receptors, such as SR-A, MARCO, CD36, LOX-1, and SREC, can bind a wide array of bacterial molecules, such as lipids, CpG DNA, and proteins (see Table 26.2 for binding specificities). SR-A is important for clearance of extracellular bacteria from the spleen and liver. MARCO expressed on alveolar macrophages is implicated in clearance of pneumococcal bacteria preventing pneumonia. C-type lectins are similarly membrane-expressed and include DC-specific intercellular adhesion molecule–grabbing nonintegrin (DC-SIGN); mannose receptor; dectin-1; dectin-2, which chiefly recognize fungal components; and MINCLE, which recognizes trehalose dimycolate (TDM), the cord factor of *M. tuberculosis.*

It has been suggested that scavenger receptors and C-type lectins are required to bind and internalize the bacillus, whereas it is primarily the TLRs that discriminate between the pathogens and initiate the necessary intracellular signaling events. It should, however, be noted that intracellular signaling events can also be triggered by other interactions, such as ligand binding to macrophage mannose receptor (MMR), dectin 1, or DC-SIGN. Far from a one-ligand, one-receptor binary mechanism of sensing and signaling, PRRs often collaborate to produce multiprotein complexes. CD14, MD2, and TLR4 collaborate for LPS sensing and signaling. Similarly, MARCO and TLR2 synergize to recognize TDM. To allow signaling, these complexes interact with adaptor proteins containing immunoreceptor tyrosine-based activation motif (ITAM)-like or Toll/interleukin-1 receptor (TIR) domain motifs. TLR signaling occurs via the adaptor proteins MyD88, TIRAP/Mal, and Trif. These molecules then orchestrate a downstream signaling cascade, which culminates in induced patterns of gene transcription that mediate innate and, ultimately, adaptive immune mechanisms that aim at combating the intracellular bacteria.

The cellular cytoplasm is monitored for presence of molecules of bacterial origin by a further group of PRRs, the nucleotide oligomerization domain protein-like receptors (NLRs). These molecules are characterized by presence of a nucleotide-binding domain and leucine-rich repeat motifs. Molecules from this group recognizing bacterial components are nucleotide-binding oligomerization domain (NOD)–containing proteins NOD1 and NOD2, NOD-like receptor P1 (NLRP1), NLRP3, and Naip5. Other cytosolic PRRs include the absent in melanoma-2-like receptor (ALR) family, cyclic guanosine monophosphate–adenosine monophosphate (cGMP-AMP) synthase (cGAS), and stimulator of interferon (IFN) genes (STING), all of which can be activated by bacterial DNA (see Table 26.2).

Engagement of NLRs and ALRs leads to activation of the multiprotein complex called the *inflammasome,* leading to cleavage of pro-IL-1β and pro-IL-18 to produce their active forms. In addition, activation of the NLRs, NOD1, and NOD2 results in inflammatory cytokine secretion. Certain PRRs are also receptive to certain endogenous “danger” signals produced by tissues undergoing stress, damage, or cell death. These signals are triggered by self-proteins, named *danger-associated molecular patterns* (DAMPs), include endogenous heat shock proteins, host nucleotides, and the chromatin component HMGB1. Therefore PRRs mediate signals not only emanating from intracellular bacteria but also from host cells damaged by the infection process. Understanding how PAMP and DAMP PRR signaling meshes to produce a coherent disease-specific output remains an exciting challenge for future research.

As already alluded to, the culmination of PRR collaborative sensing and signaling is the induction of inflammatory genes that induce innate immune mechanisms and, subsequently, the mobilization of the adaptive immune response. These include cytokines that act both locally and systemically and are important mediators of protection against intracellular bacteria via specific signaling through engagement of host cell surface receptors. Such engagement mobilizes both critical mechanisms of host protection and orchestration of adaptive immune responses.

Macrophage Training by Epigenetic Mechanisms

Following infection with, for example, *M. tuberculosis,* macrophages express elevated effector mechanisms over long periods. More recent findings have revealed that this is caused, at least in part, by epigenetic changes that are induced in MPs during infection. Similar training can also occur after vaccination with live vaccines. Moreover, BCG vaccination was shown to induce changes in methylation patterns of the *NOD2* gene in humans. These epigenetic changes led to increased production of proinflammatory cytokines by MPs, namely, TNF and IL-1β, a tendency that persisted for 3 months after vaccination. Furthermore, this enhanced innate immunity led to increased resistance not only against *M. tuberculosis* but also against other bacterial pathogens. Indeed, it has been speculated that epigenetic alterations in MPs following vaccination are responsible for the nonspecific reduction of mortality in BCG-vaccinated infants in resource-poor regions of the world.[23]

Cytokines as Mediators of Defense Against Intracellular Bacteria

We have already mentioned that a range of cytokines are induced by the signaling mechanisms that result from engagement of PRRs. These serve by both enhancing intracellular mechanisms of bacterial killing and mobilizing adaptive immune responses, representing the next layer of host defense. Because these responses allow an amplification of the initial innate immune responses, they must be carefully regulated by the host to prevent extensive tissue pathology. In fact, we might view the development of a granuloma as the sequela of a balance between bacterial killing mechanisms and the need to restrict tissue pathology orchestrated by adaptive immunity. At the onset of infection, initial cytokine secretion occurs in the cell type that initially encounters intracellular bacteria and on initiation of signaling cascades by PRRs. These small molecules can act locally and systemically to directly instruct cells, to produce antibacterial molecules, to combat intracellular infection, and to both increase numbers of immune cells and direct the composition of the cellular infiltrate that will ultimately attempt to resolve intracellular bacterial infection. Cytokines are ultimately produced by multiple cell types, including adaptive T cells, B cells, unconventional T cells, MPs, DCs, PMNs, and even epithelial and endothelial cells. We will first consider the hierarchy by which these cytokines act in the control of intracellular bacterial infection and the antibacterial mechanisms they regulate. We will then return to the generation and regulation of the cells that produce them.

IFN-γ, TNF-α, IL-12, and IL-18

By far, the cytokine with the clearest demonstrable potency against intracellular bacteria is IFN-γ. Extensive studies on the activation of antibacterial effector functions in macrophages have revealed a central role for IFN-γ. Accordingly, IFN-γ neutralization with antibodies, or deletion of the IFN-γ gene by homologous recombination, markedly exacerbates infectious diseases, such as

KEY CONCEPTS

T Cell–Mediated Mechanisms Underlying Protection

- IFN-γ- and TNF-α-mediated activation of phagocytes to kill bacteria by means of:
  - Reactive oxygen intermediate (ROI) and reactive nitrogen intermediate (RNI)
  - Delivery of lysosomal hydrolytic enzymes and antimicrobial peptides to the bacteria-containing phagosome
  - Autophagy
  - Formation and maintenance of granulomas
- T cell–mediated response controls but does not eradicate the pathogen

listeriosis, TB, or typhoid in experimental animals. For macrophages harboring intracellular bacteria, namely, *M. tuberculosis*, signaling with IFN-γ is a game-changer, summoning infected macrophages to escalate antimicrobial mechanisms. The action of TNF-α appears to augment IFN-γ and is also important in control of intracellular infection. This has been demonstrated in humans through use of blocking of TNF-α by antibodies as antiinflammatory therapy. Such treatments can activate TB in individuals with LTBI. Despite this, these potent protective effects of IFN-γ and TNF-α come at a price. The need to kill intracellular bacteria often leads to death of the host cell as collateral damage. In part, the host manages this by controlling how the cell dies. Excessive TNF-α leads to less regulated necrotic cell death, benefiting *M. tuberculosis*. For this reason elaborate host mechanisms have evolved to maintain TNF-α at optimal levels to control infection. The host enzyme leukotriene A4 hydrolase (LT4H) catalyzes synthesis of a highly proinflammatory lipid leukotriene B4. In the event of enzyme deficiency, an antiinflammatory lipid lipotoxin A4 accumulates that counteracts effects of TNF-α. Two common variant promoters control expression of *LT4H* in humans, and homozygotes are associated with either high or low inflammation. In contrast, the heterozygotes show a balanced response to TNF-α associated with resistance against TB. Such a finding strongly suggests that genetic mechanisms can maintain an optimal level of TNF-α responsiveness of cells harboring intracellular bacteria, namely, *M. tuberculosis*. A central antimicrobial mechanism stimulated by IFN-γ and TNF-α is production of reactive nitrogen intermediates (RNIs) via the induction of nitric oxide synthase (NOS)2 and reactive oxygen intermediates (ROIs) via activation of nicotinamide adenine dinucleotide phosphate (NADPH)–dependent oxidative burst. IFN-γ also promotes antimicrobial effects associated with vitamin D and induces autophagy, a mechanism that plays an important role in host defense. It is now clear that the production of IFN-γ depends on prior activation by IL-12 and/or IL-18. IL-12, in concert with TNF-α, induces a cytokine loop resulting in the production of IFN-γ, which sustains the production of IL-12 and IL-18. These observations have been extended to humans, in whom mutations that affect IFN-γ signaling cause susceptibility to *M. tuberculosis* and salmonellae, as well as to BCG and commonly nonpathogenic mycobacteria, and are termed *mendelian susceptibility to mycobacterial disease* (MSMD). The mutations are located in genes that include *IL12B* and *IL12RB,* which encode subunit β of the IL-12 cytokine and its receptor, respectively, and *IFNGR1* and *IFNGR2,* which encode the IFN-γ receptor. Further unraveling of the molecular basis of human genetic susceptibility to intracellular bacterial infection will continue to illuminate our understanding of how immunity is similarly orchestrated across multiple infectious diseases.

Proinflammatory Cytokines and Phagocyte Attraction

The recruitment of more phagocytes to the site of infection represents a vital process in the resolution of infection. Phagocyte recruitment is achieved via the secretion by MPs and endothelial cells of cytokines of the IL-1 family, TNF-α, IL-6, and chemokines. Signaling via IL-1 cognates is considered closely related to that of the TLRs because of the close homology of the cytoplasmic domains of TLRs and IL-1 family receptors. The most studied member is IL-1β, which, in synergy with chemokines and TNF-α, increases the expression of adhesion molecules on the vascular epithelium, thereby promoting extravasation of the inflammatory cell infiltrate into infected tissues. Chemokines are a family of structurally related proteins. The positions of the first two cysteine residues in the protein sequence have been used to divide chemokines into four subfamilies: CC (MIP-1β, MCP-1, MCP-2, MCP-3), CXC (MIP-2, IL-8), C (lymphotactin), and CX3C chemokines (fractalkine), where C represents cysteine and X represents any amino acid other than cysteine. These molecules are critical in controlling the migration of PMNs (IL-8) and monocytes (MCP-1, also known as CCL2) from the bloodstream to infected tissue.[24] Recently, the role of chemokines in intracellular infections has been increasingly appreciated, for example, with mice lacking the receptor for CCL2 being deficient in their ability to clear listeria infection. It has been suggested that in the early stages of infection, *M. tuberculosis* exploits a delay in the mobilization of T-cell immunity to recruit MPs to the site of infection, which preferentially serve as habitat because of a lack of local IFN-γ from T cells. Moreover, *M. tuberculosis* is thought to infect the relatively sterile lower airways. The lack of commensal bacteria could mean that *M. tuberculosis* uses cell surface phenolic glycolipid (PGL) to signal epithelial cells to produce the chemokine CCL2 in the absence of signaling via other PAMPs.[25] This mechanism then recruits MPs that are more permissive for bacterial growth than those recruited by a more "global" MyD88-dependent signaling of TLRs, requiring coengagement of PAMPs on commensal bacteria that are more abundant in the upper airways. The initial macrophage infiltrates could play an important role in early granuloma development.

Cytokine-Induced Host-Protective Mechanisms

Effector Molecules

Activation of a membrane-bound NADPH oxidase by stimulation with IFN-γ or immunoglobulin G (IgG) initiates an oxidative burst that generates the ROIs O_2^-, H_2O_2, OH^-, $^1O_2$, and •OH radical (Table 26.3).[26] In human PMNs and blood monocytes that possess myeloperoxidase, ROI activity is further augmented by the formation of hypochlorous acid. Oxidation and/or chlorination of bacterial lipids and proteins result in their inactivation and subsequent bacterial killing. The importance of ROIs in antibacterial defense is underlined by recurrent infections in patients whose phagocytes fail to generate an oxidative burst. Nitric oxide synthase 2 (NOS2) is an inducible cytosolic enzyme in professional phagocytes that delivers NO to the phagolysosome harboring bacteria while consuming O_2 and L-arginine. NO is further oxidized to NO_2^- and NO_3^-. Nitrification and/or oxidation then inactivates bacterial molecules needed for bacterial growth.[27] The formation of •NO is catalyzed by NOS2, which is promoted by both immunological stimuli, such

TABLE 26.3 **Antibacterial Effector Mechanisms of Activated Macrophages and Corresponding Microbial Evasion Strategies**

Macrophage Effector Mechanism	Microbial Evasion Strategy
Production of ROIs	Uptake via complement receptors; production of ROI detoxifying molecules (superoxide dismutase, catalase); bacterial ROI scavengers (phenolic glycolipids, sulfatides, lipoarabinomannans)
Production of RNIs	Inhibition of phagosome maturation via blockage of H^+ATP pump, indirect effect of ROI-detoxifying molecules
Autophagy, intraphagolysosomal killing	Egression into cytoplasm; Resistant cell wall
Phagosomal acidification, phagosome–lysosome fusion	Inhibition of phagosome maturation
Defensins	Modification of cell wall lipid A to resist defensins
Reduced iron supply (transferrin receptor downregulation, lipocalins)	Expression of microbial siderophores to increase iron uptake
Tryptophan degradation	Upregulation of bacterial tryptophan synthesis

ATP, adenosine triphosphate; ROI, reactive oxygen intermediate; RNI, reactive nitrogen intermediate.

as IFN-γ and TNF, and microbial products, such as LPS, lipoteichoic acid, and mycobacterial lipids. RNIs exert their bactericidal activity by destroying iron-/sulfur-containing reactive centers of bacterial enzymes and by synergizing with ROIs to form highly reactive peroxynitrite ($ONOO^-$). Despite being highly effective in killing intracellular bacteria, NO production relies on a continuous supply of L-arginine, which becomes limited because of competition with another macrophage enzyme, Arginase-1 (Arg-1). Arg-1 metabolizes L-arginine to produce urea and ornithine and demonstrates antiinflammatory activity. The competitive function of Arg-1 likely regulates collateral tissue damage caused by overexuberant RNIs. The final downstream product of NOS2 activity is citrulline, which is recycled to L-arginine by the enzymes argininosuccinate synthase (Ass1) and argininosuccinate lyase (Asl). A mouse deficient in macrophage Asl activity is unable to control mycobacterial infection, highlighting the importance of this recycling pathway. The central role of NOS2 in protection against intracellular bacteria is well established in murine models of infection. Whether NOS2 plays a similarly central role in humans is still unclear. Defensins are small lysosomal polypeptides that are microbicidal at basic pH and are particularly abundant in phagocytes. These include granulysin, present in granules of human natural killer (NK) and cytolytic T (CTLs) cells,[28] and cathelicidin, which is regulated by vitamin D in a TLR-dependent manner and is converted by cleavage to the antimicrobial peptide LL-37.

Apoptosis and Autophagy

Apoptosis is a highly regulated form of cell death that is critical for control of cell turnover, a vital process for tissue homeostasis.[29] Macrophage apoptosis also constitutes a defense mechanism, allowing removal of phagocytes containing intracellular bacteria without the need to generate significant inflammation. Apoptosis, in contrast to cellular necrosis, results in cell death without permeabilization of the host cell membrane. The process can be triggered by TNF-α signaling and augmented by IFN-γ, resulting in activation of cellular caspases, mitochondrial membrane permeability, and cytochrome c release. These processes result in cellular disintegration and generation of apoptotic bodies that are engulfed and digested by neighboring phagocytic cells. Apoptosis is protective against *L. monocytogenes* and *Salmonella* spp. and is inhibited by *M. tuberculosis,* which promotes necrotic cell death of infected cells to its benefit via mitochondrial membrane damage and by caspase-independent mechanisms during conditions of high bacterial burden in macrophages. Noninfected cells engulf bacterial antigens associated with vesicles produced by apoptotic cells. Apoptosis as a prerequisite for this pathway is induced by many intracellular bacteria, including salmonellae, mycobacteria, and listeriae. This cross-presentation pathway in infections with intracellular bacteria adds an essential function to the physiological role of apoptosis in the maintenance of tissue integrity and growth.

Upon signaling via IFN-γ, autophagy, a process common to all cells for removal of dysfunctional or damaged cellular organelles, can be harnessed to dispose of intracellular *L. monocytogenes* and *M. tuberculosis* in a process termed *xenophagy.* Signaling via members of the immunity-related guanosine triphosphatase (GTPase) family (IRG family) and the guanylate-binding protein family, TLR2 and TLR4 engagements and the active form of vitamin D_3 all act to augment xenophagy. Formation of double-membrane autophagosomes that mature analogously to the phagosomal pathway and fuse with lysosomes that degrade bacteria contained within. The importance of this process is highlighted by polymorphisms in one of the three IRG families of genes in humans, *IRGM,* being associated with susceptibility to TB. Recently, a host-encoded microRNA, miRNA-155, was shown to potentiate xenophagy during intracellular mycobacterial infection by targeting an endogenous inhibitor of autophagy, Ras homologue enriched in brain (Rheb) by suppressing Rheb expression.

Nutrient Deprivation

Deprivation of required nutrients to intracellular bacteria is also a strategy employed by the host, markedly so within infected macrophages. Tryptophan degradation is achieved by the enzyme indoleamine 2,3-dioxygenase (IDO), which degrades tryptophan to kynurenine (see Table 26.3). This reaction is induced by IFN-γ in both MPs and IFN-γ-responsive nonprofessional phagocytes and inhibits the growth of *C. psittaci* and *C. trachomatis* inside human macrophages and epithelial cells. Similarly, augmentation of NOS2 by IFN-γ and TNF-α depletes intracellular L-arginine, also required for growth of intracellular bacteria.[20]

EVASION FROM, INTERFERENCE WITH, AND RESISTANCE TO MICROBIAL KILLING

Strategies Against Toxic Effector Molecules

Many intracellular bacteria have exploited successful strategies against macrophage effector mechanisms (see Table 26.3). One mechanism of evasion is determined by the receptor that is used for pathogen entry into the host cell. Internalization via CRs inhibits the production of IL-12, a cytokine critical in facilitating macrophage activation. Engulfment by this receptor also bypasses activation of the oxidative burst, thereby avoiding ROI production. Similarly, engaging mismatch repair (MMR) and DC-specific intercellular adhesion (DC-SIGN) molecules for uptake triggers secretion of the suppressive cytokines IL-10 and TGF-β. Several intracellular bacteria also produce ROI detoxifiers, including superoxide dismutase and catalase, which nullify oxygen (O_2) and hydrogen peroxide (H_2O_2), respectively. Finally, a number of small bacterial products, such as the phenolic glycolipid and LAM of mycobacteria, scavenge ROIs. Many of the strategies used to counteract the effects of ROIs also overlap in their effects on RNIs. A modification of lipid A renders gram-negative bacteria, including salmonellae, resistant to the effects of host antimicrobial peptides.

Intraphagosomal Survival

Inhibition of phagolysosome fusion represents a major intracellular survival strategy for a number of intracellular bacteria, including *M. tuberculosis*, *Francisella* spp., *Brucella* spp., and *L. monocytogenes* (Fig. 26.2). After engulfment, these pathogens manipulate the endocytic fate of the phagosome that contains them. This is achieved in part by manipulation of Rab GTPases, proteins required for normal endocytic trafficking, positioned in the phagosome membrane.[30] Rab GTPases associated with phagosome maturation of the pathogen-containing phagosome are Rab3, -4, -5, -9, -7, -11, and -14. These proteins are associated with different maturation stages of the phagosome and chiefly orchestrate membrane fusion events to allow delivery of vesicular protein cargo to the phagosomal compartment. The mycobacteria-containing phagosome acquires Rab5a but not the late endosomal marker Rab7a, which ultimately mediates fusion of the bacterial-containing phagosome with lysosomes that contain proteolytic enzymes active at low pH. By enabling arrest of this maturation, *M. tuberculosis* maintains its compartment at an early endosomal stage. This compartment does not acidify, partly because of a paucity of vacuolar H^+ adenosine triphosphatase (ATPase); at the same time, it exchanges molecules with the plasma membrane, such as the transferrin receptor, to access iron. Activation of macrophages with IFN-γ restores the normal maturation of the mycobacterial phagosome, resulting in a drop in mycobacterial viability. Francisellae and brucellae are engulfed by phagosomes that acquire the early endosomal markers EEA1 and Rab5a.[31] The *Francisella*-containing vacuole acquires late endosomal markers, but the pathogen escapes into the cytosol by perforating the late endosomal membrane. After a transient phagosomal stage, brucellae enter compartments enclosed by endoplasmic reticulum (ER) membranes to escape delivery to phagolysosomes.

Phenotypic Plasticity of the Infected Cell

The ability of intracellular bacteria to influence the phenotypic fate of both the cell in which it dwells and the cells within lesions that form as a result of unresolved infection is becoming increasingly

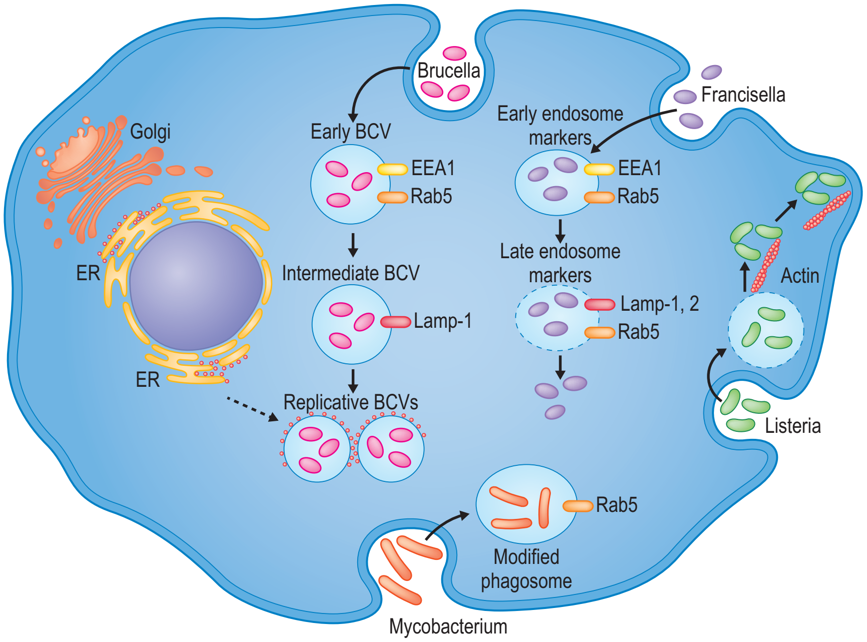


FIG 26.2 **Inhibition of Phagolysosome Fusion Represents a Major Survival Strategy for a Number of Intracellular Bacteria, Including *Mycobacterium Tuberculosis, Francisella* spp., *Brucella* spp., and *Listeria Monocytogenes.*** The mycobacteria-containing phagosome acquires Rab5 but not the late endosomal markers Lamp-1 and 2, enabling arrest of maturation of this compartment at an early endosomal stage. *Francisella* spp. and *Brucella* spp. are engulfed by phagosomes that acquire the early endosomal markers EEA1 and Rab5. The *Francisella*-containing vacuole acquires late endosomal markers but escapes into the cytosol by perforating the late endosomal membrane. A similar strategy is adopted by *L. monocytogenes.* After a transient phagosomal stage, brucellae enter compartments enclosed by endoplasmic reticulum (ER) membranes to escape delivery to phagolysosomes. BCV, Brucella-containing vacuole.

appreciated. This has already been highlighted by the tendency of *M. tuberculosis* to use its own membrane lipids to exploit host chemotactic pathways to recruit bacterial growth-permissive macrophages to the site of infection; this process subsequently allows a proliferative head-start before adaptive immunity kicks in and amplifies intracellular defences by the action of cytokines, such as IFN-γ and TNF-α. Moreover, bacterial killing must be tempered inside the granuloma to prevent destruction of host tissue. This is achieved by balancing macrophage phenotypes ranging from a phenotype highly bactericidal, termed "classically" activated, to a phenotype that is more suppressive of inflammation and is associated with wound healing and fibrosis, termed "alternatively" activated. Tipping the balance one way or the other is detrimental for the host in terms of disease.

Myeloid-derived suppressor cells (MDSCs) represent a certain stage of development of myeloid cells (both of monocytic and granulocytic lineage).[32] Although most of our knowledge stems from their suppressive role in cancer, recent evidence suggests that they play a role in control of chronic infections, such as TB.[33] They can be distinguished from canonical MPs and granulocytes by means of distinct surface markers. The granulocytic MDSCs are CD11b$^+$ LY6G^{hi} Gr1^{int}, whereas monocytic MDSCs are CD11b$^+$ LY6G^{neg} LY6C$^+$ Gr1^{hi}.

More recent findings point to host cell reprogramming resulting from intracellular infection. During intracellular infection of Schwann cells, *M. leprae* is able to downregulate genes active for the Schwann cell phenotype and upregulate genes that orchestrate differentiation to a "stem cell–like" phenotype. This stem cell–like property allows the infected cell to differentiate further to multiple mesenchymal cell states, such as skeletal cells or smooth muscle cells.[34] This ability to regress and then reprogram an infected cell phenotype could play a role in spreading infection throughout the host during leprosy.

Recently, mesenchymal stem cells (MSCs) were identified as an intracellular niche of *M. tuberculosis* in mice, and the equivalent human MSC phenotype could be readily infected *in vitro*. Because they reside in hypoxic niches and most antimycobacterial therapies are inactive in these conditions, it is feasible that MSCs could maintain the bacteria during long-term infection and could represent a protective niche from drug therapy. Dormant *M. tuberculosis* has also been detected in hematopoietic stem cells (HSCs) in mice and humans. HSCs are pluripotent, giving rise to both lymphoid and myeloid cell lineages in the blood. Clarification of the pathophysiological context of carriage of M. tuberculosis by both HSCs and MSCs is an exciting prospect.

Escape Into Cytoplasm

A successful strategy for survival inside activated macrophages is egression from the phagosome into the cytoplasm, which has been exploited by *L. monocytogenes* and the various pathogenic *Rickettsia* spp. (see Fig. 26.2).[35,36] This has the advantage of both avoiding the cellular defense mechanisms within the phagosome and providing the bacteria with a nutrient-rich environment. *L. monocytogenes* possesses several virulence factors to facilitate its escape from the phagolysosome, a pore-forming hemolysin (listeriolysin [LLO]) that acts together with a metalloproteinase, a lecithinase, and two phospholipases to efficiently promote the rupture of the phagosomal membrane and to spread to other cells. *M. tuberculosis* and *M. leprae* can also egress from the phagosome into the cytoplasm of macrophages and DCs, a behavior that is mediated by a mycobacterial protein secretion system ESX-1.[37] Bacterial virulence factors secreted by ESX-1 may also contribute to increased cell death and are not expressed by the vaccine BCG, which hence does not escape from the phagosome.

T LYMPHOCYTES AS SPECIFIC MEDIATORS OF ACQUIRED RESISTANCE

Activated macrophages act as the nonspecific executors, whereas T lymphocytes are the specific mediators of acquired resistance against intracellular bacteria. The dramatic increase in the incidence of TB and other intracellular bacterial infections in patients with AIDS illustrate the central role of T lymphocytes in protection. For instance, 15 million individuals are coinfected with HIV and *M. tuberculosis,* and HIV increases the risk of developing TB by several orders of magnitude resulting in more than 1 million TB cases annually. At the site of microbial growth, T lymphocytes not only initiate the most potent defense mechanisms available, they also focus this response to the site of encounter, thus minimizing collateral damage to the host. Although protective T-cell responses are multifactorial, they can be reduced to a few principal mechanisms (Fig. 26.3).

As previously mentioned, T cells inevitably also produce pathology through cytotoxic antimicrobial defense mechanisms. Moreover, pathogenesis of intracellular bacterial infection is highly influenced by T cells. It is therefore important that the T-cell response be tightly controlled and downregulated, when necessary. Regulatory mechanisms, including regulatory T cells (Tregs), are in place to limit immunopathology.[19]

Protective immunity involves the so-called conventional T-cell sets, CD4 αβ T cells, and CD8 αβ T cells, as well as unconventional T cells, such as γδ T cells, CD1-restricted αβ T cells, and T cells that recognize antigen in the context of other nonclassic MHC class I molecules, such as mucosal-associated invariant T (MAIT; Chapter 20) cells (see Fig. 26.3). Although these T-cell sets perform different tasks, substantial redundancy exists. Furthermore, these T-cell populations act in a coordinated way in close interaction with other leukocytes. Depending on the etiological agent and the stage of disease, the relative contribution of the different T-cell subsets to acquired resistance may vary. The conventional αβ T cells make up more than 90% and γδ T cells less than 10% of all lymphocytes in the blood and peripheral organs of humans and mice. However, γδ T cells represent a significant proportion of the intraepithelial lymphocytes in mucosal tissues, suggesting a particular role at this important port of microbial entry.

CD4 T Cells

The CD4 T-cell population can be further subdivided into distinct subsets, according to their pattern of cytokine production and expression of unique transcription factors that control patterns of gene expression (Chapter 16). At least four major subsets exist, T-helper cell-1 (Th1), Th2, Th17, and Tregs. The first two subsets were discovered over 20 years ago and have been identified in both mice and humans: Th1 cells, which overwhelmingly produce IFN-γ and IL-2, and Th2 cells, which produce IL-4, -5, and -13. The Th1 subset can also be defined on the basis of the T-bet transcription factor and the signal transducer STAT4, whereas Th2 classification is consistent with expression of the transcription factor GATA-3 and signal transducer STAT5.

Th17 cells express the retinoid orphan receptor γt (ROR-γt) transcription factor and the signal transducer STAT3. They produce the cytokines IL-17, IL-22, and granulocyte macrophage–colony-stimulating factor (GM-CSF). Cytokines of the IL-17 family are strong inducers of granulopoiesis; of proinflammatory mediators,

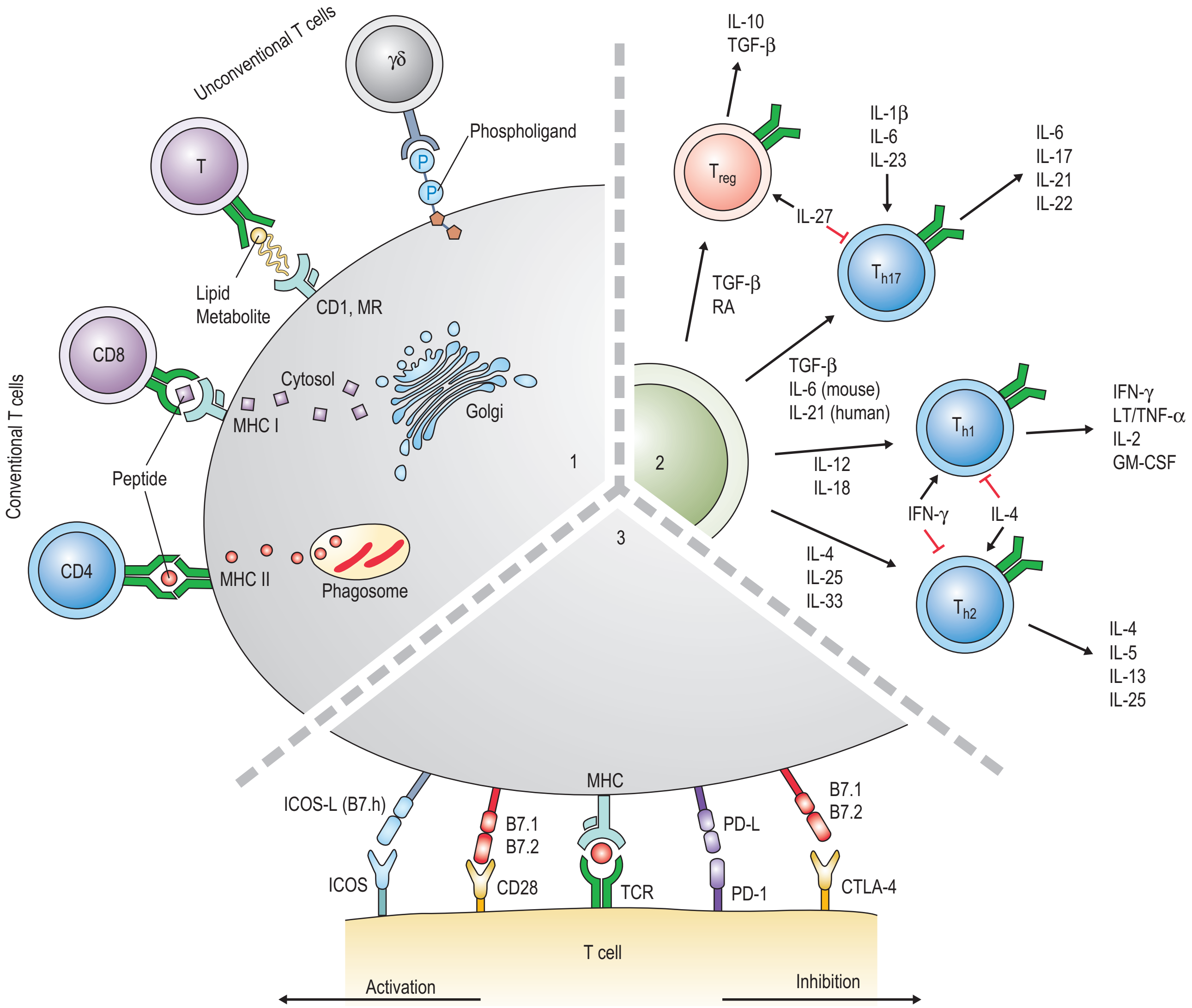

Unconventional T cells
γδ
Phospholigand
P
P
T
Lipid
Metabolite
CD1, MR
CD8
Cytosol
MHC I
Golgi
Conventional T cells
Peptide
CD4
MHC II
Phagosome
1
2
3
IL-10
TGF-β
Treg
IL-27
IL-1β
IL-6
IL-23
IL-6
IL-17
IL-21
IL-22
Th17
TGF-β
RA
TGF-β
IL-6 (mouse)
IL-21 (human)
Th1
IFN-γ
LT/TNF-α
IL-2
GM-CSF
IL-12
IL-18
IFN-γ
IL-4
IL-4
IL-25
IL-33
Th2
IL-4
IL-5
IL-13
IL-25
ICOS-L (B7.h)
MHC
B7.1
B7.2
B7.1
B7.2
PD-L
ICOS
CD28
TCR
PD-1
CTLA-4
T cell
Activation
Inhibition

FIG 26.3 **T-Cell Stimulation During Infection.** (A) Recognition of bacterial antigen by T cells. Antigen originating from intracellular bacteria is presented to conventional CD4 and CD8 T cells. Unconventional T cells including γδ T cells, mucosa-associated invariant T (MAIT) cells, and CD1-restricted T cells are also activated. Human γδ T cells recognize small molecules containing pyrophosphate residues; MAIT cells recognize bacterial metabolites, such as vitamin B_2 derivatives, in the context of major histocompatibility complex (MHC)–related (MR) gene products; CD1-restricted T cells recognize glycolipids in the context of CD1 molecules. (B) CD4 T cells can be subdivided into different T helper (Th) cells according to their cytokine expression pattern. Th1 cells are critical for protection against intracellular bacteria. They typically produce interferon-γ(IFN-γ), tumor necrosis factor-α(TNF-α), lymphotoxin (LT), interleukin (IL)-2, and granulocyte macrophage–colony-stimulating factor (GM-CSF). Th2 cells stimulate humoral immune responses via secretion of IL-4 and IL-5. Other cytokines produced by Th2 cells include IL-13 and IL-25. Th17 cells produce IL-6, -17, -21, and -22, which probably contribute to early protection. Regulatory T cells (Tregs) produce transforming growth factor β(TGF-β) and IL-10, which suppress immune responses. ICOS, inducible costimulatory molecule; PD-1, programmed death 1; PD-L, program death ligand; RA, retinoic acid; TCR, T-cell receptor. See Chapters 9 and 12 for details. (Modified from Kaufmann SHE, Parida SK. Tuberculosis in Africa: learning from pathogenesis for biomarker identification. Cell Host Microbe 2008;4:219–28, with permission of Elsevier.)

such as IL-6; and of the chemokines CXCL1, CXCL8, and CXCL6, which attract neutrophilic and eosinophilic granulocytes and prolong their survival.[38] Th17 cells, too, have limited importance for protection in murine models against primary infection with mycobacteria, salmonellae, and listeriae. However, Th17 cells can drive more rapid Th1 responses against pulmonary TB in mice after vaccination, resulting in enhanced protection. IL-17 is also required for optimally protective Th1 responses during murine *F. tularensis* infection.[39]

Despite the convenience of defining T-cell populations in terms of subsets, recent evidence suggests considerable plasticity in cytokine production by T cells. This was first suggested by demonstration that all subsets could produce IL-10, which regulates potency of T-cell responses to limit host collateral damage during immune responses. IL-10 expression might be an intrinsic control mechanism common to all T cells. However, reduction in T-cell potency also favors chronic intracellular bacterial infection. T-cell subsets may acquire the ability to produce additional cytokines by expression of additional transcription factors or by remodeling chromatin structure.[40] Future research will redefine T-cell behavior during intracellular bacterial infection.

CD8 T Cells

Infection of mice deficient in specific T-cell subsets has conclusively demonstrated a role for CD8 T cells during listeriosis and TB.[41] Furthermore, CD8 effector T cells have been identified in granulomas of patients with tuberculoid leprosy with low numbers of bacteria. The cytolytic potential of these T cells can serve two roles in infection with intracellular bacteria, namely, target cell killing or lysis of cells that are unable to control the infection, thus releasing the bacteria for phagocytosis by more activated cells. In humans, CD8 T cell–mediated killing is cell contact dependent and based on production of perforin, granzymes, and granulysin.[42] Finally, CD8 T cells are also a potent source of IFN-γ and TNF-α, thus contributing to direct activation of infected macrophages to enhance protective mechanisms (see Fig. 26.3).

CD8 T cells recognize antigenic peptides in the context of MHC class I gene products, which are responsible for presentation of antigens residing in the cytosol. Initially, therefore, it was mysterious how CD8 T cells were stimulated by intracellular bacteria, which were thought to have a uniquely restricted phagosomal residence. However, with the knowledge that many intracellular bacteria egress into the cytoplasm, one major mechanism for MHC class I processing became obvious: proteins secreted by bacteria in the cytoplasm undergo antigen processing and presentation similarly to newly synthesized proteins of viral or host origin.[35] Yet, alternative contact points for MHC class I molecules and bacterial peptides exist. Cross-presentation by noninfected antigen-presenting cells (APCs) of antigens engulfed within apoptotic blebs from infected cells represents a critical pathway to induce CD8 T cells by phagosomal bacteria (see below). One major advantage of CD8 T cells over CD4 T cells is their recognition of antigen bound by MHC class I gene products, which are expressed by almost all host cells. Thus CD8 T cells recognize professional and nonprofessional phagocytes equally well.

Unconventional T Cells

The relevance of the γδ T cells to antibacterial immunity is incompletely understood. Several studies indicate that γδ T cells are rapid producers of IL-17 at sites of bacterial implantation. Transient participation of γδ T cells in protection and a unique requirement for γδ T cells in granuloma formation have been described for murine listeriosis and TB. Murine γδ T cells appear to recognize peptides presented by nonpolymorphic MHC class I molecules, whereas human γδ T cells respond to nonpeptidic phosphorylated metabolites, notably from the isoprenoid pathway of bacterial and host origin.[43]

MAIT cells are primarily localized at mucosal sites, and current evidence suggests that they play a role in control of bacterial infections in mucosal tissues, such as lung (*M. tuberculosis*) and gut (gram-negative bacteria) tissues. Antigenic ligands include derivatives of vitamin B_2 (riboflavin), produced by many intracellular bacteria, including salmonellae and mycobacteria.[44-46]

CD1 comprises a group of nonpolymorphic MHC-related molecules that can present glycolipid antigens to unconventional T cells. In humans, group 1 CD1-restricted T cells respond to a variety of microbial glycolipids, including LAM, PIMs, mycolic acids, sulfatides, sulfoglycolipids, and lipopeptides.[47] Group I CD1 molecules are absent in mice. The group II CD1 molecule CD1d is present in both humans and mice and controls development of natural killer T cells (NKT cells) that express the NK cell marker NK1.1. Upon antigen activation, these T cells rapidly produce cytokines and are capable of producing both IL-4 and IFN-γ. Bacterial antigens recognized by NKT cells, PIMs from mycobacteria, and glycosphingolipids from *Ehrlichia* and *Sphingomonas* spp. have been identified. NKT cells also respond to host endogenous lysosomal lipids loaded onto CD1d.[48]

In summary, unconventional T cells often recognize nonpeptidic ligands of bacterial origin, emphasizing that they play a particular role in immunity against bacteria, including intracellular bacteria. Because of the highly skewed T cell receptor, these T cells are specific for a limited variety of bacterial ligands. Because of their less demanding antigen recognition and activation requirements, unconventional T cells may fill a gap between prompt innate resistance and the delayed conventional T-cell response.

T-Cell Memory and Regulation of Immune Responses

Long-term protective immunity against infectious agents relies on immune memory, which forms the basis for the success of all vaccines. Memory T cells can be divided into central memory T (T_{CM}) cells and effector memory T (T_{EM}) cells, based on differential surface phenotypic and tissue migration patterns.[49] T_{EM} cells accumulate in peripheral tissues where they express effector functions, whereas T_{CM} cells persist in lymph nodes, where they rapidly develop into T_{EM} cells after secondary antigen encounter. The tissue-resident memory T (T_{RM}) cells are localized to mucosal sites, where they provide efficient protection against invading pathogens.[50] Although it is generally accepted that memory T cells can survive in the absence of persistent antigen, little is known about the induction and maintenance of long-lasting T-cell memory in chronic infections with intracellular bacteria.[51]

B Cells

Although antibodies produced by B cells appear to play a minor biological role in infections with intracellular bacteria, many facultative intracellular bacteria spend some time outside their host cells, where they are accessible to antibodies.[52] Thus antibodies contribute to protection against salmonellae. Besides antibody production, B cells are potent APCs for soluble antigens, including lipids presented by CD1c, and secrete many cytokines otherwise associated with T cells, DCs, and macrophages. B-cell signaling via MyD88 during *S. typhimurium* infection has been associated with B-cell production of IL-10, and mice with B-cell-specific deficiency in MyD88 were more resistant to infection demonstrating that, like T cells, B cells can perform regulatory functions.[53]

KEY CONCEPTS

How Might a Vaccine Work?

- Activation of innate immunity for instruction of appropriate acquired immune responses (PRRs)
- Activation of the appropriate array of T-cell populations
- T-cell secretion of appropriate cytokine combination
- Efficient development of T-cell memory
- Efficient activation of antibacterial effector mechanisms
- Best case scenario: sterile pathogen eradication
- Second best case scenario: driving the infection "deeper" into latency, thus efficiently preventing reactivation of disease

Regulatory T Cells

Intracellular bacteria can cause detrimental inflammation and tissue damage, for example, as a result of IFN-γ- and TNF-α-mediated Th1 responses. Under normal circumstances, control mechanisms are in place to limit immunopathology. Such countermeasures are elicited as part of the ongoing immune response during infection. The main cytokines that limit inflammation and control IFN-γ production are IL-10 and TGF-β. Although macrophages and DCs produce these cytokines, their main producers are Tregs, which are the prime cells involved in immune regulation. Natural Tregs are responsive to IL-2 because of high constitutive CD25 expression and are characterized by expression of the transcription factor FOXP3.[54] Expansion of Tregs appears to be both antigen dependent and antigen independent. In addition, Tregs selectively express TLRs and can be activated, for example, by LPS and possibly other TLR ligands. This makes their immediate activation during bacterial infection a probable scenario. Although Tregs limit CD8 T-cell responses in experimental listeriosis, their general role in infections with intracellular bacteria has not been fully elucidated. Suppression of T-cell responses and anergy have been clinically documented for TB and leprosy. Although Treg functions can limit detrimental T-cell responses and immunopathology, they can also prevent elimination of bacteria, and hence are likely a key factor in promoting the persistent chronic state of infection with intracellular bacteria.

CONCLUDING REMARKS

ON THE HORIZON

- Design of biomarkers to discriminate latent infection from active tuberculosis (TB) and to predict risk of progression to active TB
- Development and clinical testing of new drugs for multidrug-resistant/extensively drug-resistant (MDR/XDR) TB and dormant *M. tuberculosis*
- Developing host-directed therapy (HDT) in adjunct to conventional drugs, notably for highly drug-resistant pathogens
- Development and clinical testing of new vaccines that protect against pulmonary TB in adults
- Reduction of the unequal burden of TB in developing and industrialized countries by public health measures, education, and socioeconomic advances

Our deepening understanding of the molecular events governing intracellular bacterial infections is allowing development of novel therapeutic and preventive approaches. The need for such interventions is becoming all the more pronounced in the face of increasing levels of antibiotic resistance of bacteria, such as *M. tuberculosis,* which render canonical drugs that specifically target bacterial molecular processes ineffective. A new approach, termed *host-directed therapy* (HDT) aims to develop new drugs or repurpose previously approved ones that are directed at host molecular processes.[55] Such approaches include monoclonal antibodies (mAbs) to neutralize cytokines, such as TNF-α or IL-6, to abrogate tissue destructive inflammation, repurposed use of licensed drugs, such as ibuprofen and verapamil, to modulate inflammation and enhance antibiotic effectiveness, respectively, and use of immunostimulatory natural molecules, such as vitamin D_3 to enhance bacterial killing by xenophagy. These approaches range from preclinical and early-stage clinical development to late-stage clinical development and offer promise to shorten the traditional duration of therapy.

It is being increasingly recognized that the interaction between intracellular bacteria and the immune system is not of the "all or nothing" type but is instead a "continuous struggle." This realization has far-reaching implications for preventive and therapeutic strategies against intracellular bacterial infections. First, vaccination against intracellular bacteria has not yet been effected satisfactorily because of the involvement of several distinct T-cell subsets with different modes of stimulation and activity profiles. Second, chemotherapy has frequently proven to be suboptimal for the sterile eradication of bacteria hidden in cellular niches. A better understanding of the complex crosstalk between cytokines, T lymphocytes, macrophages, and infected host cells will no doubt directly promote the development of improved control measures.[55]

ACKNOWLEDGMENTS

We are grateful to Mary Louise Grossman for excellent assistance and Diane Schad for the figures.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Kaufmann SHE, Rouse BT, Sacks D. The immune response to infectious diseases. Washington, D.C.: ASM Press; 2011.
2. Kaufmann SHE, van Helden PD. Handbook of tuberculosis: clinics diagnostics therapy epidemiology. 1st ed. Weinheim, Germany: Wiley-VCH; 2008.
3. WHO. Global tuberculosis report 2016. Geneva, Switzerland: WHO Press; 2016.
4. White C, Franco-Paredes C. Leprosy in the 21st century. Clin Microbiol Rev 2015;28:80–94.
5. Griffith DE, Aksamit T, Brown-Elliott BA, et al. An official ATS/IDSA statement: diagnosis, treatment, and prevention of nontuberculous mycobacterial diseases. Am J Respir Crit Care Med 2007;175:367–416.
6. Sarfo FS, Phillips R, Wansbrough-Jones M, et al. Recent advances: role of mycolactone in the pathogenesis and monitoring of *Mycobacterium ulcerans* infection/Buruli ulcer disease. Cell Microbiol 2016;18:17–29.
7. WHO. Global Health Observatory (GHO). Data: Buruli ulcer. Geneva, Switzerland: WHO; 2016.
8. Crump JA, Mintz ED. Global trends in typhoid and paratyphoid Fever. Clin Infect Dis 2010;50:241–6.

9. Pappas G, Papadimitriou P, Akritidis N, et al. The new global map of human brucellosis. Lancet Infect Dis 2006;6:91–9.
10. Pathela P, Blank S, Schillinger JA. *Lymphogranuloma venereum:* old pathogen, new story. Curr Infect Dis Rep 2007;9:143–50.
11. Limmathurotsakul D, Golding N, Dance DA, et al. Predicted global distribution of and burden of melioidosis. Nat Microbiol 2016;1: 15008.
12. Sjostedt A. Intracellular survival mechanisms of *Francisella tularensis,* a stealth pathogen. Microbes Infect 2006;8:561–7.
13. Dooling KL, Toews KA, Hicks LA, et al. Active Bacterial Core Surveillance for Legionellosis - United States, 2011-2013. MMWR Morb Mortal Wkly Rep 2015;64:1190–3.
14. Land JA, Van Bergen JE, Morre SA, et al. Epidemiology of *Chlamydia trachomatis* infection in women and the cost-effectiveness of screening. Hum Reprod Update 2010;16:189–204.
15. Wright HR, Turner A, Taylor HR. Trachoma. Lancet 2008;371: 1945–54.
16. Parola P, Paddock CD, Raoult D. Tick-borne rickettsioses around the world: emerging diseases challenging old concepts. Clin Microbiol Rev 2005;18:719–56.
17. Winslow GM, Bitsaktsis C. Immunity to the ehrlichiae: new tools and recent developments. Curr Opin Infect Dis 2005;18:217–21.
18. Biswas S, Rolain JM. Bartonella infection: treatment and drug resistance. Future Microbiol 2010;5:1719–31.
19. Dorhoi A, Reece ST, Kaufmann SH. For better or for worse: the immune response against *Mycobacterium tuberculosis* balances pathology and protection. Immunol Rev 2011;240:235–51.
20. Kaufmann SHE, Dorhoi A. Molecular determinants in phagocyte-bacteria interactions. Immunity 2016;44:476–91.
21. Yin Q, Fu TM, Li J, et al. Structural biology of innate immunity. Annu Rev Immunol 2015;33:393–416.
22. Taylor PR, Martinez-Pomares L, Stacey M, et al. Macrophage receptors and immune recognition. Annu Rev Immunol 2005;23:901–44.
23. Netea MG. Training innate immunity: the changing concept of immunological memory in innate host defence. Eur J Clin Invest 2013;43:881–4.
24. Le Y, Zhou Y, Iribarren P, et al. Chemokines and chemokine receptors: their manifold roles in homeostasis and disease. Cell Mol Immunol 2004;1:95–104.
25. Cambier CJ, Takaki KK, Larson RP, et al. Mycobacteria manipulate macrophage recruitment through coordinated use of membrane lipids. Nature 2014;505:218–22.
26. Werner E. GTPases and reactive oxygen species: switches for killing and signaling. J Cell Sci 2004;117:143–53.
27. Bogdan C. Nitric oxide and the immune response. Nat Immunol 2001;2:907–16.
28. Lieberman J. The ABCs of granule-mediated cytotoxicity: new weapons in the arsenal. Nat Rev Immunol 2003;3:361–70.
29. Rudel T, Kepp O, Kozjak-Pavlovic V. Interactions between bacterial pathogens and mitochondrial cell death pathways. Nat Rev Microbiol 2010;8:693–705.
30. Brumell JH, Scidmore MA. Manipulation of rab GTPase function by intracellular bacterial pathogens. Microbiol Mol Biol Rev 2007;71: 636–52.
31. Salcedo SP, Holden DW. Bacterial interactions with the eukaryotic secretory pathway. Curr Opin Microbiol 2005;8:92–8.
32. Gabrilovich DI, Nagaraj S. Myeloid-derived suppressor cells as regulators of the immune system. Nat Rev Immunol 2009;9:162–74.
33. Knaul JK, Jorg S, Oberbeck-Mueller D, et al. Lung-residing myeloid-derived suppressors display dual functionality in murine pulmonary tuberculosis. Am J Respir Crit Care Med 2014;190:1053–66.
34. Masaki T, Qu J, Cholewa-Waclaw J, et al. Reprogramming adult Schwann cells to stem cell-like cells by leprosy bacilli promotes dissemination of infection. Cell 2013;152:51–67.
35. Stavru F, Archambaud C, Cossart P. Cell biology and immunology of *Listeria monocytogenes* infections: novel insights. Immunol Rev 2011;240:160–84.
36. Cooper AM. Cell-mediated immune responses in tuberculosis. Annu Rev Immunol 2009;27:393–422.
37. van der Wel N, Hava D, Houben D, et al. *M. tuberculosis* and *M. leprae* translocate from the phagolysosome to the cytosol in myeloid cells. Cell 2007;129:1287–98.
38. Iwakura Y, Ishigame H, Saijo S, et al. Functional specialization of interleukin-17 family members. Immunity 2011;34:149–62.
39. Lin Y, Ritchea S, Logar A, et al. Interleukin-17 is required for T helper 1 cell immunity and host resistance to the intracellular pathogen *Francisella tularensis*. Immunity 2009;31:799–810.
40. O'Shea JJ, Paul WE. Mechanisms underlying lineage commitment and plasticity of helper CD4+ T cells. Science 2010;327:1098–102.
41. Cox MA, Harrington LE, Zajac AJ. Cytokines and the inception of CD8 T cell responses. Trends Immunol 2011;32:180–6.
42. Walch M, Dotiwala F, Mulik S, et al. Cytotoxic cells kill intracellular bacteria through granulysin-mediated delivery of granzymes. Cell 2014;157:1309–23.
43. Bonneville M, O'Brien RL, Born WK. Gammadelta T cell effector functions: a blend of innate programming and acquired plasticity. Nat Rev Immunol 2010;10:467–78.
44. Ussher JE, Klenerman P, Willberg CB. Mucosal-associated invariant T-cells: new players in anti-bacterial immunity. Front Immunol 2014;5:450.
45. Napier RJ, Adams EJ, Gold MC, et al. The role of mucosal associated invariant T cells in antimicrobial immunity. Front Immunol 2015;6:344.
46. Corbett AJ, Eckle SB, Birkinshaw RW, et al. T-cell activation by transitory neo-antigens derived from distinct microbial pathways. Nature 2014;509: 361–5.
47. Brigl M, Brenner MB. CD1: antigen presentation and T cell function. Annu Rev Immunol 2004;22:817–90.
48. Darmoise A, Teneberg S, Bouzonville L, et al. Lysosomal alpha-galactosidase controls the generation of self lipid antigens for natural killer T cells. Immunity 2010;33:216–28.
49. Lanzavecchia A, Sallusto F. Understanding the generation and function of memory T cell subsets. Curr Opin Immunol 2005;17:326–32.
50. Mueller SN, Gebhardt T, Carbone FR, et al. Memory T cell subsets, migration patterns, and tissue residence. Annu Rev Immunol 2013;31: 137–61.
51. Scott P. Immunologic memory in cutaneous leishmaniasis. Cell Microbiol 2005;7:1707–13.
52. Joller N, Weber SS, Oxenius A. Antibody-Fc receptor interactions in protection against intracellular pathogens. Eur J Immunol 2011;41: 889–97.
53. Neves P, Lampropoulou V, Calderon-Gomez E, et al. Signaling via the MyD88 adaptor protein in B cells suppresses protective immunity during *Salmonella typhimurium* infection. Immunity 2010;33:777–90.
54. Belkaid Y, Rouse BT. Natural regulatory T cells in infectious disease. Nat Immunol 2005;6:353–60.
55. Kaufmann SH, Lange C, Rao M, et al. Progress in tuberculosis vaccine development and host-directed therapies-a state of the art review. Lancet Respir Med 2014;2:301–20.

MULTIPLE-CHOICE QUESTIONS

1. Which disease caused by intracellular bacterial infection results in the HIGHEST level of global mortality?
   A. Typhoid
   B. Tuberculosis
   C. Legionnaires' disease
2. Which host cytokine molecule is the MOST important in augmenting host cell defense mechanisms against intracellular bacteria?
   A. Interleukin (IL)-4
   B. IL-17
   C. Interferon (IFN)-γ
3. Which of these host cell processes is **NOT** harnessed for protection against intracellular bacterial infection?
   A. Autophagy
   B. Necrosis
   C. Apoptosis
4. Which bacterial-derived ligand is presented to T cells via major histocompatibility complex (MHC) class I?
   A. Lipopolysaccharide
   B. Polypeptide
   C. RNA

27

Host Defenses to Extracellular Bacteria

Marcos C. Schechter, Sarah W. Satola, David S. Stephens

The human host has developed protective mechanisms to interact with the multitude of bacterial species encountered in nature. These host defenses include nonspecific mechanisms of clearance, as well as innate and specific adaptive immune responses. Partly because of these mechanisms, the vast majority of bacterial species do not cause human disease. Many bacterial species have established symbiotic or commensal relationships with the human host and colonize skin and mucosal surfaces. These commensals are generally of low virulence except in individuals whose host defenses are compromised. Given the diversity of the microbial world, a relatively few pathogenic bacterial species or subpopulations of those species have evolved virulence factors or strategies that can overcome or circumvent intact human host defense mechanisms to cause localized or systemic disease.

Important bacterial pathogens of clinical importance reside mostly extracellularly (Table 27.1). Examples are bacterial pathogens typically labeled *extracellular*, such as *Streptococcus pneumoniae*, *Streptococcus pyogenes*, *Haemophilus influenzae*, *Neisseria meningitidis*, *Neisseria gonorrhoeae*, and *Bordetella pertussis*, which are transmitted from one individual to another by close contact. Other "extracellular" bacterial pathogens, such as *Clostridium* spp. *Vibrio cholerae*, *Shigella dysenteriae*, enteropathogenic *Escherichia coli*, and *Bacillus anthracis*, are transmitted through food, water, animal, or other environmental contact. *Staphylococcus aureus* is an important extracellular pathogen for humans and can be acquired from other humans, from animals, or through environmental contact. Acquisition of these pathogenic bacteria may be transient, result in variable intervals of colonization, or rapidly cause localized or systemic disease.

Extracellular bacterial pathogens can produce acute inflammatory and purulent infectious diseases, such as meningitis, septicemia, pneumonia, urethritis, pharyngitis, inflammatory diarrhea, cellulitis, and abscesses and/or produce disease by the release of toxins. Disease associated with some extracellular bacteria (*e.g., Helicobacter pylori*) results from chronic colonization. Susceptibility to extracellular bacterial pathogens is enhanced by hereditary, acquired, or age-related defects in innate or adaptive host defenses. Resistance to extracellular bacterial pathogens or their toxins can be accentuated by chemoprophylaxis, by vaccines, and by other immune modulation processes (*e.g.,* passive immune globulin administration). Caution is urged in the interpretation of the term "extracellular." The classification of bacteria as "extracellular" and "intracellular" is primarily based on observations *in vitro* and has been challenged by some authors, as some "extracellular" bacterial species invade host cells as a part of their normal lifecycle and during steps in the disease process (*e.g., S. aureus, S. pneumoniae, S. pyogenes, N. meningitidis*).[1] Conversely, bacteria typically classified as "intracellular" can have an extracellular component to their lifecycle (*e.g., Mycobacterium tuberculosis* in cavitary lesions).

Host defense mechanisms against extracellular bacteria are a continuum. The innate and adaptive immune systems cooperate to protect the host from extracellular bacterial infections. The innate immune system senses bacteria through pattern recognition receptors (PRRs; Chapter 3). These receptors activate antimicrobial defenses and stimulate the adaptive response, all while balancing excessive immune and inflammatory responses with the need to protect against the infecting pathogen.[2]

CLINICAL PEARLS

Distinguishing Clinical Characteristics of Infections With Extracellular Bacteria

- Sterilizing immunity
- Colonization of mucosal surfaces often precedes disease
- Causes of pyogenic infections
- T-helper (Th)17 response critical in generating a neutrophilic response
- Antibodies are protective for some of the major pathogens
- Effective vaccines available for many of the major pathogens

CLEARANCE AND NONSPECIFIC HOST DEFENSES AT MUCOSAL EPITHELIAL SURFACES

Bacteria first encounter a physical barrier, which comprises skin, mucus and mucosal surfaces, and the normal microbiota, as well as nonspecific factors, such as nutrient limitation (*e.g.,* iron), and antimicrobial proteins or peptides. Intact skin and mucosal surfaces provide complex chemical and biological obstacles to extracellular bacteria and are an important line of defense preventing the invasion of these pathogens and their products.[3] Humans commonly carry extracellular pathogens asymptomatically on skin and at respiratory and gastrointestinal (GI) mucosal surfaces. To cause disease, pathogens breach or disrupt epithelial barriers. Damage to epithelial barriers as a result of trauma; coinfections; drugs, such as those used in chemotherapy; environmental factors, such as smoking, allergies, or low humidity; and catheterization and intubation circumvent these barriers and allow bacteria access to subcutaneous tissues, blood vessels, and other normally sterile sites. Additionally, acquired and genetic diseases affecting epithelial barriers are also linked to increase in infections.

The human epithelium has evolved to prevent colonization and invasion by pathogenic organisms.[3] Skin is a relatively dry, acidic (pH 5–6) barrier that contains growth-inhibiting fatty acids and antimicrobial peptides (AMPs) (see below), characteristics that

TABLE 27.1 Examples of Clinically Relevant Pathogenic Extracellular Bacteria

Species	Examples of Human Disease	Selected Mechanisms of Pathogenesis	Special Features Key to Host Infection	Examples of Susceptible Populations/Risk Factors
Staphylococcus aureus	Cellulitis, abscesses, bacteremia, endocarditis, toxic shock syndrome, osteomyelitis, pneumonia, wound infections	Protein A: promotes fibronectin binding PVL: cytotoxic α-toxin: membrane damage TSST-1: superantigen	Asymptomatic colonization, resistant to dehydration	Injection drug users, patients on hemodialysis, defects on Th17 response (Job's syndrome), Surgical procedures and skin trauma
Streptococcus pneumoniae (*Pneumococcus*)	Pneumonia, otitis media, meningitis	Capsule: prevents phagocytosis, antigenic variation Pneumolysin: cytotoxic PspA & C: inhibition of complement Neuraminidase, hyaluronidase: spread and colonization IgA_1 protease	Asymptomatic colonization, readily acquires new genes through transformation	Smokers, cerebrospinal fluid leak, asplenia, hypogammaglobulinemia, human immunodeficiency virus/acquired immunodeficiency syndrome (HIV/AIDS), unvaccinated children
Streptococcus pyogenes (group A *Streptococcus*)	Pharyngitis, cellulitis, erysipelas, toxic shock syndrome, necrotizing fasciitis, scarlet fever, rheumatic fever	Hyaluronic acid capsule, M-protein: prevents phagocytosis Streptolysin O & S: cytotoxic Streptococcal pyrogenic exotoxins C5a peptidase	High diversity of M-proteins, molecular mimicry of human antigens	School-age children, crowded-conditions (*e.g.,* military barracks), injury to lymphatic system (*e.g.,* surgical harvest of saphenous vein)
Streptococcus agalactiae (group B *Streptococcus*)	Neonatal sepsis, pneumonia and meningitis, perinatal infections, bacteremia	FbsA: fibrinogen receptor, promotes adherence Capsule β-hemolysin C5a peptidase β protein: downregulates complement	Asymptomatic colonization, acquisition by infants during birth	Neonates and infants (immunity dependent on passive transfer of maternal antibodies), diabetes mellitus
Neisseria meningitidis (*Meningococcus*)	Meningitis, bacteremia (purpura fulminans)	Capsular polysaccharide: promotes adherence and prevents phagocytosis Type IV pili: promote attachment to host cells LOS: analogues to LPS, activates TLR4 pathway IgA_1 protease fHbp: downregulates the host alternative complement pathway	Molecular mimicry of human antigens, phase and antigenic variation, asymptomatic carriage	Terminal complement deficiencies, hypogammaglobulinemia
Neisseria gonorrhoeae (*Gonococcus*)	Urogenital infections, disseminated gonococcal infection, pharyngitis	Type IV pili: promote attachment to host cells Opa protein adhesion IgA_1 protease LOS: analogs to LPS, activates TLR4 pathway	Phase and antigenic variation, molecular mimicry of human antigens	Terminal complement deficiencies, women during menstrual period (increases risk of dissemination)
Escherichia coli	Urinary tract infections, gastroenteritis, sepsis, neonatal meningitis	Capsular polysaccharide Tissue specific fimbriae Heat-labile enterotoxins: increases intestinal chloride secretion LPS: activation of TLR4	Antigenic heterogeneity of LPS and capsule	Bladder instrumentation, pregnancy
Pseudomonas aeruginosa	Ventilator-associated pneumonia, bronchiectasis	Pili and flagella: attachment to the host and formation of biofilms LPS Exotoxin A Lipases, lecithinases, elastase	Considerable adaptability to changes in environment, large genome size, biofilms	Orotracheal intubation, cystic fibrosis
Clostridium difficile	Colitis	Toxin B: cytotoxic Flagella	Endospore formation, asymptomatic carriage	Antibiotics and other disruptions of the microbiota
Haemophilus influenzae	Otitis media, pneumonia, epiglottitis, bacteremia, meningitis	LPS with phosphorylcholine Pili: adherence Capsule High-molecular weight adhesins IgA_1 protease	Phase variation of pili, asymptomatic carriage	Unvaccinated children, immunocompromised, sickle cell disease, smoking
Helicobacter pylori	Peptic ulcer disease	Urease: colonization of gastric mucosa Flagella: motile in gastric mucus CagA; bacteria-derived carcinogen	Polymorphism of CagA	Crowded living conditions, unreliable source of clean water, living with someone who has *H. pylori*
Bordetella pertussis	Whopping cough (children), chronic cough (adults)	Pertussis toxin: inhibits neutrophils, macrophages, lymphocytes Pertactin & filamentous hemagglutinin; mediates attachment	Antigenic variation of adhesins	Nonvaccinated adults and children, infants who have not completed vaccine series, adults and adolescents whose immunity has diminished

TSST-1, toxic shock syndrome toxin 1; Psp, pneumococcal surface protein; LOS, lipooligosaccharide; fHbp, Factor H-binding protein; LPS, lipopolysaccharide; CapA, cytotoxin-associated gene A; Ig, immunoglobulin; TLR, Toll-like receptor; PVL, Panton-Valentine leukocidin.

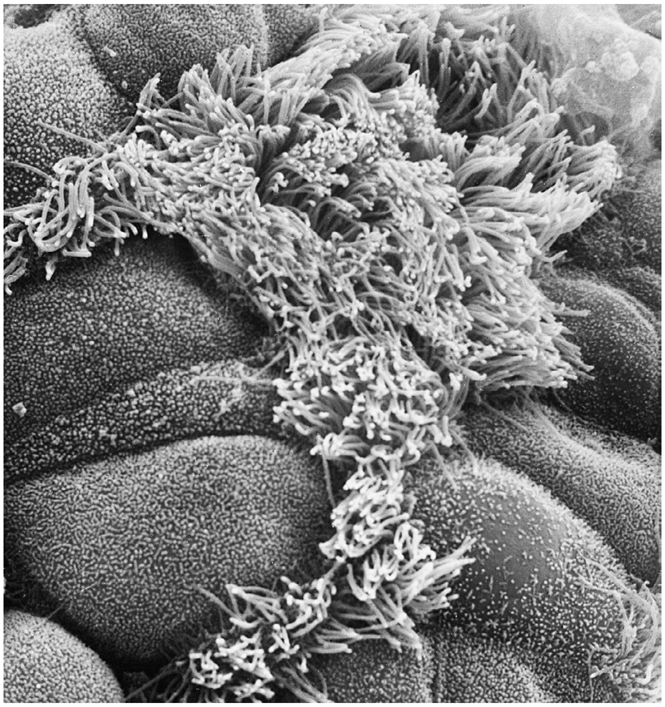


FIG 27.1 **Mucociliary Host Defense.** Scanning electron micrograph of human upper respiratory mucosa showing the ciliated and nonciliated epithelial surface (×16000).[41]

are detrimental to many bacteria. The constant desquamation of stratified epithelial surface of skin helps in the removal of microorganisms.[4] Nevertheless, a complex skin microbiome that can include bacterial pathogens has been identified, and this differs remarkably, depending on location in the body. Disruption of these physical barriers can augment pathogen tissue colonization and invasion. Infections by *S. aureus* and *S. pyogenes*, bacteria that can colonize skin, are often preceded by skin damage. Repeated trauma to skin (*e.g.*, dialysis and intravenous drug use) also enhances skin colonization with pathogens, including that by *S. aureus*.

Mucosal surfaces have additional nonspecific antibacterial defenses. The mucociliary blanket of the respiratory tract (Fig. 27.1) and the female urogenital tract (fallopian tube) move bacteria away from epithelial surfaces, as does the flushing of the urinary tract with urine, intestinal peristalsis, and the bathing of the conjunctiva with tears. Lysozyme is found in most mucosal secretions and lyses bacterial cell walls by splitting muramic acid β(1–4)-*N*-acetyl-glucosamine linkages. The acid pH of the stomach, intestinal peristalsis, and the antibacterial effect of proteolytic enzymes present in intestinal secretions are important GI tract host defenses against many pathogenic bacteria. The GI mucosa has a layer of mucus that acts as a physical shield to bacteria. Mucus is rich in mucin, glycoproteins that limit pathogen binding to other host molecules necessary for mucosal adhesion. Additionally, the mucus layer may function more than as a physical barrier by acting as a diffusion barrier to concentrate antimicrobial proteins at the appropriate epithelial cell surface. The glycocalyx, an extracellular layer of the apical surface of mucosal cells composed of carbohydrates, also protects cells against bacterial attachment.

Bacterial attachment and colonization of mucosal surfaces can be inhibited by bacterial binding to human cellular antigens present in secretions, such as ABO blood group antigens. Cell adhesion and extracellular matrix molecules, such as fibronectin and proteoglycans, can also inhibit or enhance bacterial binding to epithelial surfaces. The Tamm-Horsfall glycoprotein, found in urine, can bind avidly to a variety of bacteria and facilitate clearance. Proteins, such as lactoferrin (Lf), present at mucosal surfaces, bind iron, an important requirement for bacterial growth. This action may reduce microbial proliferation, but some mucosal pathogens bind Lf and remove iron from the molecule for growth.

Normal Microbiota as Host Defense

The human microbiome is now recognized as a major host defense against bacterial pathogens by providing “colonization resistance,” maintaining a balance of commensals to pathogens, and by priming the immune system (Chapter 14).[5] Altering or disrupting the normal microbiota by antibiotics facilitates the expansion of enteric pathogens as *Clostridium difficile* and *Salmonella typhimurium* or selection of antibiotic-resistant members of the microbiome. Similarly, changes in human physiology, for example, exposure of skin to elevated temperatures and humidity, chronic stress, host immune suppression, or active behavioral changes, such as smoking, can cause a commensal-to-pathogen switch. Recent studies have demonstrated that certain resident microbiota can resist pathogen colonization and infection. For example, matched volunteers were inoculated with *Haemophilus ducreyi* into the arms, and the subsequent infection either resolved or resulted in formation of abscesses; characterization of the skin microbiome before, during, and after the experimental inoculation showed that the microbiomes of those with pustule formation and of those with resolved infection were distinct and influenced the course of the *H. ducreyi* infection.[6]

The interaction of the microbiome with the immune system is also important for defense against extracellular pathogens. Normal microbiota facilitate a high level of priming of the immune system by maintaining high levels of major histocompatibility complex (MHC) class II molecule expression on macrophages and other antigen-presenting cells (APCs). PRRs (see below) are traditionally known to recognize microbial molecules during infection; however, ligands for PRRs are abundantly produced by the resident microbiota during normal colonization. The integrity of the intestinal epithelial layer is dependent on activation of Toll-like receptors (TLRs; see below) by normal microbiota.[4] Stimulation of TLR-5 has been shown to increase resistance to *E. faecium* infection in a murine model. Activation of nucleotide oligomerization domain 1 (NOD1) receptors by gut resident microbiota is necessary for priming of the innate immune system. Additionally, resident microbiota produce such factors as bacteriocins, lantibiotics, and phenol-soluble modulin (PSM), which function in a similar manner to that of host-derived antimicrobial proteins and peptides (APPs; see below), suggesting an important host defense strategy against pathogen colonization.[7] Importantly, members of the resident microbiota can cause disease, particularly with loss of epithelial integrity and translocation to a different host tissue.

Antimicrobial Peptides and Antimicrobial Proteins

Pathogens colonizing or invading epithelial surfaces are confronted with APPs, which can be produced both by the host and the microbiota (Table 27.2).[7] In addition to pathogen killing, APPs control host physiological functions, such as inflammation, angiogenesis, and wound healing. They also limit pathogen colonization and shape the composition of the host microbiome

TABLE 27.2 Antimicrobial Proteins (AMPs) Against Extracellular Bacteria

AMP	Tissue/Cell Sources	Mechanism of Action	Target Organisms
α-defensins	Small intestines, Paneth cells	Membrane disruption; inhibits complement activation; chemoattracts dendritic cells	Gram-positive bacteria Gram-negative bacteria
β-defensins	Large intestines, skin, respiratory tract epithelial cells	Membrane disruption; lipid II binding	Gram-positive bacteria Gram-negative bacteria
Cathelicidin (LL37)	Large intestine, skin, lung, urogenital tract	Membrane disruption	Gram-positive bacteria Gram-negative bacteria
RNases	Skin, intestine, respiratory epithelia, placenta	Unknown	Gram-positive bacteria Gram-negative bacteria
Psoriasin (S100A7)	Skin, urogenital tract	Unknown	*Escherichia coli*
Calprotectin (S100A8-A9)	Abscesses/neutrophils	Metal Chelation	*Staphylococcus aureus*
C-type lectins	Small intestines	Peptidoglycan recognition	Gram-positive bacteria
Bactericidal/permeability-increasing protein (BPI)	Neutrophils, epithelial cells	Neutralizes lipopolysaccharide	Gram-negative bacteria
Lysozyme	Skin, body fluids, tears, intestinal Paneth cells	Degrades peptidoglycan	Gram-positive bacteria Some gram-negative bacteria activity
Dermcidin	Sweat glands	Membrane disruption	Gram-positive bacteria Gram-negative bacteria
Peptidoglycan recognition proteins	Liver, intestines, skin, neutrophils	Activates bacterial two-component systems; targets peptidoglycan	Especially active against gram-positive bacteria Gram-negative bacteria
Phospholipase A2	Tears, intestines	Hydrolysis of bacterial phospholipids	Gram-positive bacteria

TABLE 27.3 Immune Microbial Pattern Recognition Molecules

Toll-like receptors (TLRs)
Nucleotide-binding oligomerization domain (NOD), caterpillar proteins, peptidoglycan recognition proteins (PGRPs)
RNA helicases/PkR
Complement proteins: C1q, C1 inhibitor
Antimicrobial peptides
Collectins and surfactants
C- and S-lectins: mannose-binding lectin, L-ficolin

and are both constitutively present and inducible by infection and injury.[7] Humans produce two main classes of APPs: defensins and cathelicidins. Defensins, for example, are expressed in skin, intestines, and the respiratory tract and have activity against gram-positive and gram-negative bacteria. Interestingly, APPs expression can be altered by epithelial injury. Keratinocytes of inflamed psoriatic lesions produce increased levels of certain APPs, and patients with such lesions rarely have secondary bacterial infections. In contrast, keratinocytes from patients with atopic dermatitis have dampened APP production in response to inflammation, and colonization and superinfections of the skin by *S. aureus* are common. Not surprisingly, successful pathogens have developed several mechanisms to counteract APPs. Gonococci, for example, counteract APPs with efflux pumps.

RECOGNITION OF EXTRACELLULAR BACTERIA AND ACTIVATION OF THE IMMUNE SYSTEM

Immune pattern recognition molecules (Table 27.3) are a major arm of the innate immune system and are released or expressed by a range of host cells, including lymphocytes, macrophages, or tissue histiocytes, dendritic cells (DCs), polymorphonuclear leukocytes or neutrophils (PMNs), and epithelial cells. The discovery and characterization of specific pattern recognition molecules (see below) has revolutionized our understanding of the initial specific events occurring between microbes and human cells. The identification and function of these molecules is rapidly expanding and include TLRs, NOD and caterpillar proteins, RNA helicases, complement proteins, antimicrobial peptides, collectins and surfactants, C-lectins, and S-lectins, such as mannose-binding lectin and L-ficolin.

Pattern Recognition Receptors

Innate immune recognition relies on the detection of unique molecular structures found on microorganisms by host PRRs.[8,9] TLRs and NOD-like receptors (NLRs) are the best studied PRRs (Table 27.4). TLRs (TLR1–11) (Chapter 3) are found on macrophages, neutrophils, and other host cells. These receptors recognize a variety of microbial ligands or pathogen-associated molecular patterns (PAMPs), including lipoproteins, lipopolysaccharide (LPS), flagellin, and nucleic acids produced by gram-negative and/or gram-positive bacteria. Alterations (polymorphisms) in TLRs (*e.g.,* TLR4) and other pattern recognition molecules are associated with susceptibility or severity of specific infections (*e.g.,* sepsis).[10] TLR expression can be regulated by type I interferons (IFNs) and by microRNAs (miRNAs), and the dysregulation of TLRs can be involved in acute and chronic inflammatory diseases and cancer.[11]

The NLRs are a family of intracellular receptors, some of which function as PRRs.[8] NOD1 and NOD2 are well characterized as PRRs to extracellular pathogens, such as *S. flexneri.* Importantly, in concert with TLR signaling, NLR can respond to a variety of PAMPs by forming the inflammasome complex. Inflammasome activation generates interleukin-18 (IL-18) and the active form of IL-1, an important step in the immune response to many bacteria. In addition to PAMPs, danger-associated molecular patterns (DAMPs), typically associated with activation of the innate immune system upon injury by noninfectious mechanisms, can be released both by the host and bacteria and result in amplification of the inflammatory response.

TABLE 27.4 Pattern Recognition Receptors (PRRs) Recognize Pathogen-Associated Molecular Patterns (PAMPs) From Various Bacteria

PRRs	PAMPs	Microbes
Toll-like receptor (TLR)2/1	Triacyl lipoproteins	Bacteria
TLR2/6	Diacyl lipoproteins	Mycoplasma
	Lipoteichoic acid	Gram-positive bacteria
TLR2	Peptidoglycan	Gram-positive bacteria
	Porins	Bacteria (*Neisseria*)
TLR4	Lipopolysaccharide (LPS)	Gram-negative bacteria
TLR5	Flagellin	Flagellated bacteria (*Helicobacter pylori, Salmonella*)
TLR7/8	RNA	Group B *Streptococcus*
TLR9	CpG-DNA	Bacteria (*Salmonella*)
	DNA	Bacteria (*Staphylococcus* at low MOI)
TLR11	Not determined	Uropathogenic bacteria
Nucleotide-binding oligomerization domain (NOD)1	Meso-diaminopimelic acid	*H. pylori, Bacillus* spp., *Campylobacter jejuni, Pseudomonas aeruginosa*
NOD2	Muramyl dipeptide (MDP)	*Streptococcus pneumoniae, Staphylococcus aureus, Salmonella typhimurium*
NOD-like receptor (NLR) P3	Whole pathogens	*S. aureus*
	Toxins, LPS, MDP, and RNA	Bacteria
NLRP1	MDP	Bacteria
NLRP1b	Microbial toxin	*Bacillus anthracis*
NLRC4	Flagellin	*P. aeruginosa*

MOI, multiplicity of infection.

Complement

Complement, a series of more than 20 proteins, is activated by microbial surfaces (alternative complement [AP] cascade) or via antibody or by the mannose-binding lectin system (Chapter 21). Complement activation leads to microbial lysis and the release of opsonins and chemoattractant molecules for phagocytic cells.[13] The classical complement pathway (CP) can be initiated either by antibody binding to cell surface epitopes or by antibody-independent autocatalytic activation of C1 to form C1q. Initiation of the APs by bacterial products or mannose-binding protein leads to the direct deposition of the C3b complex on the bacterial surface. Complement activation results in generation of opsonins (as C3b), anaphylatoxins (as C3a), and activation of the late components of the complement pathway, which results in the formation of a membrane attack complex (MAC). Gram-positive extracellular pathogens resist the bacteriolytic action of the MAC as a result of a thick peptidoglycan layer, which impedes the insertion of the MAC C5b-9 complex. Gram-negative bacteria can resist the MAC through structural alterations in their LPS (the possession of O antigen keeps the MAC at a distance from the bacterial surface) or by masking or deleting the epitope(s) responsible for binding bactericidal antibody.

Dendritic Cells

DCs sample live bacteria at the mucosal surface, traffic to mucosal lymphoid tissue, and induce B cells to produce bacteria-specific immunoglobulins.[4] In most tissues, DCs are at low level of activation and are immature, but upon activation, they take up and process antigen. DCs are rich in PRRs (*e.g.,* TLRs), and microbe–PRR interaction has a key role in shaping the T-cell response.[14] For example, TLR5 stimulation by bacterial flagellin can induce a Th17 response and B-cell immunoglobulin A (IgA) production. Skin contains a major supply of tissue DCs (Langerhans cells), and their involvement in combating skin and soft tissue infections must be considered along with their function and contribution to stimulating immunity during vaccination. Limited information is available regarding the role of DCs in host resistance to extracellular bacteria, but some studies have examined the interaction between bacteria and DCs. For instance, Unkmeir et al. studied the interaction of serogroup B meningococci with DCs.[12] Infection of DCs by meningococci resulted in a significant and rapid production of proinflammatory cytokines and chemokines, including TNF-α, IL-6, and IL-8 through a lipooligosaccharide (LOS)–dependent mechanism.[12] Murine studies provide some insight into the mechanisms extracellular bacteria use to avert phagocytosis by DCs. For example, *S. suis* polysaccharide capsule reduces bacterial adhesion to the DC plasma membrane.[15] Once bacteria are internalized, they must resist degradation in the DC autophagolysosomes.

KEY CONCEPTS

Host Defenses and Immune Response at Epithelial Surfaces to Extracellular Bacteria

- Clearance and nonspecific host defenses at skin and mucosal surfaces
  - Epithelial barriers
  - Antibacterial factors (fatty acids, antimicrobial peptides, lysozyme, phospholipase A_2)
  - Mucociliary activity
  - Normal microbiota
  - Adherence blocking molecules
- Specific immune defenses at mucosal surfaces
  - Innate immune mechanism
  - Immunoglobulins
  - Phagocytosis at mucosal surfaces
- Mucosa-associated lymphoid tissue (MALT), gut-associated lymphoreticular tissue (GALT), bronchus-associated lymphoid tissue (BALT)

Macrophages

Phagocytic cells, macrophages, and PMNs are also present at mucosal surfaces (Fig. 27.2). These cells express PRRs and migrate to mucosal surfaces by chemotaxis and diapedesis between epithelial cells. Macrophages are also encountered after crossing the epithelial barrier. Specialized epithelial M cells of mucosal surfaces are key sites for antigen sampling, including viruses, and bacteria and macrophages surround these sites.[16] However, enteroinvasive pathogens, such as *Shigella*, can resist macrophages. *Shigella* induce macrophage apoptotic death by direct interaction of the bacterial protein IpaB with IL-1β-converting enzyme.

Polymorphonuclear Leukocytes

In areas of epithelial inflammation, PMNs can be recruited to mucosal and skin surfaces. PMNs are more effective in the presence of specific immune defenses, such as antibody and complement components. PMNs express PRRs and have both oxygen-dependent and oxygen-independent mechanisms of killing (Chapter 3) (Fig. 27.3). Activated neutrophils can release granule proteins with direct antibacterial action (*e.g.,* bactericidal/permeability-increasing [BPI]) or degradative activity (*e.g.,*

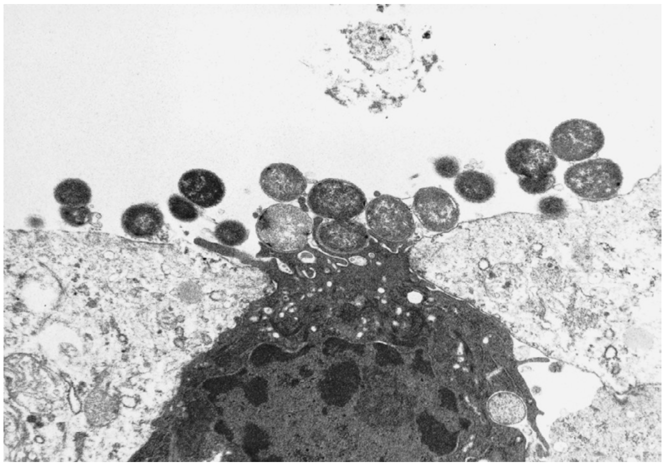


FIG 27.2 Bacterial Phagocytosis at Mucosal Surfaces. Transmission electron micrograph of phagocyte engulfing *Neisseria meningitidis* at a human respiratory epithelial mucosal surface (×19 000).[41]

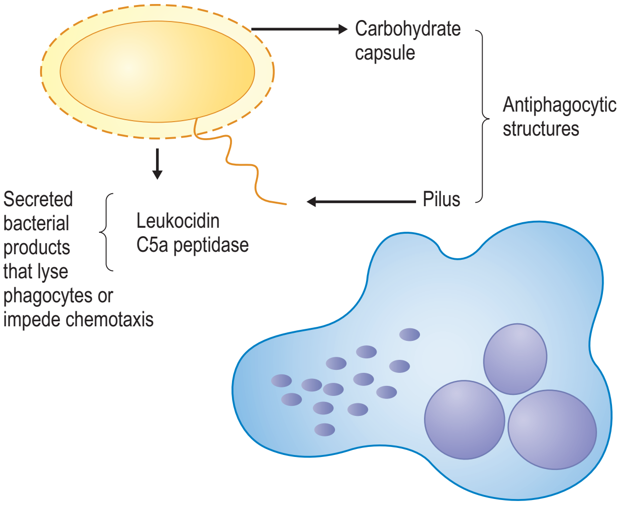

Carbohydrate capsule
Antiphagocytic structures
Pilus
Secreted bacterial products that lyse phagocytes or impede chemotaxis
Leukocidin
C5a peptidase

FIG 27.3 Bacterial Resistance to Polymorphonuclear Leukocytes (PMNs) in the Extracellular Environment. The two principal mechanisms of bacterial resistance to PMN killing. These consist of resistance to phagocytosis as a result of bacterial surface components (*e.g.*, capsule or pili) and the action of extracellular proteins that can lyse PMNs (*e.g.*, leukocidins) or decrease chemotaxis (*e.g.*, C5a peptidase). Bacteria growing in biofilms may be more protected from PMNs than are bacteria growing in the planktonic state.

elastase) and chromatin containing the antibacterial histone H2A.[17] These released compounds work together to form extracellular fibers, termed *neutrophil extracellular traps* (NETs), which can trap and kill gram-positive and gram-negative bacteria and degrade their virulence factors as well. NETs have been observed in instances of acute inflammation (experimental dysentery and spontaneous appendicitis) and provide a mechanism for reducing bacterial spread at sites of acute infection. The importance of PMNs in host defense against extracellular pathogens can best be highlighted by the increased frequency of bacteremias and other life-threatening infections in patients with neutropenia or those individuals with neutrophil deficits (*e.g.*, chronic granulomatous disease, Chediak-Higashi syndrome, or specific granule deficiency, to name a few). Bacterial infections associated with phagocytic dysfunction are described elsewhere (Chapter 22).

Innate Lymphoid Cells

Innate lymphoid cells (ILCs), a heterogeneous group of cells of the innate immune system, have lymphoid morphology but lack the capacity for rearrangement of the antigen receptors, a cardinal feature of the cells of the adaptive immune system.[18] Conventional natural killer (NK) cells, better known for generating inflammatory cytokines and cytotoxic activity against malignant cells and cells infected by viruses, seem to have a role in the defense against bacterial pathogens. Murine study data suggest that conventional NK-cell activation by lung macrophages is protective against *S. aureus* pneumonia.[19] A subset of ILCs, known as *NK-like cells*, produces IL-22 and has been found in mucosal sites, where these ILCs appear to have a protective activity against bacterial pathogens. IL-22 derived from these cells modulates AMP expression by epithelial cells. NK-like cells have minimal cytotoxic activity and are not strong producers of IFN-γ, core characteristics of conventional NK cells.[20]

Lymphocytes

Th1 cells are characterized by IFN-γ and function to activate macrophages to phagocyte and kill pathogens. While this mechanism of pathogen elimination is primarily directed against pathogens with a predominant intracellular lifecycle, Th1 cells are relevant for typical extracellular bacteria as pneumococcus and *S. aureus*.[21] The neutrophilic response to extracellular bacteria is primarily coordinated by Th17 cells.[22] Animal models have suggested that the Th17 response is central for protection against a wide variety of gram-positive and gram-negative bacteria. For example, Th17 response has been shown to induce nasopharyngeal clearance of *Pneumococcus* in both animal models and in children. Differentiation toward the Th17 subtype appears to be favored by strong antigenic signals and broad activation of pathogen recognition receptors. IL-17 and IL-22, the signature interleukins (ILs) of the Th17 response, promote AMP secretion by epithelial cells, neutrophil migration, and epithelial integrity. The increased susceptibility of subjects with Job's syndrome to *S. aureus* infections demonstrates the importance of the Th17 response in humans.

One in five cells in the intestinal epithelium is a lymphocyte. Mucosa-associated lymphoreticular tissue (MALT) comprises intraepithelial lymphocytes (IELs), lamina propria lymphocytes, and lymphoid follicles (*e.g.*, Peyer patches) and is sometimes divided into gut-associated (GALT), bronchial-associated (respiratory tract) (BALT), and genitourinary tract lymphoid tissues (Chapter 20).[23] Lymphocytes are important for homeostatic regulation and the maintenance of immune response against microbes at mucosal surfaces, including "extracellular" bacteria. These cells express PRRs (*e.g.*, TLRs), have constitutive cytotoxic activity, secrete chemokines and cytokines important in regulation and host defense, and act in concert with mucosal epithelial cells and exocrine glands.

Innate T cells represent a heterogeneous group of cells that possess T-cell receptors (TCRs) and are restricted to MHC-like molecules.[24,25] Unlike other T cells, innate T cells gain effector capacity before exiting the thymus and therefore can respond more readily to stimuli, including infections. This has led to the idea that innate T cells provide a bridge between the innate and adaptive immune systems during infections. Invariant natural killer T (iNKT) cells, a subset of innate T cells, are restricted to the MHC-like molecule CD1d. iNKT cells can be directly activated by bacterial pathogens as a result of binding of cell wall components of gram-negative bacteria to CD1d. Interestingly, cytokines secreted by APCs that have encountered bacterial pathogens can increase iNKT-cell accumulation of self-lipid antigen for CD1d activation. Mucosa-associated invariant T (MAIT) cells are restricted by MR1, a MHC-like receptor that binds molecules derived from the microbial riboflavin synthesis pathway.[24,26] Both gram-positive and gram-negative bacteria have been shown to activate MAIT cells. Murine models of intraperitoneal inoculation of gram-negative bacteria suggest the MAIT cells are important for the early clearance of pathogens.[25,26]

Immunoglobulins

Immunoglobulins (Igs), principally secretory IgA and IgG, are present at mucosal surfaces and in mucosal secretions. Important in the generation of these immunoglobulins at mucosal surfaces is the dissemination of IgA and IgG class–committed B- and T-helper (Th) cells with specificity to an antigen encountered and processed at one mucosal site to local and distant mucosal sites. Protective mucosal antibodies against bacteria may be derived from prior colonization, vaccines, or shared cross-reactive antigens on normal flora. Mucosal Igs may neutralize bacterial toxins, facilitate phagocytosis or bactericidal activity, inhibit bacterial adherence ligands, or sterically hinder other events necessary for bacterial colonization and invasion. Many extracellular bacterial pathogens (*N. meningitidis*, *N. gonorrhoeae*, *H. influenzae*, certain streptococci) colonize and/or infect mucosal surfaces where protective IgA_1 antibodies could become available.[27] These pathogens secrete an IgA_1 protease that cleaves IgA_1, thereby inactivating the molecule. IgA_1 protease can also recognize other substrates, notably lysosomal-associated membrane protein 1 (LAMP-1), which are important in host defense. Bacterial infections associated with abnormal immunoglobulin production or function are summarized in Chapter 34.

MECHANISM OF IMMUNE EVASION AND DISEASE BY EXTRACELLULAR BACTERIA

To colonize human epithelial and mucosal surfaces, bacteria must overcome the local host defense mechanisms described above. After navigating these defenses, adhesion to host cells is usually the first important step for bacterial pathogens (Fig. 27.4). Initial attachment of bacteria to human epithelial cells is, in part, mediated by pili, fimbriae, or other bacteria ligands or adhesins, and close adherence of bacteria to the human cell-surface receptors involves the cell wall, outer membrane proteins, LPS, and other bacterial surface structures. The attachment of bacteria to human epithelial cells prevents elimination of bacteria from the host. Attachment can also induce host cell pathways leading to cytoskeletal rearrangements, such as elongation and branching of the microvilli, the accumulation of actin, and calcium efflux, which facilitates close adherence and invasion of epithelial cells by normally "extracellular" bacteria, especially at sites with fluid movement. Strains of *E. coli* that successfully colonize the bladder and cause renal infection possess pili that allow adhesion to the renal epithelium.[28,29] Type IV pili are fundamental for attachment of gonococci to the male reproductive tract and play a role on the attachment of meningococci to vascular endothelial cells.[28,30] Meningococcal pili also facilitate twitching motility and microcolony formation, which allows the penetration of mucus and provides initial attachment. The pneumococcal CbpA surface protein promotes mucosal adhesion and dissemination.[29]

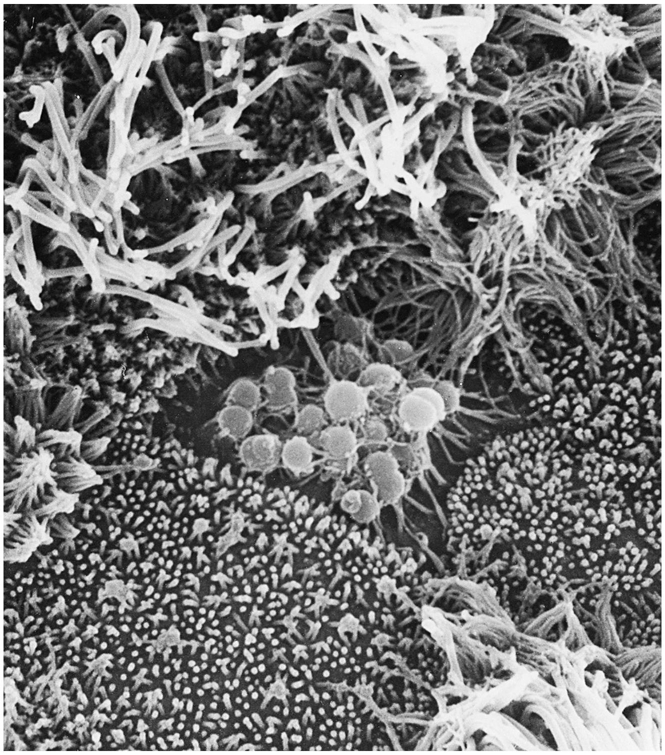


FIG 27.4 Colonization and Adherence of Extracellular Bacteria at Mucosal Surfaces. Scanning electron micrograph of *Neisseria meningitidis* adherence and microcolony formation of a human upper respiratory mucosa (×16 250).[41]

Bacteria utilize several mechanisms to avert the host immune response to bacterial surface antigens (see Table 27.1). Phase variation of adhesins is a mechanism of immune evasion common to pathogenic *Neisseria* spp. *Meningococcus*, for example, utilizes phase variation of the adhesion protein Opa and type IV pili during the process of colonization of human upper respiratory mucosal surfaces.[31] Sialylation of LPS, a potent inducer of host inflammatory response, is an example of bacterial "hiding" of surface antigens. For example, sialylation of lipooligosaccharide, a molecule analogous to LPS, in meningococci has been shown to increase resistance to CP and AP complement-mediated killing by decreasing the deposition of C3b and IgM on the cell surface, irrespective of capsular phenotype.

Many pathogenic extracellular bacteria interact with components of the complement system to induce negative regulation of the complement pathway.[29] Binding of human factor H (hfH) by meningococci factor H-binding protein (fHbp) downregulates the host AP and helps the organism to evade host innate immunity and is now included in the new serogroup B vaccines.[32] Proteolytic degradation of IgA_1 present in the urogenital and respiratory tracts is used to avert the action of the humoral system.[33] The elaboration of superoxide dismutase and catalase can reduce the efficacy of oxygen (O_2)–dependent killing of bacteria, but the high levels of O_2 radicals that accumulate in PMNs probably overcome these bacterial enzymes, as evidenced by the susceptibility of *S. aureus* to intraleukocytic killing. Several extracellular bacteria possess polysaccharide-rich capsules that resist phagocytosis. Polysaccharide capsule antigens can mimic human antigens. Antigen mimicry can lead to autoantibodies, as in the case of rheumatic fever and glomerulonephritis after *S. pyogenes* infection. Antigen mimicry can also dampen the immune response to bacterial antigens, as in the case of serogroup B *Meningococcus*. Other microbial surface structures, such as the pili of the gonococcus, can "stiff-arm" neutrophils, keeping them at a distance. A number of pyogenic bacteria (*e.g., S. aureus*) secrete leukocidins, which lyse phagocytes. Other pathogens (*e.g.,* group A streptococci) inhibit chemotaxis of neutrophils through the elaboration of enzymes (*e.g.,* C5a peptidase) that proteolytically cleave chemotactic signals. Some bacteria possess mechanisms to prevent opsonization by changing surface antigens.[30] Many bacteria form biofilms, which shield these microorganisms from host defense molecules and antibiotics.[34] Leukocytes that invade *S. aureus* biofilms exhibit impaired phagocytosis and decreased ability to kill bacteria. In addition, biofilm matrices can protect bacteria from antibody-mediated phagocytosis.

As previously noted, many "extracellular" bacteria have an intracellular component to their lifecycle. The intracellular environment provides protection from proteins of the complement system, Igs, and nonspecific barriers to infection present in the epithelia.[28] The entry of bacteria into epithelial cells provides access to nutrients and protection from host defenses, allows protected multiplication, and leads to shedding of organisms back to the mucosal surface, to facilitate transmission and further spread of the infection on the epithelium. Attachment can also initiate epithelial cell apoptosis or toxin-mediated cell death and lead to the breakdown of the epithelial barrier.

HOST RISK FACTORS FOR LOCAL AND SYSTEMIC INVASION BY EXTRACELLULAR PATHOGENS

Bacteria that breach mucosal and skin barriers and reach submucosal tissues of sites, such as pulmonary alveoli or the middle ear and/or the bloodstream, induce immune responses, including cytokine release, phagocytosis, complement activation, antibody release or production, and other local or systemic induction of the inflammatory cascade (Fig. 27.5). The survival of bacteria following colonization of the epithelium and access to the bloodstream depends on the integrity of the host immune response (including variability caused by genetic polymorphisms) and on the ability of the bacteria to resist this host immune response. Host factors that increase the risk for the development of systemic disease as a result of extracellular bacteria include polymorphisms in innate immune mechanisms, the absence of bactericidal or opsonizing antibodies, deficiencies in the complement pathways, and an absence of or reduction in neutrophil function or levels (see Table 27.1).

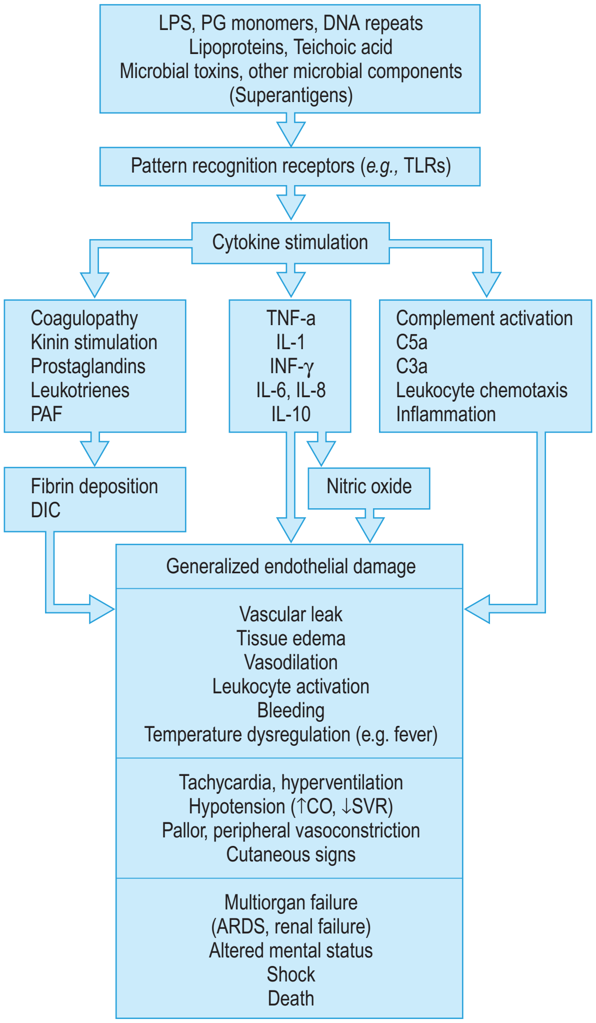

LPS, PG monomers, DNA repeats
Lipoproteins, Teichoic acid
Microbial toxins, other microbial components
(Superantigens)
Pattern recognition receptors (*e.g.,* TLRs)
Cytokine stimulation
Coagulopathy
Kinin stimulation
Prostaglandins
Leukotrienes
PAF
TNF-a
IL-1
INF-γ
IL-6, IL-8
IL-10
Complement activation
C5a
C3a
Leukocyte chemotaxis
Inflammation
Fibrin deposition
DIC
Nitric oxide
Generalized endothelial damage
Vascular leak
Tissue edema
Vasodilation
Leukocyte activation
Bleeding
Temperature dysregulation (e.g. fever)
Tachycardia, hyperventilation
Hypotension (↑CO, ↓SVR)
Pallor, peripheral vasoconstriction
Cutaneous signs
Multiorgan failure
(ARDS, renal failure)
Altered mental status
Shock
Death

FIG 27.5 Inflammatory cascade initiated during sepsis.

Complement deficiencies, either congenital or acquired, increase the risk for invasive bacterial diseases (Chapter 21). Because C3 plays a critical role in the complement cascade, congenital C3 deficiency or conditions that reduce C3 (*e.g.,* systemic lupus erythematosus, cirrhosis, nephritis, C3 nephritic factor) increase the risk for invasive disease due to pyogenic bacteria, such as *S. pneumoniae* and *N. meningitidis.* Mannose-binding lectin (MBL) is a plasma opsonin that initiates complement activation. *MBL* gene polymorphisms are found in children with meningococcal and pneumococcal sepsis. Properdin deficiency, leading to defective AP killing, is also associated with severe and recurrent meningococcal infections. Terminal complement deficiencies (C5–C8) are also associated with recurrent invasive bloodstream meningococcal and gonococcal infections, indicating an important role for insertion of the complement MAC in the bactericidal activity of human serum against pathogenic *Neisseria.* In adults, 10–20% of invasive meningococcal disease has been associated with a defect in the complement system.

In infants, antibacterial activity wanes as levels of passively transferred maternal antibody fall. This waning of antibody is correlated with the highest incidence of several "extracellular" pyogenic bacterial diseases (caused by *S. pneumoniae*, *N. meningitidis*, *H. influenzae* type b) in young children. During childhood and adolescence, levels of bactericidal antibodies rise and rates of these diseases decline. Specific antibodies are acquired through carriage and through cross-reacting epitopes on other commensal species. For example, cross-reactive antibodies to *N. meningitidis* are acquired by colonization with commensal *Neisseria* spp. (*e.g., Neisseria lactamica*) and unrelated bacteria (*e.g., Enterococcus faecium*, *Bacillus pumilus*, and *E. coli*). The lack of bactericidal antibodies against a strain recently acquired in the upper respiratory tract is an important risk factor for invasive meningococcal disease.

In addition to defects in innate immunity, Igs, and complement deficiencies, human genetic polymorphisms are associated with an increased risk or severity of bacterial diseases. For example, FcγIIa (CD32) receptor polymorphisms, Fcγ-receptor III (CD16), MBL, TLR4, tumor necrosis factor (TNF) promoter region polymorphisms, plasminogen activator and inhibitor expression, and hereditary differences in cytokine induction influence susceptibility to meningococcemia. Each of these polymorphisms can influence the course of invasive bacterial infection by influencing the response of the inflammatory cascade.

THERAPEUTIC PRINCIPLES

Sepsis

- Early effective antibiotic therapy associated with improved outcomes
- Lack of "source control" (*e.g.,* removal of infected lines, drainage of abscesses) linked to poor outcomes even in the presence of effective antibiotics
- Intensive and supportive care, management of fluid, electrolytes, and respiratory function
- Insulin for glucose control—unknown mechanism of protection; neutrophils have impaired function in the presence of hyperglycemia, insulin can have antiapoptotic effects

DELETERIOUS HOST RESPONSES

Inflammation and Autoimmunity

The host immune response can be the leading cause of tissue injury in the acute phase of an infection. Brain edema and infarcts, which are devastating consequences of pyogenic meningitis, occur as a result of the host inflammatory response. Corticosteroids are currently recommended as an adjunctive therapy to pneumococcal meningitis, an acute pyogenic infection. The use of small molecules targeting specific immunological pathways is an area of ongoing research. As acute infections are initially characterized by an inflammatory response followed by an antiinflammatory response, the timing of use of these compounds is of the utmost importance.[35,36] Infections can lead to autoimmune diseases. Molecular mimicry between a bacterial antigen and a host protein is a mechanism of generation of autoantibodies. Examples include rheumatic fever and glomerulonephritis after *S. pyogenes* infections, reactive arthritis following *Chlamydia trachomatis* urethritis, and the Guillain-Barré syndrome following *Campylobacter jejuni* enteritis. Molecular mimicry can also limit selection of epitopes for vaccine development.

Sepsis

Septicemia remains a leading cause of death in the United States and accounts for several billion dollars in health care expenditure.[35,36] Both gram-negative and gram-positive bacteria can rapidly multiply in the bloodstream and trigger sepsis and septic shock (Table 27.5). Septic shock is a result of an initial and widespread systemic proinflammatory response, resulting in hypotension, organ failure, and death. The later phase of sepsis is also characterized by an antiinflammatory response. Although survival of patients with the acute phase of sepsis has improved, advances on treatment and prevention of death, which can be secondary to nosocomial infections, have been slower. These secondary infections are often caused by less virulent organisms, likely as a result of "immunoparalysis" (as a result of an exaggerated antiinflammatory reaction) and also breaching of the physical barriers to infection as a result of invasive medical procedures (*e.g.,* intravenous lines, intubation, and bladder catheterization). The systemic inflammatory cascade of sepsis is initiated by recognition of PAMPs by PRR, both in extracellular (*e.g.,* LPS) and intracellular environments (*e.g.,* DNA fragments) (see Table 27.3) The severity of sepsis is also influenced by polymorphic alleles of genes involved in the inflammatory cascade.[35]

TABLE 27.5 Examples of Extracellular Bacteria Causing Sepsis by Presumed Site of Infection

Lung	*Streptococcus pneumoniae*, *Haemophilus influenzae*, *Pseudomonas aeruginosa*
Abdomen	*Escherichia coli*, mixed anaerobic infections
Urinary tract	*E. coli*, *Klebsiella* spp., *Enterobacter* spp., *Enterococcus* spp.
Soft tissue	*Streptococcus pyogenes*, *Staphylococcus aureus*, polymicrobial
Intravenous line	*S. aureus*, *Candida* spp.
Other	*Neisseria meningitides*

KEY CONCEPTS

Definitions of Sepsis and Septic Shock (Third International Consensus Definitions for Sepsis and Septic Shock)[42]

- Sepsis—life-threatening organ dysfunction causes by dysregulated host response to infections (including nonbacterial pathogens)
- Septic shock—subset of patients with sepsis with increased risk of death due to profound circulatory, cellular, and metabolic abnormalities
- SIRS (systemic inflammatory response syndrome)—nonspecific response to infectious and noninfectious insults; current guidelines recommend use of qSOFA (Quick Sequential Organ Failure Assessment Score) score in lieu of SIRS criteria, given its higher predictive value for sepsis

The morbidity and mortality of bacteremia and sepsis have been directly correlated with the initial levels of proinflammatory cytokines and the amount of circulating bacterial components. Indeed, the severity of gram-negative sepsis has been equated with high levels of endotoxin, increased levels of cytokines, and excessive activation of the AP. Disseminated intravascular coagulation, which often accompanies gram-negative sepsis, is caused by excessive activation of the coagulation cascade and downregulation of the fibrinolytic system associated with high levels of LPS. Levels of natural anticoagulants in the vasculature, such as antithrombin and protein C, are often low in gram-negative sepsis. The onset and severity of disseminated intravascular coagulation may be influenced by genetic polymorphisms in plasminogen activation or inhibition. The generalized, altered vascular endothelial lining facilitates thrombosis and thrombocytosis. Although much remains to be learned about the mechanisms by which gram-negative and gram-positive bacteria and microbial products trigger sepsis, significant advances have been made recently, particularly with endotoxin-mediated sepsis. Advances during the past decade include the identification of certain LPS–host protein interactions that result in delivery of LPS to host cell receptors and gene activation events that result in elevated expression of a diverse array of proinflammatory and antiinflammatory mediators (Fig. 27.6).

For example, TLR4 signaling requires an accessory protein, myeloid differentiation protein-2 (MD-2), which binds directly to endotoxin.[37] The key point, however, is that the LPS engagement

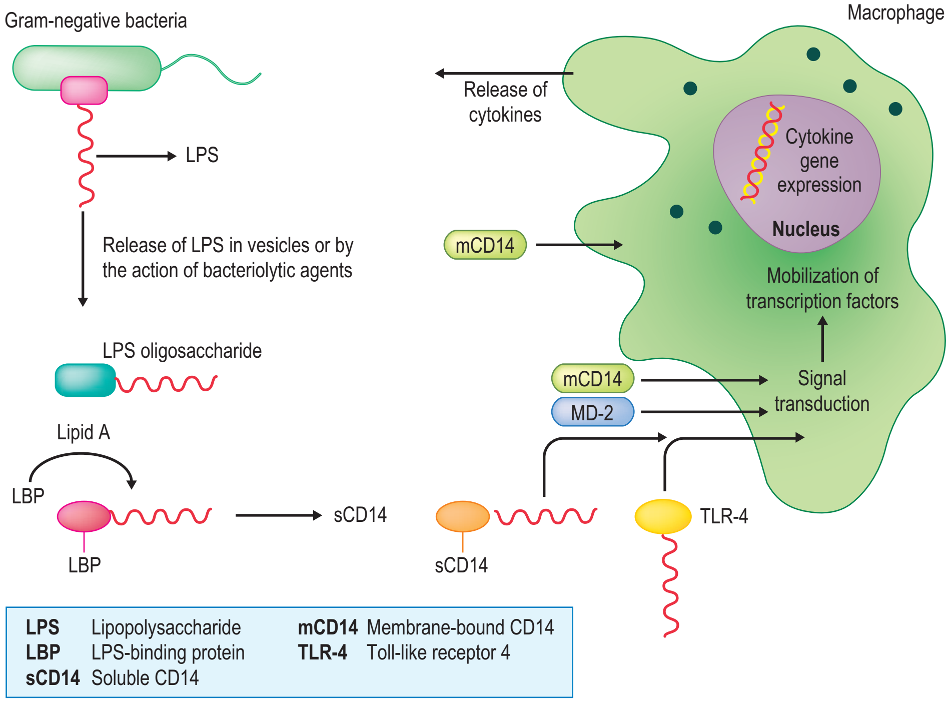

Gram-negative bacteria
LPS
Release of LPS in vesicles or by the action of bacteriolytic agents
LPS oligosaccharide
Lipid A
LBP
LBP
sCD14
sCD14
TLR-4
Release of cytokines
Macrophage
Cytokine gene expression
Nucleus
mCD14
Mobilization of transcription factors
mCD14
MD-2
Signal transduction
LPS Lipopolysaccharide
LBP LPS-binding protein
sCD14 Soluble CD14
mCD14 Membrane-bound CD14
TLR-4 Toll-like receptor 4

FIG 27.6 Lipopolysaccharide (LPS) Triggering of Cytokine Production by Macrophages. Steps known or presumed to be necessary for LPS triggering of proinflammatory cytokines.

of MD-2/TLR4 on host cells, particularly macrophages, triggers intracellular signaling events that ultimately, through NF-κB and other pathways, result in cytokine gene activation and production of cytokines (TNF-α, IL-1, IL-6, IL-8, IFNs). Other TLRs (*e.g.*, TLR2) play a critical role in the recognition of lipoproteins, and the recognition of these components is a likely a key determinant in the development of septic shock seen with gram-positive infections. TLR5 recognizes bacterial flagella, and such recognition is of importance in the host response to motile bacteria. Some human pathogens (*e.g.*, *H. pylori*) produce flagellin molecules that do not engage TLR5. TLR9 has been shown to recognize bacterial DNA CpG dinucleotides.[14] Taken together, the clinical syndrome of septic shock represents a series of interactions of bacterial products in the vascular space with pattern recognition molecules on serum proteins (Lipopolysaccharide-binding protein [LBP] and soluble CD14) and with host cell receptors (MD-2/TLR4 and TLR2, other TLRs) leading to signaling events and release of transcriptional factors that modulate cytokine gene expression. These events also trigger other events in the inflammatory cascade, leading to activation of the coagulation, complement, and kinin pathways. Superantigens can activate large pools of non–antigen-specific T cells by binding to MHC class II molecules and the TCR Vβ region outside the peptide-binding domain. The result is a cytokine storm with clinical manifestations of sepsis.

Early and effective antimicrobial therapy is the primary goal in the treatment of sepsis. In contrast to the initial phase of sepsis characterized by the release of TNF-α, IL-1, IL-6, and IFN-γ, an antiinflammatory response may predominate during the latter phase. The clinical failures of antiinflammatory therapeutic mediators (antiendotoxin antibodies, TNF-α, antagonists of IL-1, or platelet-activating factor) in sepsis suggests that this hypoinflammatory state encountered in many patients at presentation could be an additional target of immune modulation.

CLINICAL RELEVANCE

Signs of Septicemia

- Shaking chills, spiking fevers, or hypothermia (<36°C)
- Tachycardia, hyperventilation
- Pallor (peripheral vasoconstriction) and acrocyanosis, but 10–20% are flushed "warm shock"
- Nausea/vomiting, diarrhea
- Hypotension <90 mm Hg or >40 mm Hg decrease from baseline, in 20–35% of cases
- Cardiac output, decreased systemic vascular resistance (SVR)
- Cutaneous signs: purpura fulminans, petechiae, palpable purpura, ecthyma gangrenosum
- Change in mental status
  - Signs may be more subtle in elderly and uremic (renal failure) patients
  - White blood cell (WBC) count >12 000 cells/mm^3 or <4000 cells/mm^3 with >10% immature (band) forms
  - Oliguria (<20 mL/h of urine)

ENHANCEMENT OF IMMUNE RESPONSES TO EXTRACELLULAR BACTERIA (VACCINES AND IMMUNOMODULATION)

Vaccines are the most affordable and cost-effective health intervention for enhancing the immune response to extracellular bacterial and other microbial pathogens (Chapter 90). The use of vaccines to prevent diphtheria, pertussis, tetanus, and infections by *N. meningitidis*, *H. influenzae* b (Hib), and *S. pneumoniae* is a major public health success. Vaccines are also used for prevention of disease caused by *Salmonella typhi*, *Vibrio cholera*, and *Bacillus anthracis.* The efficacy of vaccines to extracellular bacteria is most often correlated with enhanced bactericidal antibodies, opsonic antibodies, and/or neutralizing antibodies, both systemically and at mucosal surfaces. Enhancement of these immune mechanisms can provide protection even in immunocompromised individuals (although immune responses are diminished in these individuals). For example, vaccination with meningococcal capsular vaccines protect patients with hereditary complement deficiencies by enhancing opsonophagocytic activity. However, vaccines have limitations in terms of long-lived immune responses, safety issues, and poor responses in certain populations (extremes of age) when infections with extracellular bacteria are most common. Advances in genetic engineering, immunology, molecular pathogenesis, vaccine adjuvants, and delivery systems are resulting in the development of new vaccines and vaccine approaches that enhance the immune response to extracellular bacterial pathogens (*e.g.*, protein–polysaccharide conjugate vaccines for Hib, *S. pneumoniae*, and *N. meningitidis*).

The conjugation of bacterial polysaccharides to carrier proteins, such as diphtheria or tetanus toxins, has been a major advance in stimulating immune responses to saccharide bacterial antigens. Polysaccharide capsules, for example, when used alone, are "T-independent" antigens; they do not require the presence of T cells to induce an immune response, and generate IgM as the dominant antibody produced (Chapter 6). Failure to induce memory and failure of affinity maturation follow polysaccharide immunization. Thus polysaccharides are poorly immunogenic in infants, older adults, and those with impaired antibody production—groups most susceptible to encapsulated bacterial pathogens. Covalent linkage of the polysaccharide to a carrier protein converts the polysaccharide to a thymus-dependent antigen generating IgG anticapsular antibodies and memory B cells. Because these vaccines induce vigorous mucosal immune responses, they also provide "herd" protection. A major (and unanticipated) result of vaccination with the Hib, meningococcal, and pneumococcal conjugate vaccines is the interruption of mucosal carriage, decreased transmission, and herd protection. These vaccines are now used as part of the routine immunization series in all age groups.

The presentation to CD4 T cells of an antigen by MHC class II molecules is critical for an immune response and influences the amount of antibody, the affinity of that antibody, and the duration of response. CD4 T cell subsets influence the qualitative and quantitative features of an immune response to vaccines and bacterial antigens. As new subsets of CD4 T cells are described, research is necessary to determine their role in existing vaccines and strategies to exploit them in new preparations (Chapter 16).[38] Mucosal vaccination is an attractive strategy for protection against extracellular bacteria as mucosal colonization often precedes invasive disease. The hope is to generate effective immune responses at the mucosa and avoid the need for parenteral injection.[39] In addition, recent advances in the application of systems biology to define molecular signatures that correlate with and predict vaccine immunity have greatly enhanced our understanding of immune responses to vaccination.[40]

Considerable progress is being made in the development of new vaccine adjuvants and immune modulators (see Chapter 90 for updated information). Aluminum salts, used since the 1930s in many vaccines against bacteria, induce a >90% Th2 response. As noted above, bacterial toxins as conjugates can be used to enhance immunogenicity. In addition, saponin adjuvants, liposomes, CpG DNA repeats, and TLR agonists are in use or under development as adjuvants. Monophosphoryl lipid A (MPL), long known to be an effective adjuvant in animal models, is now approved for human use, as an oil-in-water emulsion containing squalene. Cytokines, such as IL-1, IL-2, IL-12, IL-18, and granulocyte macrophage–colony-stimulating factor (GM-CSF), also modify and enhance immune responses to vaccines. IL-12, for example, induces strong Th1 shifts, and GM-CSF is a comigrating signal for DCs and stimulates antigen processing and presentation. Antigen recognition and processing in macrophages is critical to determining T-cell responses and can be manipulated by selected adjuvants. Immune modulation is being evaluated not only for the enhancement of bacterial vaccines but also as adjunct therapy for serious bacterial infections, such as sepsis. Vaccines and specific immunotherapeutic approaches, such as cytokines, may also find use against chronic tissue-damaging inflammatory reactions created by persistent extracellular bacteria (*e.g.*, *Helicobacter*) and autoimmune reactions that may be induced by cross-reactive bacterial antigens (*e.g.*, *C. jejuni* and Guillain-Barré syndrome).

TRANSLATIONAL RESEARCH OPPORTUNITIES

An important challenge for the next decade will be to take the rapidly expanding basic discoveries in innate immunity, systems biology, and response to bacterial antigens into clinical applications. The design and use of bacterial vaccines through the assessment of innate immune molecular signatures after vaccination, both for general use and for subpopulations of nonresponders, is one example. A second is the continued development of small-molecule inhibitors or enhancers that specifically target innate immune pathways to modulate bacterial immune responses. A third is the control of mucosal immune responses to prevent or eliminate colonization by bacterial pathogens. A fourth is the understanding of role of the microbiome in shaping the immune response to pathogens and vaccines and its therapeutic potential for both infections and noninfectious diseases. Fecal microbiota transplantation for *C. difficile* colitis is an early example of therapeutic use of the microbiome. Finally, the development of new therapies for acute bacterial sepsis may be based on improved understanding and control of the immune responses in sepsis.[36]

ON THE HORIZON

- Tailoring vaccine design based on assessment of innate immune molecular signatures
- Small molecule inhibitors or enhancers specifically targeting innate immune pathways
- Identification of immune responses that prevent or eliminate mucosal bacterial pathogen colonization
- Development of new therapies modulating immune response in sepsis
- Defining microbial community and metagenome changes after antibiotic treatment
- Managing disease based on the human microbiome

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Silva MT. Classical labeling of bacterial pathogens according to their lifestyle in the host: inconsistencies and alternatives. Front Microbiol 2012;3:71.
2. Medzhitov R. Recognition of microorganisms and activation of the immune response. Nature 2007;449:819–26.

3. Moens E, Veldhoen M. Epithelial barrier biology: good fences make good neighbours. Immunology 2012;135:1–8.
4. Nestle FO, Di Meglio P, Qin JZ, et al. Skin immune sentinels in health and disease. Nat Rev Immunol 2009;9:679–91.
5. Eloe-Fadrosh EA, Rasko DA. The human microbiome: from symbiosis to pathogenesis. Annu Rev Med 2013;64:145–63.
6. van Rensburg JJ, Lin H, Gao X, et al. The human skin microbiome associates with the outcome of and is influenced by bacterial infection. MBio 2015;6:e01315–15.
7. Gallo RL, Hooper LV. Epithelial antimicrobial defence of the skin and intestine. Nat Rev Immunol 2012;12:503–16.
8. Kumar S, Ingle H, Prasad DV, et al. Recognition of bacterial infection by innate immune sensors. Crit Rev Microbiol 2013;39:229–46.
9. Olive C. Pattern recognition receptors: sentinels in innate immunity and targets of new vaccine adjuvants. Expert Rev Vaccines 2012;11:237–56.
10. Beutler B, Jiang Z, Georgel P, et al. Genetic analysis of host resistance: Toll-like receptor signaling and immunity at large. Annu Rev Immunol 2006;24:353–89.
11. Quinn SR, O'Neill LA. A trio of microRNAs that control Toll-like receptor signalling. Int Immunol 2011;23:421–5.
12. Unkmeir A, Kämmerer U, Stade A, et al. Lipooligosaccharide and polysaccharide capsule: virulence factors of *Neisseria meningitidis* that determine meningococcal interaction with human dendritic cells. Infect Immun 2002;70:2454–62.
13. Ricklin D, Hajishengallis G, Yang K, et al. Complement: a key system for immune surveillance and homeostasis. Nat Immunol 2010;11:785–97.
14. Kawai T, Akira S. The role of pattern-recognition receptors in innate immunity: update on Toll-like receptors. Nat Immunol 2010;11:373–84.
15. Papadopoulos A, Gorvel JP. Subversion of mouse dendritic cell subset function by bacterial pathogens. Microb Pathog 2015;89:140–9.
16. Sansonetti PJ, Phalipon A. M cells as ports of entry for enteroinvasive pathogens: mechanisms of interaction, consequences for the disease process. Semin Immunol 1999;11:193–203.
17. Brinkmann V, Reichard U, Goosmann C, et al. Neutrophil extracellular traps kill bacteria. Science 2004;303:1532–5.
18. Spits H, Artis D, Colonna M, et al. Innate lymphoid cells—a proposal for uniform nomenclature. Nat Rev Immunol 2013;13:145–9.
19. Small CL, McCormick S, Gill N, et al. NK cells play a critical protective role in host defense against acute extracellular *Staphylococcus aureus* bacterial infection in the lung. J Immunol 2008;180:5558–68.
20. Fuchs A, Colonna M. Natural killer (NK) and NK-like cells at mucosal epithelia: mediators of antimicrobial defense and maintenance of tissue integrity. Eur J Microbio Immunol (Bp) 2011;1:257–66.
21. Cole J, Aberdein J, Jubrail J, et al. The role of macrophages in the innate immune response to Streptococcus pneumoniae and Staphylococcus aureus: mechanisms and contrasts. Adv Microb Physiol 2014;65:125–202.
22. Peck A, Mellins ED. Precarious balance: Th17 cells in host defense. Infect Immun 2010;78:32–8.
23. Yoshikai Y. The interaction of intestinal epithelial cells and intraepithelial lymphocytes in host defense. Immunol Res 1999;20:219–35.
24. Gao Y, Williams AP. Role of innate T cells in anti-bacterial immunity. Front Immunol 2015;6:302.
25. Gold MC, Lewinsohn DM. Co-dependents: MR1-restricted MAIT cells and their antimicrobial function. Nat Rev Microbiol 2013;11:14–19.
26. Napier RJ, Adams EJ, Gold MC, et al. The role of mucosal associated invariant T cells in antimicrobial immunity. Front Immunol 2015;6:344.
27. Mulks MH, Plaut AG. IgA protease production as a characteristic distinguishing pathogenic from harmless neisseriaceae. N Engl J Med 1978;299:973–6.
28. Ribet D, Cossart P. How bacterial pathogens colonize their hosts and invade deeper tissues. Microbes Infect 2015;17:173–83.
29. Pan X, Yang Y, Zhang JR. Molecular basis of host specificity in human pathogenic bacteria. Emerg Microbes Infect 2014;3:e23.
30. Criss AK, Seifert HS. A bacterial siren song: intimate interactions between Neisseria and neutrophils. Nat Rev Microbiol 2012;10:178–90.
31. Rouphael NG, Stephens DS. *Neisseria meningitidis*: biology, microbiology, and epidemiology. Methods Mol Biol 2012;799:1–20.
32. McNeil LK, Zagursky RJ, Lin SL, et al. Role of factor H binding protein in *Neisseria meningitidis* virulence and its potential as a vaccine candidate to broadly protect against meningococcal disease. Microbiol Mol Biol Rev 2013;77:234–52.
33. Mahalingam S, Lidbury BA. Antibody-dependent enhancement of infection: bacteria do it too. Trends Immunol 2003;24:465–7.
34. Roilides E, Simitsopoulou M, Katragkou A, et al. How Biofilms evade host defenses. Microbiol Spectr 2015;3(3).
35. Hotchkiss RS, Karl IE. The pathophysiology and treatment of sepsis. New Engl J Med 2002;348:138–50.
36. Leentjens J, Kox M, van der Hoeven JG, et al. Immunotherapy for the adjunctive treatment of sepsis: from immunosuppression to immunostimulation. Time for a paradigm change? Am J Respir Crit Care Med 2013;187:1287–93.
37. Needham BD, Trent MS. Fortifying the barrier: the impact of lipid A remodelling on bacterial pathogenesis. Nat Rev Microbiol 2013;11: 467–81.
38. Kumar P, Chen K, Kolls JK. Th17 cell based vaccines in mucosal immunity. Curr Opin Immunol 2013;25:373–80.
39. Lycke N. Recent progress in mucosal vaccine development: potential and limitations. Nat Rev Immunol 2012;12:592–605.
40. Nakaya HI, Pulendran B. Vaccinology in the era of high-throughput biology. Philos Trans R Soc Lond B Biol Sci 2015;370(1671).
41. Stephens DS, Hoffman LH, McGee ZA. Interaction of *Neisseria Meningitidis* with human nasopharyngeal mucosa—attachment and entry into columnar epithelial cells. J Infect Dis 1983;148:369–76.
42. Singer M, Deutschman CS, Seymour CW, et al. The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). JAMA 2016;315:801–10.

MULTIPLE-CHOICE QUESTIONS

1. With regard to the use of corticosteroids as adjunctive therapy for acute infections:
   - **A.** Corticosteroids are immunosuppressive and should be avoided during acute infections.
   - **B.** Corticosteroids are generally recommended for the treatment of sepsis.
   - **C.** Corticosteroids are never contraindicated during acute infections.
   - **D.** Corticosteroids are beneficial in certain types of acute infection, presumably because of their antiinflammatory effect.

2. A 18-year-old college freshman presents to the hospital with meningococcal meningitis. Upon review of his medical history, you learn that he had disseminated gonococcal infection at the age of 16 years. Should you consider investigation for an immunodeficiency? If so, which one? Should you offer any prophylactic therapies for this patient?
   - **A.** No investigation for an immunodeficiency is warranted. Meningococcal vaccination should be offered.
   - **B.** These infections suggest a T-helper (Th)17 deficiency, and there is no approved prophylactic therapy for this disorder.
   - **C.** This patient should have diagnostic workup for terminal complement deficiencies. Meningococcal vaccination should be offered.
   - **D.** This patient should have diagnostic workup for terminal complement deficiencies. If confirmed, stem cell transplantation can be offered.

3. The normal microbiota maintains a balance with the host by avoiding activation of pathogen recognition receptors. Pathogen recognition receptors are activated by pathogenic bacteria only.
   - **A.** False. Colonizers and pathogens have pathogen-associated molecular patterns, and both can activate pattern recognition receptors.
   - **B.** True. Colonizers are devoid of pathogen-associated molecular patterns.
   - **C.** True. Colonizers have pathogen-associated molecular patterns but do not activate pathogen recognition receptors.
   - **D.** False. Colonizers and pathogens have pathogen-associated molecular patterns, and both can activate pattern recognition receptors, but colonizers only do so when invading host tissues.

28

Host Defenses to Spirochetes

Nicolás Navasa, Erol Fikrig, Juan Anguita

Spirochetes constitute a unique group of bacteria that inhabit many different environments, such as soil, arthropods, and mammals. These microorganisms cause numerous human illnesses, including syphilis and Lyme disease. Spirochetes share a typical spiral shape and a distinctive flat-wave morphology. Cellular dimensions vary over a wide range, with a diameter between 0.09 to 0.75 μm and lengths that range from 3 to 500 μm. They are motile organisms with a multilayered outer membrane that encapsulates a peptidoglycan layer surrounding their inner membrane. Motility is enabled by the presence of endoflagella located in the periplasmic space. These axial filaments are arranged in a bipolar fashion and extend toward the opposite end of the cell (Fig. 28.1). The viability of the organism is dependent on an intact outer membrane, which can be damaged by variations in osmolarity, antibodies, or complement resulting in the loss of intracellular components and ultimately death of the bacterium.

Several spirochetal species are able to induce disease in mammals (Table 28.1). Lyme disease was first discovered in 1976 as an illness affecting a cluster of children in Lyme, Connecticut. Several manifestations of the disease, however, had been known for over a century in parts of Europe. In 1982, the agent of Lyme disease was identified as *Borrelia burgdorferi.*[1] In contrast, *Treponema pallidum* subspecies *pallidum* is the causative agent of venereal syphilis, a disease that has been recognized for over 5 centuries, although its agent was not identified until 1905.[2]

The genomes of *B. burgdorferi* and *T. pallidum* have been sequenced.[3,4] Despite similar ancestry and morphological features, these spirochetes have striking differences at the genetic level, which may account for the differences in their lifecycles, environmental adaptations, and the diseases they cause. Both *B. burgdorferi* and *T. pallidum* have relatively small genomes compared with other microorganisms. *B. burgdorferi* has one of the most complex genomes known among prokaryotes, with a single linear chromosome and 21 plasmids, the largest number of plasmids of any characterized prokaryote. Of the 21 plasmids, nine are circular and 12 are linear. Furthermore, less than 10% of *B. burgdorferi* plasmid-coding regions are found in other microorganisms, including spirochetes, which underscore the uniqueness of this spirochete among microorganisms. Unlike *B. burgdorferi, T. pallidum* contains a single circular chromosome with no plasmids. Yet, 476 open reading frames (ORFs) in *T. pallidum* have orthologous genes in *B. burgdorferi,* and almost 60 of these orthologous genes encode proteins of unknown biological function that are specific to spirochetes.

In contrast to other spirochetes, *B. burgdorferi* and *T. pallidum* do not contain lipopolysaccharide (LPS). Lipoproteins are the major immunogens of *B. burgdorferi* and most likely *T. pallidum,* and thus they are their dominant proinflammatory agonists. Five percent of the chromosomal ORFs and 14.5–17% of the functionally complete ORFs contained in the plasmids of *B. burgdorferi* encode putative lipoproteins. Interestingly, only 2.1% of the *T. pallidum* ORFs encode putative lipoproteins. The abundant lipoprotein coding potential of *B. burgdorferi* suggests that these molecules may be important for the survival of the spirochete. In fact, as a result of environmental changes, expression of several *B. burgdorferi* lipoproteins changes throughout the lifecycle of the spirochete.[5] For instance, the lipoprotein outer surface protein (Osp) A is expressed at high levels in the gut of the unfed tick, but upon feeding OspA expression is downregulated, and the expression of OspC increases 90-fold.[5] It is thought that many Osps have adhesive functions, and it has been shown that OspA is involved in the attachment of *B. burgdorferi* to the gut of the tick through an interaction with the tick receptor TROSPA.[6] Furthermore, OspC may be necessary for the migration of *B. burgdorferi* from the gut of the tick to the salivary glands and for its survival in the mammalian host. Changes in lipoprotein expression are primarily regulated by the alternative sigma factors, RpoS and RpoN[7] and other unknown mechanisms.

CLINICAL MANIFESTATIONS

Lyme Disease

B. burgdorferi sensu stricto (*B. burgdorferi*) is the etiological agent of Lyme disease in the United States and other parts of the world, whereas *B. afzelii* and *B. garinii* are agents of Lyme disease that are restricted to Europe and Asia. Infection with Lyme disease–causing spirochetes has also been observed in Japan, Russia, and China. In the United States, transmission of the spirochete is through the hard-bodied ticks of the *Ixodes* complex, mainly *I. scapularis* and *I. pacificus,* whereas *I. ricinus* is responsible for the transmission of the spirochete in much of Eurasia. According to the Centers for Disease Control and Prevention (CDC), Lyme disease is the most common tickborne disease in the United States and the fifth most common nationally notifiable disease.[8] The disease is concentrated in the North Eastern seaboard, with further presence in the northern Midwest and the West of the Continental United States. In Europe, diagnosed Lyme disease accounts for more than 65 000 cases annually, which tend to concentrate in countries of Central and Eastern Europe. However, the disease is present to different degrees throughout much of Europe. In both the United States and Europe, it is widely believed that the disease is underreported and that the suspected real number of cases may be much higher.

An early hallmark of infection is the appearance of a skin rash, known as erythema migrans (Fig. 28.2), which appears at the inoculation site, often during the first week of infection (stage I Lyme disease), as a result of local inflammatory responses. Other symptoms secondary to the local inflammatory responses occurring during stage I can affect more distal sites and may include fever, headache, malaise, myalgia, and/or arthralgia. Hematogenous dissemination of the spirochete is stage II Lyme disease and results in colonization of different tissues and/or organs and presentation of a range of symptoms, such as conduction system abnormalities, meningitis, and acute arthritis. Joint inflammation appears in 60% of untreated individuals in the United States and predominately affects large joints, especially the synovium of the knee. Some untreated individuals develop stage III Lyme disease, which is generally characterized by prolonged infection with the spirochete. Late-stage symptoms may include chronic arthritis, neuroborreliosis, or cutaneous lesions, such as acrodermatitis chronica atrophicans. The range of symptoms that appear upon infection in Europe and the United States vary in relative terms: arthritis and carditis are

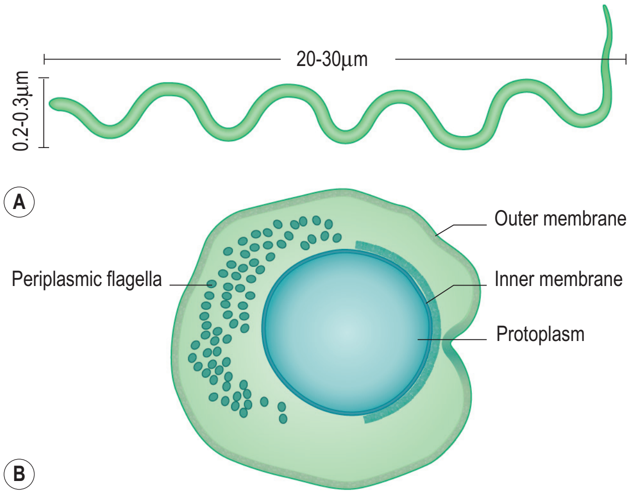


FIG 28.1 *Borrelia burgdorferi* structure is characterized by a distinctive flat-wave morphology consisting of approximately 18 bends and a length of 20–30 µm (A). A cross-section of this spirochete reveals the endoflagella, which are responsible for the unique morphology and motility of this organism (B).

TABLE 28.1 Major Diseases Caused by Spirochetes*

Disease	Agents	Distribution	Transmission	Symptoms
Lyme disease	*Borrelia burgdorferi*	North America, Europe	Tick engorgement	Development of a skin rash known as *erythema migrans,* accompanied by other symptoms, such as malaise, myalgia, and/or arthralgia. Symptoms can progress to include carditis and arthritis. Persistent infection can result in chronic arthritis, neuroborreliosis, or cutaneous symptoms (acrodermatitis chronica atrophicans).
	B. garinii	Asia, Europe		
	B. afzelii	Asia, Europe		
	B. andersonii	North America		
	B. japonica	Japan		
	B. lusitaniae	Southern Europe		
	B. valaisiana	Europe, Ireland, UK		
	B. mayonii	North America		
	B. miyamotoi	North America		
Relapsing fever	*B. hermsii*	Western USA	Tick engorgement	Clinical manifestations of infection include high-density spirochetemia, high fever, myalgias, and arthralgias and can even include cerebral hemorrhage and fatality.
	B. turicatae	Southwestern USA, Mexico		
	B. parkeri	Western USA		
	B. mazzotti	Central America		
	B. venezuelensis	Central America		
	B. duttonii	Sub-Saharan Africa		
	B. crocidurae	North Africa, Middle East		
	B. persica	Middle East, Central Asia		
	B. hispanica	Iberian peninsula, North Africa		
	B. latyschewii	Iran, Iraq, Eastern Europe		
	B. caucasia	Iraq, Eastern Europe		
Venereal syphilis	*Treponema pallidum pallidum*	Worldwide	Sexual contact	Disease progresses from a primary lesion (chancre) to a secondary eruption and then to a latent period, and if left untreated, tertiary symptoms may appear.
Endemic syphilis or Bejel syphilis	*T. pallidum endemicum*	Eastern Mediterranean region, West Africa	Nonsexual skin contact	Symptoms begin with a slimy patch inside the mouth, followed by blisters on the trunk and limbs. Bone infection in the legs soon develops, and in the later stages, lumps may appear in the nose and on the soft palate of the mouth.
Yaws	*T. pertenue*	Humid equatorial countries	Nonsexual skin contact	Destructive lesions of the skin and bones, which is rarely fatal but can be debilitating.
Pinta	*T. carateum*	Mexico, Central America, South America	Nonsexual skin contact	Dark-colored skin lesions found on those areas of the body that are exposed to sunlight. Eventually, the skin lesions become discolored.
Leptospirosis	*Leptospira interrogans*	Worldwide	Urine from an infected animal	Symptoms include fever, headache, chills, nausea and vomiting, eye inflammation, and muscle aches. In more severe cases, the illness can result in liver damage and kidney failure.

*Spirochetes are the causative agents of many diseases, which can have social as well as lasting health-related consequences.

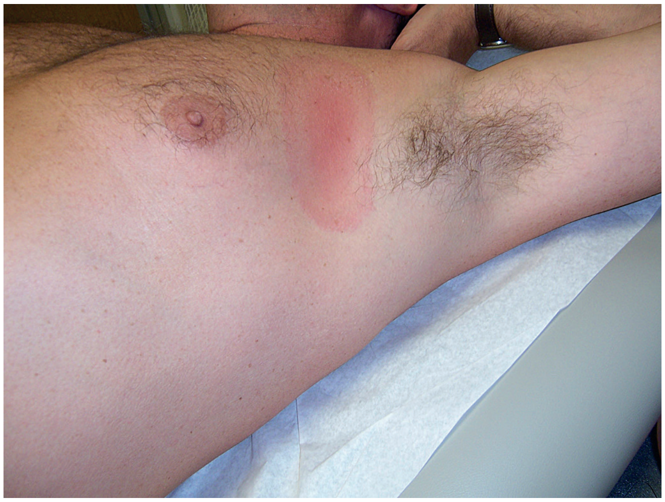


FIG 28.2 *Erythema migrans* caused by infection with *Borrelia burgdorferi,* the Lyme disease agent. (Courtesy of Gary Wormser, MD.)

more commonly found in the United States, whereas patients in Europe tend to show higher incidence of skin and nervous system involvement. This may be attributed to the heterogeneity of the *B. burgdorferi sensu lato* genospecies that cause the disease. In the United States, in the large majority of cases, *B. burgdorferi sensu stricto* is involved, whereas in Europe infections by *B. afzelii* and *B. garinii* predominate.

Diagnosis

A detailed clinical history and comprehensive physical examination are critical for the accurate diagnosis of Lyme disease. Appropriate laboratory testing, however, is a valuable diagnostic aid. A two-tiered approach is standard for the serodiagnosis of Lyme disease. Using this approach, serum is first tested for the presence of *B. burgdorferi*–specific antibodies by using a sensitive method, such as enzyme-linked immunosorbent assay (ELISA) or immunofluorescent assay (IFA). Sera testing negative for antibodies generally need not be tested further; however, those found to be positive or equivocal are further evaluated by the more specific immunoblotting for immunoglobulin M (IgM) and IgG antibodies. The detection of at least two of three specific bands in IgM or five of 10 specific bands in IgG is considered positive. Patients with early Lyme disease often report to a physician during the first few days of infection, at which time a detectable humoral response may not have developed. Consequently, the two-tiered approach is considered highly sensitive during the later stages of the disease (>90%) and less sensitive during very early infection. Furthermore, a positive serological test, particularly IgG, is evidence of exposure to *B. burgdorferi,* but it does not necessarily indicate active infection. All serological tests must then be evaluated in conjunction with a clinical assessment by the attending physician. Other diagnostic methods, such as culture and polymerase chain reaction (PCR) detection of *B. burgdorferi,* may be very useful to detect active infection, particularly of skin, joints, and the central nervous system (CNS). Culture, however, is generally limited to research laboratories, and the sensitivity and specificity of PCR can vary greatly among testing centers.

Treatment

During early stages of Lyme disease, such as that during which *erythema migrans* is present, oral administration of doxycycline (100 mg twice daily) or amoxicillin (500 mg three times a day) for approximately 2 weeks is recommended. Doxycycline has the advantage of being effective against *Anaplasma phagocytophilum,* which may also be transmitted by ticks. In areas where *B. burgdorferi* infection is prevalent, some experts recommend antibiotic therapy for individuals who served as hosts to *I. scapularis* ticks that were attached longer than 40–48 hours—the time required for transmission of the spirochete. It is extremely difficult, however, to consistently make accurate determinations of the species of tick and the degree of engorgement. Furthermore, randomized double-blind clinical trials involving individuals who were bitten by *I. scapularis* ticks led to the conclusion that antibiotic treatment of all individuals who had vector ticks removed is probably not warranted.

Venereal Syphilis

Infection with the agent of syphilis, *T. pallidum* subspecies *pallidum,* occurs worldwide. *T. pallidum* is an obligate human parasite that is almost exclusively transmitted when contact with infectious exudates from lesions of the skin and mucous membranes of infected individuals occurs. Clinically, this treponemal infection is first characterized by the formation of a hardened and painless ulcer at the initial site of infection. This primary lesion, called a *chancre,* forms after invasion of the bloodstream by the spirochete. Four to 6 weeks after infection, the edges of the chancre roll inward and upward and, in most cases, a secondary eruption appears, often accompanied by a rash on the palms of the hands and the soles of the feet. The secondary manifestations resolve within weeks to a year after the development of a vigorous cell-mediated immune response. Long periods of latency followed by late lesions of skin, bone, and viscera, as well as the cardiovascular system and the CNS, can occur despite clearance of the majority of the treponemes, which coincides with the resolution of the primary syphilitic lesion.

Diagnosis

Much like Lyme disease, the diagnosis for syphilis is based on the clinical presentations of the disease and serological tests. In addition, dark-field microscopy can be used for the identification of *T. pallidum* in the serous exudates of the chancre. This approach, however, is limited by the number of live treponemes in the exudates and by the presence of nonpathological treponemes in oral and anal lesions; as such, negative examinations on three independent days are required before a lesion is considered negative for *T. pallidum.*

Infection with *T. pallidum* leads to the production of nonspecific antibodies, which is the basis for other diagnostic tests, such as the traditional nontreponemal serological tests, including the Venereal Disease Research Laboratory (VDRL) and rapid plasma reagin (RPR) tests. Because these tests are nonspecific, false-positive reactions can occur as a result of pregnancy, autoimmune disorders, or infections. Therefore treponemal-specific tests, which detect antibodies to various antigens of *T. pallidum,* are often used to confirm the results of a nonspecific test. Interestingly, treponemal-specific tests are just as sensitive as nontreponemal tests; however, they are much more difficult and expensive to perform, which limits their use. These tests include, but are not limited to, the enzyme immunoassay (EIA) test for *T. pallidum*-specific IgG, *T. pallidum* hemagglutination test (TPHA), fluorescent treponemal antibody-absorption test (FTA-ABS), and ELISA.

Treatment

Susceptibility to infection with *T. pallidum* is universal, although only 30% of exposures with lesions teeming with the spirochete result in infection. Infection results in gradual development of immunity against *T. pallidum* and often against heterologous treponemes as well. However, treatment with long-acting penicillin subverts the development of immunity against *T. pallidum*. Administration of 2.4 million units in a single intramuscular dose the day that the primary, secondary, or latent syphilis is diagnosed is effective at killing the spirochetes. For people who are allergic to penicillin, there are alternative treatments, such as doxycycline.

HOST DEFENSES TO *B. burgdorferi*

KEY CONCEPTS

Protective Versus Pathological Responses to Borrelia burgdorferi

The early immune response to *B. burgdorferi* is necessary to control spirochetal burden; however, by itself, it is not sufficient to resolve infection.

Phagocytosis is a key element of the innate immune response, which is involved in the elimination of the bacteria while also contributing to the proinflammatory output of macrophages

The T cell–mediated response appears to be involved in pathology arising from infection.

A T cell–independent B-cell response is sufficient to resolve infection with *B. burgdorferi.*

B. burgdorferi virulence is attributed, in part, to the evolution of the spirochete's sophisticated tactics to evade killing mechanisms during all stages of the immune response: the first stage, beginning with transmission into the host via tick engorgement, at which time the spirochete is exposed to serum complement and cellular immunity; and the second and third stages, that is, hematogenous dissemination to and colonization of peripheral sites, at which time the host is producing specific antibodies. Both the innate and adaptive immune responses elicited by *B. burgdorferi* are discussed, with the supposition that these responses are required for efficient bacterial clearance, while acknowledging that unnecessarily prolonged or intense responses may contribute to pathology arising from infection. Indeed, predisposition to infection could be the result of one or more monogenic traits that confer primary immunodeficiencies; whether or not this is the case remains to be determined, but human studies have shown that responses to *B. burgdorferi* are diminished in individuals with specific mutations in or diminished expression of innate immune cell receptors (nucleotide-binding oligomerization domain-2 [NOD2] and Toll-like receptor-1 [TLR1]).

Innate Immune Responses

The initial recognition of pathogens with cells of the host relies on a complex interplay between pathogen recognition receptors (PRRs) and bacterial constituents, and this initiates a cascade of responses leading to the upregulation of chemokines and cytokines, adhesion molecules, and other effector molecules (Chapter 3). This response is often initiated by endothelial and/or epithelial cells, resulting in the recruitment of different types of innate immune cells and their activation. Each PRR recognizes a specific structure that is present in a group or groups of microorganisms and that distinguishes them from the more specific recognition of antigens by the T- and B-cell receptors. The recognition of patterns instead of specific antigens provides the innate immune system with a rapid way to respond to infecting organisms until the more specific response mediated by T and B cells develops.

A distinct innate cellular infiltration profile has been observed in *B. burgdorferi*–infected tissues. Based on studies using mice, the inflamed heart appears to predominately comprise macrophages, with smaller amounts of T and invariant natural killer T cells (iNKT cells). However, neutrophils are the primary innate immune cell found in the joints. In this tissue, resident myeloid cells act as the initiators of type I interferon (IFN) production upon encounter with *B. burgdorferi,* whereas endothelial cells and joint fibroblasts expressing adhesion/ activation markers amplify the response and serve as the major source of disease-promoting chemokines. This tissue-specific tropism complicates our understanding of the host defense against the spirochete, but this suggests that macrophages are more efficient at preventing successful colonization of *B. burgdorferi* because carditis is a less frequent manifestation of Lyme disease. These early defense responses also have a role in determining the type, strength, and duration of the more specific adaptive immune response, beyond their direct bacterial-killing capability.

Early Pathogen Recognition

The interaction of *B. burgdorferi* lipoproteins with complexes formed by TLRs 1 and 2 initiates a series of signaling cascades that results in the production of proinflammatory cytokines (IL-1β, tumor necrosis factor [TNF], interleukin [IL]-12, and IL-18, among others), chemokines (IL-8, monocyte chemoattractant protein [MCP]-1, keratinocyte chemoattractant [KC]), metalloproteinases, adhesion molecules (E-selectin, vascular cell adhesion molecule-1 [VCAM-1] and intercellular adhesion molecule-1 [ICAM-1]),[9] and type I IFNs. These interactions are critical for generating an effective spirochete-clearing inflammatory response, as demonstrated by experiments using mice deficient for TLR1, TLR2, or their adaptor molecule, MyD88. These mice had significant increases in *B. burgdorferi* burdens after infection. Moreover, humans with decreased TLR1 expression are hyporesponsive to the original OspA-based Lyme disease vaccine, and macrophages from these subjects also have diminished inflammatory responses to the lipoprotein. Besides TLR1/TLR2 complexes, other members of this family of receptors participate in, and amplify, the inflammatory response to *B. burgdorferi.* These include TLR5, TLR7, TLR8, and TLR9, which recognize different bacterial constituents. Importantly, the development of inflammatory arthritis in the absence of TLRs or MyD88 was the first suggestion that alternative pathways trigger the inflammatory response to the spirochete. Other PRRs that have been shown to be involved in the response to the spirochete include NOD-like receptors. For the full inflammatory response to take shape upon recognition of *B. burgdorferi,* most of the interactions of *B. burgdorferi* constituents and PRRs occur within the phagolysosome. Therefore phagocytosis is a hallmark of the innate immune cell recognition of the spirochete.

The interaction of several cell types with the spirochetes also involves several integrins belonging to the β_1, β_2, and β_3 groups. Integrins are involved in the adhesion of cells to a variety of

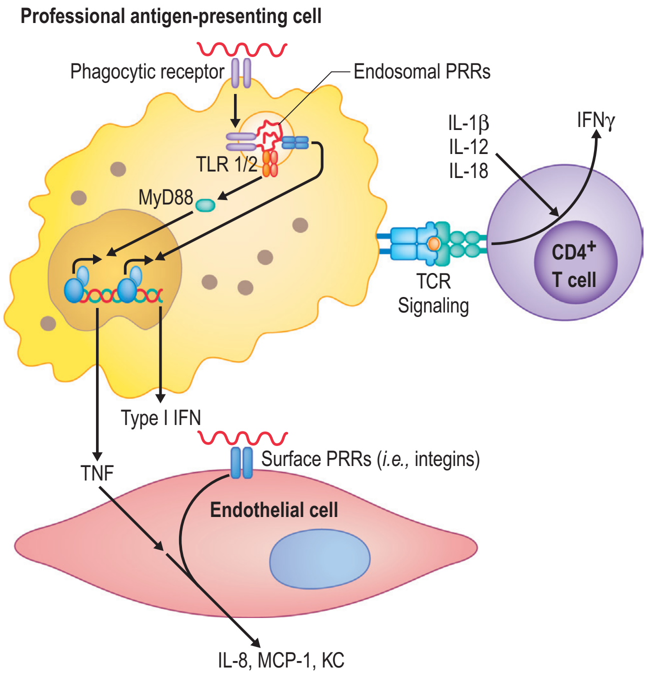


FIG 28.3 The interaction of *Borrelia burgdorferi* with pattern recognition receptors (PRRs) in innate immune cells and endothelial cells mediates the inflammatory response to the spirochete. Phagocytosis induces Toll-like receptor (TLR)–driven proinflammatory cytokine production as well as antigen presentation by professional antigen-presenting cells (APCs), which leads to the activation of CD4 effector T cells, marked by the production of IFN-γ. Likewise, PRR- and tumor necrosis factor (TNF) receptor signaling lead to the upregulation of chemokines by endothelial cells. Overall, these responses lead to increased activation and recruitment of innate immune cells in sites of infection.

ligands and mediate essential cellular processes, including attachment and cell migration. Some integrins have also been associated with the phagocytosis of microorganisms. For the most part, study of the interaction between the spirochete and integrins has focused on their role aiding the adhesion of *B. burgdorferi* to cells and the colonization of tissues and as receptors that contribute to signals that induce the production of proinflammatory factors (Fig. 28.3).

Phagocytic Cell Recruitment and Spirochetal Clearance

The recruitment of phagocytic cells and other cell types into sites of infection is mediated by the production of chemokines, increased vascular permeability, and upregulated expression of cell adhesion molecules in endothelial cells. *B. burgdorferi* induces the upregulation of these factors in different cell types. Chemokine production at sites of pathology in disease-susceptible C3H/HeJ mice and disease-resistant C57BL/6 J mice shows that inflammation is related to increased production of neutrophil and monocyte–macrophage chemokines, KC and MCP-1, respectively. In patients, the production of chemokines, especially IL-8, during the initial response to *B. burgdorferi* correlates well with the onset of symptoms known to occur during the early stages of infection, and this suggests that their production is increased during the beginning of the infection to recruit phagocytic cells, which are involved in the initial clearance of the spirochete.

Little is known about the molecular events or receptors that mediate *B. burgdorferi* phagocytosis despite the importance of this cardinal mechanism of pathogen clearance. MyD88-mediated signals substantially mediate the phagocytosis of *B. burgdorferi.* However, MyD88-mediated phagocytosis occurs independently of any known *B. burgdorferi*-recognizing TLRs. The mechanism of MyD88-mediated phagocytic uptake is therefore unclear. Furthermore, the analysis of MyD88-deficient macrophages shows that although reduced, phagocytosis is not absent in these cells. Therefore the uptake of *B. burgdorferi* by phagocytic cells seems to be mediated by more than one receptor with MyD88-dependent and MyD88-independent components. Complement receptor 3 (CR3) is a β_2 integrin, which, in cooperation with the GPI-anchored CD14, acts as a phagocytic receptor for *B. burgdorferi* independently of TLR or MyD88-induced signals.[10] Phagocytosis plays a major role in the pathogenesis of Lyme disease, not only through the control of bacterial numbers but also through modulation of the potency and quality of proinflammatory cytokine induction. Indeed, as opposed to MyD88-mediated phagocytosis, which is proinflammatory, the internalization of *B. burgdorferi* by CR3 tempers the inflammatory response of macrophages. Therefore the presence of alternative phagocytic mechanisms has nonredundant physiological consequences during infection with the spirochete.

Complement

The complement system is a key component of the innate immune system (Chapter 21). It comprises a collection of serum proteins and cell surface receptors that are involved in the early response to pathogens, including *B. burgdorferi.*[11] Destruction of microorganisms via complement involves the formation of a pore in the microbial cell membrane by the membrane attack complex (MAC), which results in the lysis of the organism. Three different pathways elicit complement activation: the classical (antigen/antibody–mediated) pathway (CP), the lectin pathway, and the alternative (pathogen surface) pathway (AP). These pathways converge at the level of C3 convertase, a protease that cleaves complement component C3 into C3a and C3b. As a result, C3b can (i) bind to the surface of the bacteria and facilitate internalization of the spirochete via opsonization; or (ii) it can bind C3 convertase and facilitate the deposition of downstream components onto the surface of the spirochete resulting in the formation of MAC and lysis of the cell.

B. burgdorferi activates the CP and AP of the complement cascade.[12] Moreover, the activation of complement has been associated with dramatic decreases in spirochetal numbers in different tissues of infected mice, indicating the importance of the complement system early in *B. burgdorferi* infection.

The members of the *B. burgdorferi sensu lato* group, which includes *B. burgdorferi sensu stricto, B. garinii,* and *B. afzelii,* have evolved a variety of mechanisms enabling them to escape complement-mediated lysis, including the expression of complement regulator-acquiring surface proteins (CRASPs). Of these CRASPs, the Erp (OspEF-related protein) family of outer membrane proteins serve as binding sites for the complement inhibitor factor H and factor H–like protein 1 (FHL-1).[13] The interaction of factor H with these proteins recruits a protease (factor I) that cleaves and inactivates the complement serum proteins, C3b and C4b. Cleavage of these two complement proteins prevents the deposition of downstream components onto the surface of the spirochete, thereby halting the formation of the membrane attack complex. *B. burgdorferi* also express a CD59-like molecule on the outer membrane that can inactivate MAC and prevent complement-mediated lysis.[14] It has been speculated that most borreliae are able to evade complement-mediated lysis, and recently, a novel protein expressed by *B. hermsii* (a relapsing fever spirochete) has been discovered that affords protection to spirochete by inactivating C3b.[15] More recently, it has been shown that *B. burgdorferi* expresses a protein, BBK32, on its surface, which prevents activation of the CP by blocking activation of the C1 complement complex.[16]

Adaptive Immune Responses

T Cell–Mediated Responses

Upon antigen presentation by macrophages, dendritic cells (DCs), or B cells, naïve CD4 T cells are activated and differentiate into effector T cells. Effector CD4 T cells are classified on the basis of their cytokine production profile, which determines their mode of action and downstream effects, and include Th1 (IFN-γ-producing), T-helper [Th]2 (IL-4, IL-5, IL-13), Th17 (IL-17), or regulatory T cells (Tregs; IL-10) (Chapter 16). Interaction of *B. burgdorferi* antigen with TLRs induces the production of IL-12, which drives the differentiation of CD4 T cells into Th1 effector cells. Th1 cells are regulators of the cell-mediated inflammatory reactions, which are characterized by macrophage activation, including phagocytosis, or the induction of opsonizing IgG antibodies. Although IFNγ and Th1 CD4 T cells have been shown to be protective during cardiac inflammation with *B. burgdorferi,* joint inflammation is independent of this effector cell type in mice. However, in patients with Lyme disease, Th1 cells dominate in the synovial fluid, and the severity of arthritis directly correlates with increased levels of Th1 cells in the synovium.[17] Whether neutrophilic infiltration during joint infection with the spirochete is influenced by Th effector cells that more directly affect this cell type, such as Th17 cells (through the production of IL-17) remains to be elucidated.

B Cell–Mediated Responses

Antibodies are specific and powerful effector molecules of the adaptive immune response (Chapter 15). Once antibodies bind to their specific foreign antigen, they confer protection to the host by using a variety of effector mechanisms: antibodies contribute to adaptive immunity by neutralizing microbes or their products, activation of complement, and opsonization, which leads to phagocytosis of microorganisms. The activation of B cells and their differentiation into antibody-secreting plasma cells often requires an interaction with Th cells, and this interaction controls isotype switching as well as somatic hypermutation. However, in response to *B. burgdorferi,* T cell–independent humoral responses also confer protection to the host. Thus mice deficient in T cells, CD40L, and MHC class II infected with *B. burgdorferi* mount a protective antibody response, which, upon passive serum transfer, affords protection to severe combined immune deficient (SCID) mice from homologous challenge. However, mice that lack both B and T cells developed severe arthritis and carditis in response to infection with *B. burgdorferi.*

The role of antibodies in controlling *B. burgdorferi* infection may be more important during the hematogenous dissemination phase, when they are easily accessible, than after the spirochetes have colonized tissues. *B. burgdorferi* is able to evade the humoral response by interacting with the extracellular matrix of the mammalian host via attachment to decorin, a major component of the extracellular matrix. *B. burgdorferi* attaches to decorin with a ligand-binding lipoprotein known as *decorin-binding protein A* (DbpA). A deficiency in decorin reduces the incidence of Lyme arthritis in mice, suggesting that the interaction of DbpA with the extracellular matrix provides a protective niche for *B. burgdorferi,* preventing humoral-mediated bacterial killing. Once the spirochete is in the joints and the heart, its clearance may be more dependent on cellular responses (*e.g.,* macrophages in the heart) than on antibodies. In fact, the lack of IFNγ-mediated activation of macrophages has profound consequences on murine cardiac inflammation, even in the presence of strong antibody responses. Furthermore, the bacterial clearance potential of infected mouse sera administered in newly infected mice is lost when administered 4–8 days after infection, potentially the result of the colonization of tissues into which antibodies are less able to penetrate.

B. burgdorferi can also avoid clearance by antibodies through antigenic diversity. *B. burgdorferi* differentially expresses outer membrane antigens under pressure from the immune response, which might contribute to the ability of the spirochetes to persist in the host. A mechanism that is potentially essential for spirochetal immune escape is the recombination that takes place at the *vls* locus,[18] located near the right telomere of the linear plasmid lp28-1. *B. burgdorferi* is also able to evade antibody responses during transmission from the tick by attachment to the tick salivary protein, Salp15, which interacts with the lipoprotein, OspC, and protects the spirochete from antibody-mediated killing.[19]

HOST DEFENSES TO *T. pallidum*

KEY CONCEPTS

Protective Versus Pathological Responses to Treponema pallidum

The role of the early innate response to *T. pallidum* is still poorly understood because immunogenic cell surface proteins able to activate pattern recognition receptors remain to be found.

It is likely that the cell-mediated immune response to *T. pallidum* is involved in the development of pathology following infection with the spirochete. Likewise, this response seems to be involved in the resolution of infection with *T. pallidum.*

The protection that the humoral response affords to the host is unclear. Furthermore, no definite antigens have been isolated from the spirochete because of the inability to cultivate this organism *in vitro.*

T. pallidum is known colloquially as the "stealth pathogen" because of its denuded outer membrane, which comprises mostly nonimmunogenic transmembrane proteins, whereas the highly immunogenic lipoproteins are contained within the periplasmic space.[20] This molecular architecture, coupled with the ability to generate antigenic variants, is responsible for the treponeme's remarkable ability to cause persistent infection with relatively few organisms.[21] We have yet to elucidate many of the specific immunological phenomena that occur as a result of infection with the treponeme and that are inherent to the clinical manifestations of the disease. Our limited understanding of immunological phenomena during infection mainly results from our inability to culture these organisms *in vitro* and to reproducibly infect an animal model. Consequently, our understanding of the immune responses to this pathogen is not nearly as detailed as our knowledge of those elicited in response to infection with *B. burgdorferi.*

Innate Immune Responses

Similar to *B. burgdorferi* lipoproteins, treponemal lipoproteins appear to be the major proinflammatory agonists during treponemal infection through engagement with their cognate receptors, TLR1/2 and CD14. The inflammatory milieu established by treponemal lipoproteins is a principal driving force for immune cell recruitment to *T. pallidum*–infected tissues. The importance of this immune system compartment during the response to *T. pallidum* is further supported by the systemic upregulation of innate immune cells during treponemal dissemination and by demonstration that macrophages are principal effectors of treponemal clearance during infection.[22]

Early Pathogen Recognition

Because of their location in the skin, DCs (*i.e.,* Langerhans cells), likely mediate the initial responses to treponemes at the primary syphilitic lesion. Because of the denuded outer membrane of *T. pallidum,* extremely high treponeme to DC ratios (500 : 1 and higher) are required for observable levels of phagocytosis *in vitro.* Thus unhindered replication of spirochetes at the initial site of infection probably underlies the vigorous cell-mediated immune response, as well as chancre appearance, which follows antigen presentation to and differentiation and activation of CD4 T cells.

Phagocytic Cell Recruitment and Spirochetal Clearance

It is not until after the recruitment of macrophages and their activation via CD4 T cell–derived IFN-γ, that the majority of spirochetes are killed and resolution of syphilitic lesions occurs. Activated macrophages readily phagocytose antibody-opsonized treponemes through Fc receptor–mediated uptake. Recently, it has been shown that MyD88 mediates the opsonophagocytosis of *T. pallidum* by macrophages without affecting the production of opsonic antibodies. As result, MyD88-deficent mice infected with *T. pallidum* showed a defect in bacterial clearance that resulted in worse inflammation. These results reveal a previously unsuspected link between MyD88 signaling and FcR-mediated phagocytosis. However, as with DCs, uptake via a direct PAMP–PRR interaction does not occur readily, and because of treponemal antigenic variation, not all spirochetes will be eliminated via opsonophagocytosis despite the vigorous inflammatory response. The relatively few remaining spirochetes are able to cause persistent infection.

Complement

The immunoprotection afforded by human syphilitic serum is, in large part, a result of the activation of the complement cascade by bactericidal antibodies and spontaneous hydrolysis of C3. Thus like *B. burgdorferi, T. pallidum* activates the CP and AP of the complement cascade, and there is a considerable amount of evidence for an important role of these pathways in syphilitic lesion resolution during human infection with treponemes. In contrast to *B. burgdorferi,* there is no evidence indicating that *T. pallidum* has evolved mechanisms to evade complement-dependent killing, suggesting that these complement pathways have a larger role in controlling treponemal infection than in the case of *B. burgdorferi.*

During experimental syphilis, immunization with purified outer membrane vesicles (OMVs) isolated from *T. pallidum* results in complement-dependent bactericidal activity.[23] More recently, immunization with OMVs led to the isolation of a bactericidal monoclonal antibody (mAb), M131, which provides partial protection to experimental syphilis. M131 binds to a phosphorylcholine surface epitope of *T. pallidum* and is the first demonstration of such an antigen on the surface of the spirochete.[24]

Adaptive Immune Responses

T Cell–Mediated Responses

The infiltration of T cells (CD4 and CD8) and macrophages into the primary and secondary syphilitic lesion facilitates local clearance of the majority of treponemes and lesion resolution via a vigorous cell-mediated immune response characteristic of a delayed-type hypersensitivity or Th1 response.[25,26] CD4 T cells are the principal T-cell subset found in the lesions and are believed to promote macrophage activation and subsequent treponeme clearance through IFN-γ secretion. The role of CD8 cells is not clear, but throughout the course of primary lesion development during experimental syphilis, their proportion related to CD4 cells increases. Once the treponemes are cleared and the lesion heals, a latency stage ensues, which is characterized by long-lasting, protective T-cell memory to *T. pallidum* antigens.

B Cell–Mediated Responses

Many functional activities of human syphilitic serum originate from B-cell responses to *T. pallidum.* Infection with *T. pallidum* invokes a humoral immune response early in the course of infection, which strengthens as the number of recognizable antigens increases during the progression of infection. A polymorphic gene family of *T. pallidum, T. pallidum* repeat (*tpr*), has been recently identified as related to the major surface protein (*msp*) genes of *T. denticola.*[27] A single member of the *tpr* family, tprK, serves as an antigen for opsonizing antibodies, suggesting that this protein is a surface antigen of *T. pallidum.* However, there are only inferential indications that surface antigens of *T. pallidum* exist. So, although candidate *T. pallidum* surface proteins have been advanced on the basis of porin activity or homology with a surface protein of *T. denticola,* there is no direct evidence identifying specific surface antigens on *T. pallidum.* Furthermore, like most pathogenic organisms, *T. pallidum* has evolved various mechanisms to escape host killing. TprK is an immunogen with seven discrete variable regions differing among isolates of the spirochete.[27] In fact, the antibody response to *T. pallidum* is directed against these variable regions, which leads to immune selection of new TprK variants; thus antigenic variation is involved in the reinfection of hosts despite robust immune responses.[27,28]

Translational Research

ON THE HORIZON

- Efficient diagnostic methods that are required to rapidly identify infection and procure antimicrobial treatments
- Development of new generation vaccines for Lyme disease, including those that target the tick vector and that can prevent other tick-borne infections
- Attention to communities at risk for contracting syphilis with investments in prophylaxis and efficient treatments

The challenge in the next 5–10 years is to develop effective diagnostic methods for both Lyme disease and syphilis and to devise new preventive measures. For Lyme disease, advances in genetic manipulation, as well as other means to study structure/

function of key components of the spirochete, could lead to development of single or combination vaccines. Further, the identification of host factors that mediate a protective response and those that contribute to inflammation may result in a better understanding of the disease. Overall, the identification of the elements that mediate specific tissue tropisms for the bacterium and their interaction with local/infiltrating cellular components can permit the design of targeted therapies in conjunction with antimicrobial treatments.

A significant effort is being made to find new antigenic determinants that can be the base for a vaccine to prevent infection with *B. burgdorferi*. While the search for spirochetal antigens continue, a new wave of research has focused on finding antigenic determinants that can prevent the efficient attachment of the tick vector to the mammalian host, which, in turn, could dampen the ability of the arthropod to transmit *B. burgdorferi* and perhaps other pathogens.

For syphilis, the study of the microorganism is hampered by the difficulties associated with its culture and manipulation. Although advances have been made in our understanding of the pathology associated with infection, a significant challenge in the near future is to fully understand this pathogen. An area in which advances can be made in the next decade consists of better education of populations that are at risk of acquiring the infection. Because these populations are usually associated with low-income communities, a significant investment needs to be made to enable persons at risk to better evade the transmission of the pathogen, through education, better diagnostics, and rapid antimicrobial treatment.

CONCLUSIONS

Spirochetes are a phylogenetically ancient and distinct group of microorganisms. Because of their propensity to cause diseases in humans, *B. burgdorferi* and *T. pallidum* are the two most-well-studied spirochetes. However, the inability to culture *T. pallidum in vitro* has made researching this spirochete very difficult, and as a consequence, the immune response that follows infection with this spirochete remains poorly understood.

The etiological agents of Lyme disease and syphilis are similar in that they have relatively small genomes, surviving only in association with a host and eliciting inflammatory disease, but genomic comparison clearly shows that *T. pallidum* and *B. burgdorferi* are not closely related. It seems likely that despite some similarities, these spirochetes evolved independently from a more complex ancestor, resulting in differences in their life-cycles, environmental adaptations, and the pathology associated with their infection. Therefore it is not surprising to learn of the differences in the host immune responses to *B. burgdorferi* and *T. pallidum*. Both the host response to these spirochetes and their infectivity determine the extent of pathology following infection.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Burgdorfer W, Barbour AG, Hayes SF, et al. Lyme disease-a tick-borne spirochetosis? Science 1982;216:1317–19.
2. Krause RM. Metchnikoff and syphilis research during a decade of discovery, 1900-1910. Development of an animal model and a preventive treatment set the stage for progress. ASM News 1996;62:307–10.
3. Fraser CM, Casjens S, Huang WM, et al. Genomic sequence of a Lyme disease spirochaete, *Borrelia burgdorferi*. Nature 1997; 390:580–6.
4. Fraser CM, Norris SJ, Weinstock GM, et al. Complete genome sequence of *Treponema pallidum,* the syphilis spirochete. Science 1998;281:375–88.
5. Anguita J, Hedrick M, Fikrig E. Adaptation of *Borrelia burgdorferi* in the tick and the mammalian host. FEMS Microbiol Rev 2003;27:493–504.
6. Pal U, Li X, Wang T, et al. TROSPA, an Ixodes scapularis receptor for *Borrelia burgdorferi*. Cell 2004;119:457–68.
7. Hubner A, Yang X, Nolen DM, et al. Expression of *Borrelia burgdorferi* OspC and DbpA is controlled by a RpoN-RpoS regulatory pathway. Proc Natl Acad Sci USA 2001;98:12724–9.
8. Center for Disease Control and Prevention. 2016. Available at: https://www.cdc.gov/lyme/stats/index.html.
9. Guerau-de-Arellano M, Huber BT. Chemokines and Toll-like receptors in Lyme disease pathogenesis. Trends Mol Med 2005;11:114–20.
10. Hawley KL, Olson CM Jr, Iglesias-Pedraz JM, et al. CD14 cooperates with complement receptor 3 to mediate MyD88-independent phagocytosis of *Borrelia burgdorferi*. Proc Natl Acad Sci USA 2012;109:1228–32.
11. Lawrenz MB, Wooten RM, Zachary JF, et al. Effect of complement component C3 deficiency on experimental Lyme borreliosis in mice. Infect Immun 2003;71:4432–40.
12. Kochi SK, Johnson RC. Role of immunoglobulin G in killing of *Borrelia burgdorferi* by the classical complement pathway. Infect Immun 1988;56:314–21.
13. Madar M, Bencurova E, Mlynarcik P, et al. Exploitation of complement regulatory proteins by *Borrelia* and *Francisella*. Mol Biosyst 2015;11:1684–95.
14. Pausa M, Pellis V, Cinco M, et al. Serum-resistant strains of *Borrelia burgdorferi* evade complement-mediated killing by expressing a CD59-like complement inhibitory molecule. J Immunol 2003;170:3214–22.
15. Hovis KM, McDowell JV, Griffin L, et al. Identification and characterization of a linear-plasmid-encoded factor H-binding protein (FhbA) of the relapsing fever spirochete *Borrelia hermsii*. J Bacteriol 2004;186:2612–18.
16. Garcia BL, Zhi H, Wager B, et al. *Borrelia burgdorferi* BBK32 inhibits the classical pathway by blocking activation of the C1 complement complex. PLoS Pathog 2016;12:e1005404.
17. Gross DM, Steere AC, Huber BT. T helper 1 response is dominant and localized to the synovial fluid in patients with Lyme arthritis. J Immunol 1998;160:1022–8.
18. Zhang JR, Hardham JM, Barbour AG, et al. Antigenic variation in Lyme disease borreliae by promiscuous recombination of VMP-like sequence cassettes. Cell 1997;89:275–85.
19. Ramamoorthi N, Narasimhan S, Pal U, et al. The Lyme disease agent exploits a tick protein to infect the mammalian host. Nature 2005;436:573–7.
20. Cox DL, Chang P, McDowall AW, et al. The outer membrane, not a coat of host proteins, limits antigenicity of virulent *Treponema pallidum*. Infect Immun 1992;60:1076–83.
21. Salazar JC, Hazlett KR, Radolf JD. The immune response to infection with *Treponema pallidum,* the stealth pathogen. Microbes Infect 2002;4:1133–40.
22. Salazar JC, Cruz AR, Pope CD, et al. *Treponema pallidum* elicits innate and adaptive cellular immune responses in skin and blood during secondary syphilis: a flow-cytometric analysis. J Infect Dis 2007;195:879–87.
23. Blanco DR, Champion CI, Lewinski MA, et al. Immunization with *Treponema pallidum* outer membrane vesicles induces high-titer complement-dependent treponemicidal activity and aggregation of *T. pallidum* rare outer membrane proteins (TROMPs). J Immunol 1999;163(5):2741–6.

24. Blanco DR, Champion CI, Dooley A, et al. A monoclonal antibody that conveys in vitro killing and partial protection in experimental syphilis binds a phosphorylcholine surface epitope of *Treponema pallidum*. Infect Immun 2005;73:3083–95.
25. McBroom RL, Styles AR, Chiu MJ, et al. Secondary syphilis in persons infected with and not infected with HIV-1: a comparative immunohistologic study. Am J Dermatopathol 1999;21:432–41.
26. Van Voorhis WC, Barrett LK, Koelle DM, et al. Primary and secondary syphilis lesions contain mRNA for Th1 cytokines. J Infect Dis 1996;173:491–5.
27. Centurion-Lara A, Castro C, Barrett L, et al. *Treponema pallidum* major sheath protein homologue Tpr K is a target of opsonic antibody and the protective immune response. J Exp Med 1999;189:647–56.
28. Giacani L, Molini BJ, Kim EY, et al. Antigenic variation in *Treponema pallidum:* TprK sequence diversity accumulates in response to immune pressure during experimental syphilis. J Immunol 2010;184:3822–9.

MULTIPLE-CHOICE QUESTIONS

1. The major immunogens of *Borrelia burgdorferi* and *Treponema pallidum* are:
   - **A.** Lipopolysaccharide molecules present on the outer membrane
   - **B.** Sugar coated-proteins that decorate their surface
   - **C.** A group of lipoproteins that constitute Toll-like receptor (TLR)1/TLR2 ligands
   - **D.** Both spirochetes, which are considered "stealth pathogens" because of their lack of immunogens

2. The phagocytosis of spirochetes:
   - **A.** Allows the elimination of the bacteria before they have a chance to induce a proinflammatory response
   - **B.** Allows the interaction of pathogen recognition receptors with their ligands within the phagolysosome t modulate the inflammatory response
   - **C.** Depends exclusively on the presence of opsonins that target the bacteria to professional antigen-presenting cells
   - **D.** Occurs only on tissues where the pathogens reside, exclusively upon activation of the adaptive immune system

3. The generation of a protective antibody response against *B. burgdorferi:*
   - **A.** Occurs even in the absence of T-cell responses
   - **B.** Is targeted exclusively to variable regions of the vlsE protein
   - **C.** Is always accompanied by a T-helper type 2 response
   - **D.** Occurs within the first 3 days of infection and is the basis for the successful diagnosis of patients

4. T-cell responses against *T. pallidum:*
   - **A.** Are always weak because of the low level of infection in patients
   - **B.** Induce a type I delayed hypersensitivity response that helps the elimination of the bacteria
   - **C.** Are always composed of CD4 T cells that exclusively aid the development of antibody responses
   - **D.** Can be used as diagnostic and prognostic tools for the disease

29

Host Defenses to Fungal Pathogens

Allison K. Lord, Jatin M. Vyas

Advances in modern medicine have significantly improved the prognosis and quality of life for patients with a wide spectrum of diseases, including many forms of cancer. However, opportunistic fungal pathogens have exploited these modern immunosuppressive and invasive medical interventions. Invasive fungal infections (IFIs) represent a significant cause of morbidity and mortality among immunocompromised patients. Patients at highest risk for IFIs include solid organ or hematopoietic stem cell transplant (HSCT) recipients on intensive immunosuppressive regimens, patients with hematological malignancies, and other patients that are immunocompromised as a result of various clinical conditions and treatments (*e.g.,* human immunodeficiency virus/acquired immunodeficiency syndrome [HIV/AIDS], advanced age, recent surgery, etc.) (Fig. 29.1).[1,2] The clinical consequences of these pathogenic fungi include superficial disease, allergic disease, and IFIs.

KEY CONCEPTS

Invasive Fungal Infections

- Invasive fungal infections are systemic life-threatening infections that are estimated to cause 1.5 million deaths annually.
- *Aspergillus, Candida, Cryptococcus,* and *Pneumocystis* are opportunistic fungi that cause the majority of invasive fungal infections.
- Risk factors for invasive fungal infections include bone marrow or solid organ transplantation, intensive immunosuppressive regimens, hematological malignancies, HIV/AIDS, advanced age, recent surgery, and other clinical conditions and treatments that cause immunosuppression.

Superficial fungal infections of the skin and nails affect 25% of the global population (≈1.7 billion people) and give rise to various conditions, including athlete's foot and ringworm of the scalp. Fungi also commonly cause mucosal infections of the oral and genital tracts; vulvovaginal candidiasis occurs at least once in 50–75% of women in their childbearing years.[3,4] Superficial fungal infections of the skin and mucosa can become chronic, but they are rarely life threatening.

IFIs occur when fungal pathogens invade the bloodstream, resulting in a systemic life-threatening infection that affects multiple organs. IFIs are estimated to cause 1.5 million deaths annually, with four genera accounting for most of the infections: *Cryptococcus, Candida, Aspergillus,* and *Pneumocystis.* Mortality resulting from infections with these fungal organisms exceeds that caused by tuberculosis or malaria.[4] As a result, IFIs have emerged as an escalating and worrisome clinical problem. There is an urgent need to define the mechanisms of host defense against fungal pathogens with the goal of developing novel therapeutics as well as improving diagnostic and preventative procedures.

One of the greatest limiting factors in the treatment of IFIs is lack of effective and prompt diagnostic tools. Although *Candida* is capable of growing in conventional blood cultures, most other invasive fungal organisms, including *Cryptococcus, Aspergillus,* and *Histoplasma*, are typically not recovered by adding blood to growth media. There has been a heavy reliance on clinical diagnosis and a few biomarkers. Fungal cell wall carbohydrates can be detected in the bloodstream of some patients. β-1,3 glucan, galactomannan, and glucuronoxylomannan (GXM) are three carbohydrates that can be analyzed clinically. β-1,3 glucan is a common fungal cell wall carbohydrate found in *Candida albicans* and many other fungi (with the notable exception of *Cryptococcus neoformans*). Galactomannan is typically associated with *Aspergillus fumigatus.* This fungal biomarker can be measured in blood and bronchoalveolar lavage fluid. GXM is an abundant fungal-derived carbohydrate produced by *C. neoformans.* GXM can also be found on the surface of *Candida gattii,* a closely related fungal organism that can cause disease in immunocompetent individuals. Although GXM appears to be a useful biomarker for *Cryptococcus,* its widespread use is limited by technical and economic issues. Finally, a urine test for histoplasmosis has been established. Again, the antigen for this test is a polysaccharide derived from the cell wall of this fungal organism. Many of these biomarkers suffer from poor sensitivity (galactomannan) and, in some cases, poor specificity (*e.g.,* testing urine for *Histoplasma* antigen).

Current therapeutic approaches for IFIs include the prompt institution of antifungal agents (including polyenes, azoles, and echinocandins) and even more important, the reversal of underlying host immune defects, such as neutropenia or high doses of immunosuppressive therapy. The first-line therapy for the treatment of invasive aspergillosis (IA) is voriconazole, an extended-spectrum azole. Despite its *in vitro* efficacy, voriconazole only demonstrated a survival rate at 12 weeks of 70.8% compared with 57.9% in the amphotericin B group.[5] Although voriconazole has become the gold standard for treatment, its high mortality rate indicates that the host immune response plays a critical role in determining host outcome. Despite clinical vigilance, prevention, diagnosis, and treatment of IFIs remain a significant challenge because of lack of rapid and reliable diagnostic methods and limited treatment options. Discovery of novel diagnostic tools and antifungal therapies is crucial to ameliorating this public health burden.

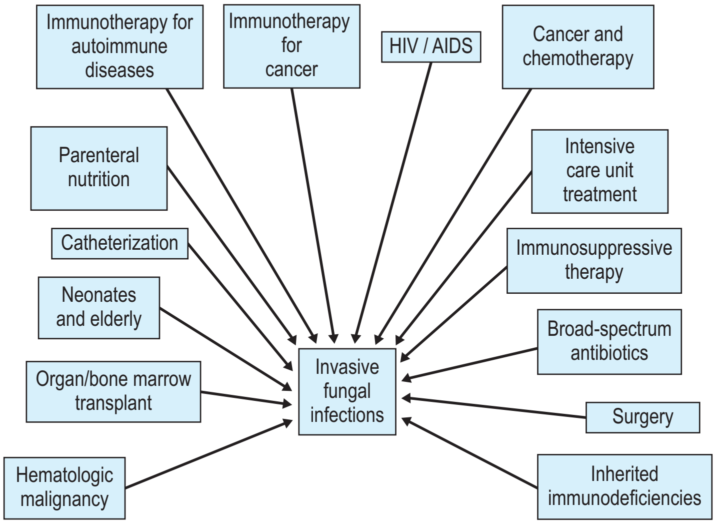

Immunotherapy for autoimmune diseases
Immunotherapy for cancer
HIV / AIDS
Cancer and chemotherapy
Parenteral nutrition
Intensive care unit treatment
Catheterization
Immunosuppressive therapy
Neonates and elderly
Broad-spectrum antibiotics
Invasive fungal infections
Organ/bone marrow transplant
Surgery
Hematologic malignancy
Inherited immunodeficiencies

FIG 29.1 Major risk factors for developing invasive fungal infections. (Adapted from Karkowska-Kuleta J, Kozik A. Cell wall proteome of pathogenic fungi. Acta Biochim Pol 2015; 62: 339–51.)

CLINICALLY RELEVANT FUNGAL ORGANISMS

Although it is estimated that there are in excess of 5 million distinct species of fungi, only a handful are considered significant to human health: *Aspergillus* spp., *Candida* spp., and *Cryptococcus* spp. cause the majority of IFIs in the United States and Europe. Invasive infections caused by these pathogens in immunocompromised patients are serious and often life threatening.

Aspergillus fumigatus

Aspergillus fumigatus is a ubiquitous airborne mold found in soil, air, food, and decaying organic material. Humans routinely inhale *A. fumigatus* conidia, but the microorganism is rapidly eliminated by the innate immune system in immunocompetent individuals. In immunocompromised patients, *A. fumigatus* can cause IA, a severe and usually fatal infection. It is estimated that IA occurs in 200 000 patients annually.[4] In cases of widespread infection, IA mortality rates approach 90%, with favorable responses to antifungal therapy observed in <30% of patients.[5] In particular, IA has emerged as a significant cause of mortality among patients who have undergone hematopoietic stem cell transplantation and solid organ transplantations. Advanced age, respiratory viruses, graft-versus-host disease (GvHD), prolonged glucocorticosteroid use, cytomegalovirus (CMV) infection, and iron overload are also common risk factors for IA.[6] Given that *A. fumigatus* spores are inhaled, IA most frequently manifests as an invasive pulmonary disease, but the pathogen can enter the bloodstream and infect multiple organs. Less frequently, IA affects skin, sinuses, and the central nervous system (CNS). Clinical characteristics of IA vary, depending on the organ(s) affected.

Aspergillus can cause a condition called *chronic pulmonary aspergillosis* (CPA), which causes progressive destruction of lung tissue. CPA is most common in patients with underlying lung disease (*e.g.,* asthma) and affects 3 million people worldwide. CPA is estimated to have 15% mortality, often as a result of a massive pulmonary hemorrhage within the first 6 months of diagnosis.[4]

Candida albicans

Candida species are ubiquitous commensal fungi that normally colonize human mucosa and skin. There are over 165 species of *Candida,* but only a few are known to cause human disease. *C. albicans* accounts for the majority of infections, followed by *C. glabrata.* Less frequently, *C. parapsilosis, C. tropicalis, C. krusei,* and others cause invasive candidiasis.[7] In immunocompromised individuals, *Candida* can become pathogenic and infect host tissues. For example, *Candida* is commonly associated with skin infections, as well as oral and vaginal thrush. However, *Candida* can become invasive (invasive candidiasis), entering the bloodstream (candidemia) and other organs (*e.g.,* bladder/kidneys, liver, spleen, gut, etc.), resulting in a life-threatening infection. Despite the availability of antifungal therapy, the crude mortality rate for candidemia exceeds 50%, indicating that the immune system is a necessary partner for successful treatment of this infection.[8] Every year, it is estimated that invasive candidiasis affects 250 000 people worldwide and causes more than 50 000 deaths.[9] As with other opportunistic fungal infections, patients on immunosuppressive regimens (*i.e.,* transplant recipients) and patients with hematological cancers are at high risk for invasive candidiasis. Other major risk factors include recent surgery, broad-spectrum antibiotic therapy, and central vascular catheters. Disruptions to the balance of *C. albicans* in the gut are associated with increasing severity of Crohn disease and ulcerative colitis (UC). Moreover, disseminated candidiasis may originate in the gut.

Cryptococcus neoformans

During the 1980s to the 1990s, cryptococcosis emerged as a major cause of morbidity and mortality among persons with HIV/AIDS. The majority of cryptococcosis cases are caused by *C. neoformans,* a fungus that is widely distributed in soil. The major route of exposure to *C. neoformans* is inhalation of airborne organisms into the lungs. In healthy individuals, the immune system is effective at clearing the pathogen, but in immunocompromised individuals (especially those with HIV/AIDS), *C. neoformans* can cause an IFI. Interestingly, *C. neoformans* targets the CNS, causing meningoencephalitis, a life-threatening condition.

HOST DEFENSE AGAINST FUNGI BY EPITHELIAL CELLS

Given the ubiquity of many fungal organisms, humans inhale a diverse species of fungi regularly. Moreover, the microbiota has a profound influence on human health. Early work on this complex collection of microorganisms focused on bacteria, but increasing evidence indicates that the mycobiome plays an important role in human health. Indeed, sequencing data from the respiratory sputum of normal volunteers showed over 40 distinct species of fungi.[10] Interestingly, sputum from patients with allergies demonstrated a distinct mycobiome. Although there is an association, it remains unclear as to whether specific species of fungi are causal for allergic responses. Growth of commensal fungi is limited by release of antimicrobial peptides and colonization by other microbial flora that compete for nutrients. Disruptions to this delicate balance can have serious pathological consequences.

Skin and the epithelial mucosa of the respiratory, gastrointestinal (GI), and genitourinary tracts represent the first line of defense against fungal pathogens, acting as a physical barrier against infection. There is evidence that epithelial cells provide more than just barrier function by triggering an immune response. Several groups have demonstrated that oral epithelial cells infected with fungi produce proinflammatory cytokines and chemokines *in vitro.*[11,12] Similarly, stimulation of human corneal epithelial cells with *A. fumigatus* hyphae activates Syk signaling and leads to production of inflammatory cytokines (interleukin-1β [IL-1β] and IL-6) and chemokines (IL-8 and CXCL1).[13] Hyphal formation is critical for inducing a strong response by epithelial cells.[14] Recently, a fungal toxin termed *candidalysin* has been identified that is a cell permeable peptide that induces a strong inflammatory response within epithelial cells. Production of candidalysin by *C. albicans* is required for T-helper 17 (Th17) responses.

C. albicans is a commensal fungus in the human GI tract, but its abundance is limited by intestinal bacteria. Antibiotics that disrupt the balance of bacterial microbiota favor overgrowth of *Candida* species. It is thought that patients with fever and neutropenia who develop invasive candidemia result from translocation of fungi from the GI tract. Emerging data connect gut homeostasis and immune responses to fungi in the lung.[15] Many questions remain about how epithelial cells discriminate fungal morphology and which proteins mediate epithelial recognition and activation in response to pathogenic fungi. Compared with host defense mechanisms exerted by myeloid cells, much less is known about host–pathogen interactions in epithelial mucosa.

THE FUNGAL CELL WALL

The fungal cell wall contains many of the relevant pathogen-associated molecular patterns (PAMPs) and epitopes for the immune response. PAMPs are recognized by cells of the host innate immune system and are often targets for antifungal agents.

The ultrastructure of fungal organisms is similar to mammalian cells. This feature has thwarted the development of a broad armamentarium of antifungal agents. However, fungi also possess a cell wall, a structure not found in mammalian cells. The fungal cell wall not only provides the organism with mechanical strength to withstand diverse environmental conditions but also maintains plasticity to permit cell growth and division and formation of different cell types throughout the life cycle of a fungal organism. Availability of nutrients, stress, hypoxia, and hypercarbia are environmental cues to the modification of the cell wall. The major constituents of the fungal cell wall are chitin, glucans, and glycoproteins. Chitin is a structurally important component of the fungal cell wall located closest to the plasma membrane. The composition of the outer layer varies, depending on the fungal species, morphotype, and growth stage. Branched β-1,3 glucan cross-links to chitin and is covalently linked to other polysaccharides (*e.g.,* galactomannan and β-1,6 glucan). The fungal cell wall composition differs dramatically between conidia and hyphae, which leads to differential recognition and response by the immune system. Many fungal species produce melanin, a natural pigment found on the cell wall and is associated with the enhanced virulence of many pathogenic fungi. Melanin protects the organism from both environmental stressors and interferes with host defense mechanisms. Specifically, fungal melanin has been associated with decreased phagocytosis, an altered cytokine response, and reduces susceptibility to antifungal drugs.[16] In some organisms, such as *A. fumigatus,* a hydrophobic rodlet layer composed of hydrophobins makes the conidia waterproof and shields underlying carbohydrates that may induce inflammation.

Fungal dimorphism is a critical feature of virulence for both yeast and molds. *C. albicans* exist in three forms that have distinct shapes: blastospores (also known as *yeast cells*), pseudohyphal cells, and true hyphal cells. Yeast cells are typically 5 μm in diameter with a round to ovoid shape. Pseudohyphae resemble elongated yeast cells that remain attached to one another. These fungi usually grow in a branching pattern that may facilitate scavenging for nutrients at distal sites. True hyphal cells are long and highly polarized, with parallel sides and no obvious constrictions between cells. Similarly, *A. fumigatus* conidia are waxy hydrophobic ovoid-shaped cells of approximately 3–5 μm in diameter. *A. fumigatus* can filament to generate hyphae that are >40 μm in length. For both *C. albicans* and *A. fumigatus,* this dimorphism is a virulence factor as organisms that fail to do so are avirulent.

THE INNATE IMMUNE RESPONSE TO FUNGAL PATHOGENS

Once a fungal pathogen has invaded the host system, innate immune cells, such as macrophages/monocytes, neutrophils, and dendritic cells (DCs), phagocytose and degrade the organism (Chapter 3). In immunocompetent individuals, physical barriers (*e.g.,* epithelial cells, mucous, skin) are effective at preventing infection. In the respiratory tract, macrophages routinely neutralize fungal pathogens that make their way to the alveoli. Understanding the rules that govern the recognition and response to these fungal pathogens provide important insights into how these organisms cause disease in immunocompromised individuals.

Recognition of fungal pathogens by innate immune cells is mediated by pattern recognition receptors (PRRs). Engagement of PRRs and PAMPs triggers a cascade of molecular events that coordinates phagocytosis and degradation of the pathogen. Additionally, release of cytokines, reactive oxygen species (ROS) production, and presentation of fungal antigens to the adaptive immune system aids in controlling infection.

KEY CONCEPTS

Innate Immunity to Fungal Pathogens

- Fungi express highly conserved pathogen-associated molecular patterns (PAMPs) that are recognized by pathogen recognition receptors (PRRs) expressed on host phagocytes.
- Dectin-1 is a C-type lectin receptor (CLR) that acts as a PRR to recognize β-1,3 glucan expressed on the cell walls of *Candida, Aspergillus,* and other fungal pathogens.
- Toll-like receptors (TLRs) recognize fungal cell wall components and nucleic acids.
- Engagement of PRRs and PAMPs initiates signaling that coordinates secretion of cytokines, reactive oxygen species (ROS) production, and presentation of fungal antigens to the adaptive immune system to facilitate elimination of the pathogen.

Recently engulfed pathogens become enclosed in membrane delineated compartments called *phagosomes,* which traffic toward the lysosome, as directed by modifications to the phagosomal membrane proteins and changes to the intraphagosomal environment. Proteins recruited to the phagosomal membrane are specific to its contents. Phagolysosomes intersect with class II major histocompatibility complex (MHC) molecules and permit loading of pathogen-specific peptides. T cells are then activated in an antigen-specific manner to augment the immune response and generate long-term immunity.

Role of Neutrophils

Neutrophils are the most critical cell in the host defense against fungal pathogens. Patients with neutropenia or acquired defects in neutrophil-mediated killing are at higher risk of developing IFIs, including IA. Although phagocytosis is a critical feature in controlling infection, other mechanisms also limit the spread of infection and serve to kill invading organisms. Neutrophils rely on multiple mechanisms for killing, including granule-dependent killing, nicotinamide adenine dinucleotide phosphate (NADPH) oxidase–dependent killing, and neutrophil extracellular trap (NET) formation (Chapter 22).

Phagocytosis of fungal organisms by neutrophils triggers production of antimicrobial ROS. Generation of ROS by the NADPH oxidase complex is a result of engagement of surface receptors, including dectin-1, and signaling adaptors, including caspase recruitment domain family member 9 (CARD9) and Syk. Loss of ROS production has a clear phenotype in *Aspergillus* infections. Patients with chronic granulomatous disease (CGD) fail to control *Aspergillus* hyphae, which leads to invasion and metastatic spread of disease. In contrast, conidial forms of *Aspergillus* do not require ROS. The data for *Candida* infections are less compelling. There appears to be modest reduction in killing of serum-opsonized *C. albicans,* but killing of unopsonized organisms is unimpaired.

In addition to cytotoxic granules and ROS production, neutrophils are capable of NET formation. This relatively newly discovered cytotoxic mechanism delivers a web-like structure composed of chromatin and histones. For *Candida,* it appears that NET formation contributes to fungal killing.[17] However, in *Aspergillus,* data from studies do not support a clear role for NETs.

Role of Macrophages

Macrophages also play a critical role in neutralizing fungal organisms. These phagocytic cells are much longer lived than neutrophils. Macrophages are recruited to the site of infection by chemotaxis. They are also responsible for patrolling interfaces with the environment, including the lung and the GI tract. In contrast to neutrophils, macrophages express class II MHC and can activate T cells. Upon activation by fungi through PRRs and PAMPs, macrophages produce potent inflammatory cytokines, including tumor necrosis factor α (TNF)-α and IL-6. Engagement of the inflammasome within these cells leads to copious production of IL-1β. Macrophages receive assistance from activated T cells that secrete interferon-γ (IFN-γ), which can activate a host of genes to improve the antifungal response from macrophages.

Role of Dendritic Cells

In contrast to macrophages, DCs can stimulate naïve T cells. Moreover, DCs are capable of taking up both yeast and conidia. DC subsets are characterized by expression of specific surface markers. Plasmacytoid DCs (pDCs) expressing specific surface markers phagocytose *A. fumigatus* conidia and spread over hyphae. pDCs incite an immune response by release of proinflammatory cytokines, including IFN-α and TNF-α. pDCs are typically found in the spleen but will migrate to the lungs following challenge by *A. fumigatus* conidia. Depletion of pDCs has been shown to result in increased mortality, indicating a nonredundant role for pDCs in host defense against *A. fumigatus.*

C. albicans epithelial infections appear to recruit pDCs. These cells reorganize and become more concentrated in the T-cell zones of lymph nodes. Draining lymph nodes of mice infected with *Candida* versus control mice showed that pDCs were the predominant DC subset after vaginal infection. pDCs are involved in the induction of Th1 responses to *C. albicans* vaginal infection.

PATTERN RECOGNITION RECEPTORS

Recognition of fungal pathogens is mediated by PRRs expressed by innate immune cells, including DCs and myeloid cells (*e.g.,* macrophages, monocytes, and neutrophils). The interaction between PRRs and PAMPs expressed on the fungal cell wall triggers downstream signaling, which elaborates the host immune response and facilitates elimination of the pathogen. The major PRRs involved in antifungal immunity are Toll-like receptors (TLRs) and C-type lectin receptors (CLRs). Other relevant PRRs include nucleotide-binding oligomerization domain (NOD)–like receptors (NLRs), and retinoic acid-inducible gene I (RIG-I)–like receptors (RLRs).

Toll-Like Receptor

Toll-like receptors (TLRs) are transmembrane receptors that recognize a broad range of microbial ligands, including fungal and bacterial cell wall components, bacterial and viral nucleic acids, and bacterial lipoproteins (Fig. 29.2) (Chapter 3). TLR1, -2, -4, -5, -6, and -10 are expressed on the cell surface, whereas TLR3, -7, -8, and -9 are expressed on intracellular membranes. Intracellular TLRs recognize nucleic acids derived from fungi, bacteria, and viruses and signal through MyD88 or TRIF. MyD88 signaling triggered by TLRs is essential for host fungal defense; mice lacking MyD88 are more susceptible to IFIs. However, humans with MyD88 signaling defects do not have increased incidence of fungal infections. To date, several TLRs have been implicated in fungal immunity in mice, including TLR2, TLR4, TLR7, and TLR9.[18] The importance of TLRs has been demonstrated in human biology. In humans, polymorphisms in TLR1, TLR3, TLR4, and TLR6 are associated with increased susceptibility to IFIs (especially IA).[19]

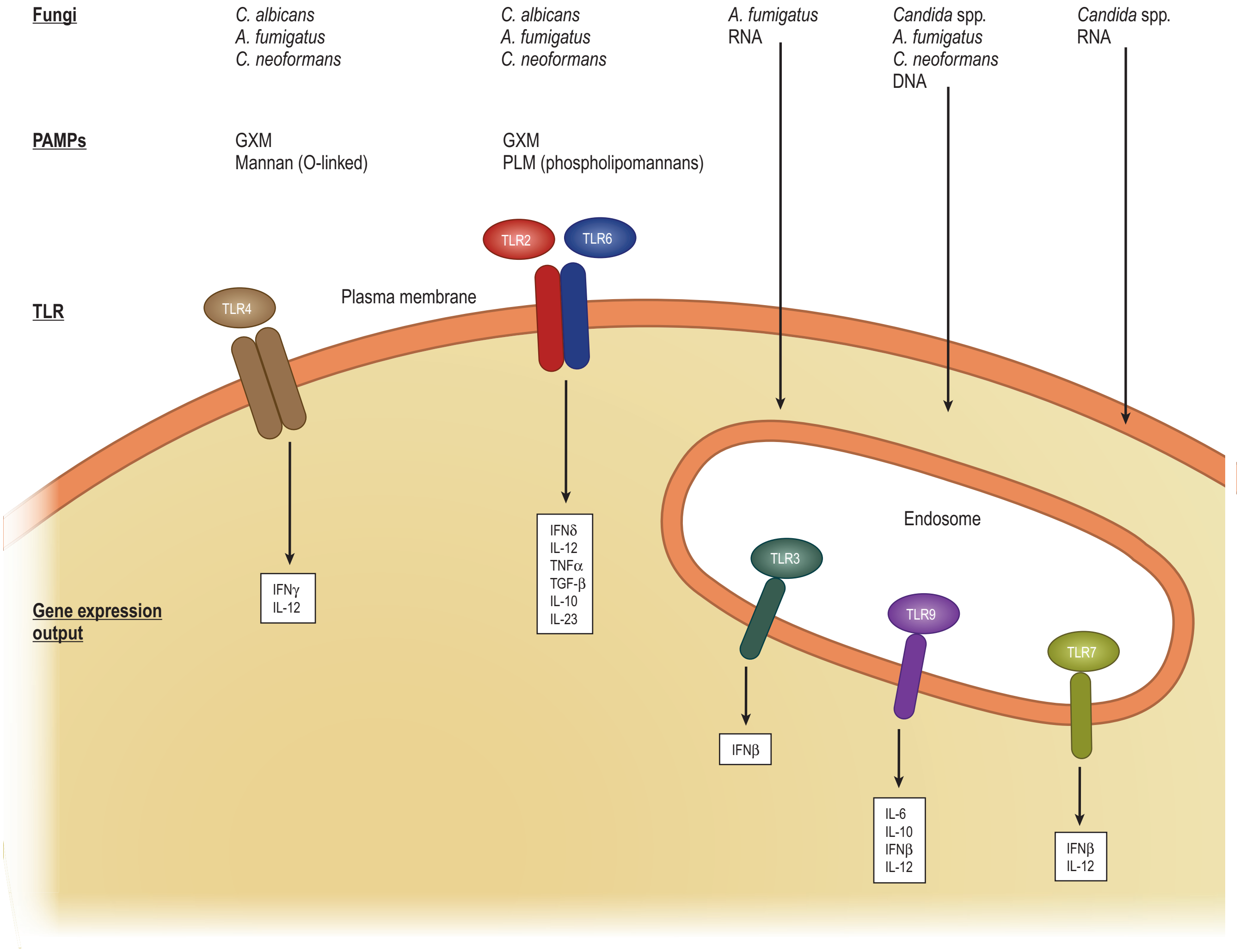

Fungi
C. albicans
A. fumigatus
C. neoformans
C. albicans
A. fumigatus
C. neoformans
A. fumigatus
RNA
Candida spp.
A. fumigatus
C. neoformans
DNA
Candida spp.
RNA
PAMPs
GXM
Mannan (O-linked)
GXM
PLM (phospholipomannans)
TLR
TLR2
TLR6
TLR4
Plasma membrane
Endosome
TLR3
TLR9
TLR7
Gene expression output
IFNγ
IL-12
IFNδ
IL-12
TNFα
TGF-β
IL-10
IL-23
IFNβ
IL-6
IL-10
IFNβ
IL-12
IFNβ
IL-12

FIG 29.2 **Toll-Like Receptors (TLRs) and Fungal Immunity.** Surface and endosomal TLRs recognize fungal pathogen-associated molecular patterns (PAMPs) resulting in downstream signaling that promotes production of cytokines. (Adapted from Bourgeois C, Kuchler K. Fungal pathogens-a sweet and sour treat for Toll-like receptors. Front Cell Infect Microbiol 2012;2:142.)

In contrast, TLR9 polymorphisms are more tightly associated with allergic bronchopulmonary aspergillosis.

Surface-disposed TLR2 recognizes β-glucans of several fungal species (*C. albicans, A. fumigatus,* and *C. neoformans*). TLR2 also interacts with phospholipomannans and linear β-1,2-oligomannoside structures found on the cell wall of *C. albicans.* TLR2 can heterodimerize with TLR1 or TLR6 to recognize GXM, a cell wall component of *C. neoformans.* TLR2 is critical for early recruitment and the killing abilities of neutrophils.[19] TLR2-deficient mice infected with *Pneumocystis jiroveci* display more severe symptoms, increased fungal burden, and less TNF-α production.

TLR4 recognizes O-linked mannan expressed on the cell wall of *C. albicans* and through downstream signaling, stimulates production of TNF-α. Mice deficient in TLR4 do not appear to have increased susceptibility to disseminated candidiasis compared with wild-type mice. In contrast, killing of *Aspergillus* is impaired in $TLR4^{-/-}$ mice in a corneal inflammation model, although recruitment of immune cells to the infection, remains unaffected.

TLR9 recognizes intracellular microbial ligands, such as CpG-rich fungal DNA. Dectin-1–dependent Syk activation triggers phagosomal acidification, which permits cleaved TLR9 to traffic to the phagosomes. Moreover, dectin-1 regulates TLR9-dependent changes in gene expression.[20] Interestingly, TLR9 deficient macrophages have increased fungicidal activity and production of antiinflammatory cytokines. Thus TLR9 appears to modulate the inflammatory response by downregulating cytokine production.[21]

C-Type Lectin Receptor

C-type lectin receptors (CLRs) comprise a diverse family of proteins characterized by a conserved C-type lectin domain (CTLD) that recognizes a variety of ligands. CLRs that are recognized to play a role in fungal immunity include dectin-1, dectin-2, mannose receptor, Mincle, and dendritic cell–specific intercellular adhesion molecule–grabbing nonintegrin (DC-SIGN) (Fig. 29.3). However, dectin-1 is the best-characterized CLR associated with fungal immunity. Dectin-1 is a type II transmembrane protein expressed primarily on myeloid cells. It is considered the major mammalian cell surface receptor for β-1,3 glucan and β-1,6 glucan, carbohydrates widely expressed on the cell wall of many fungal organisms. Furthermore, dectin-1

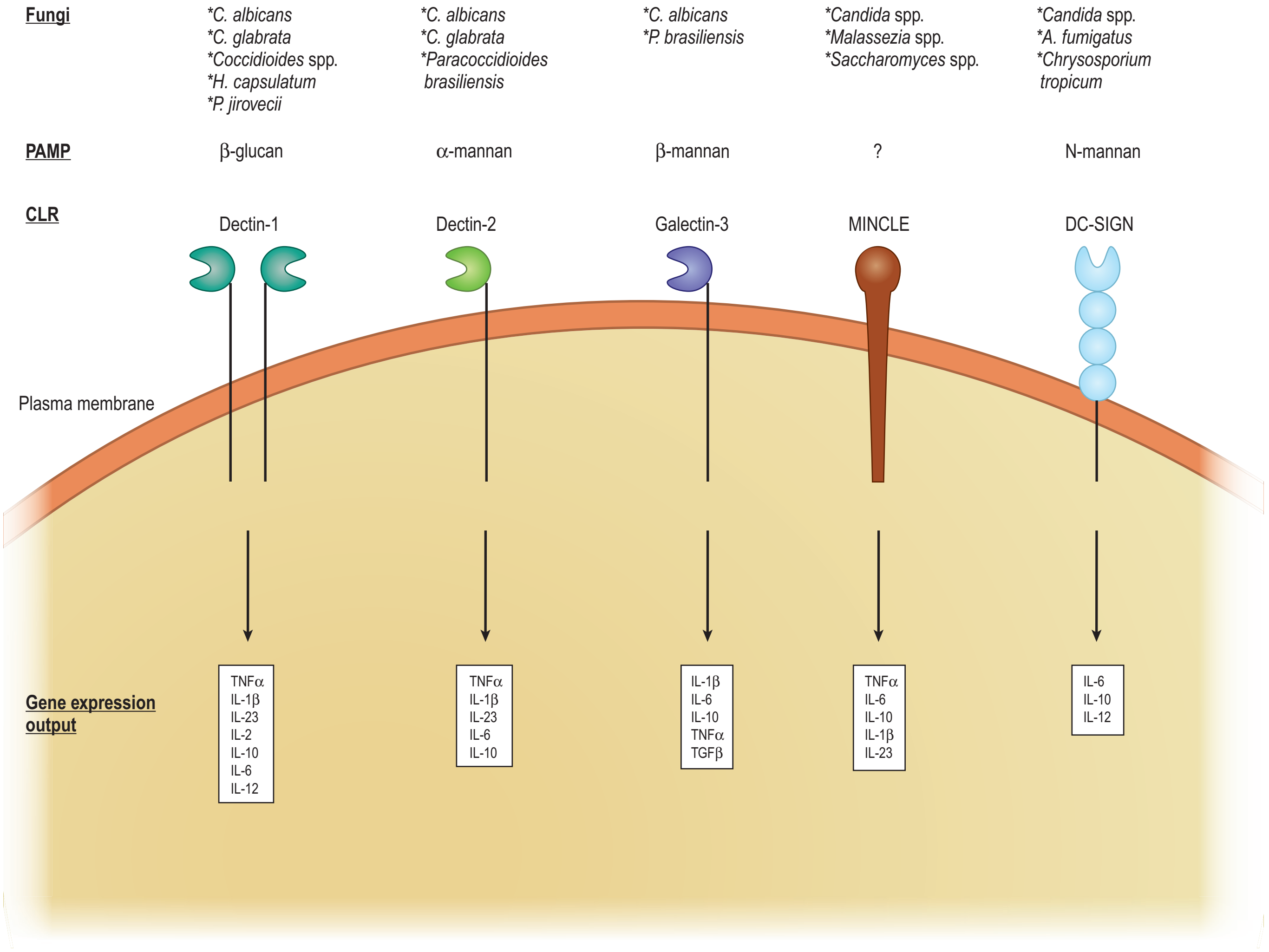

Fungi
*C. albicans
*C. glabrata
*Coccidioides spp.
*H. capsulatum
*P. jirovecii
*C. albicans
*C. glabrata
*Paracoccidioides brasiliensis
*C. albicans
*P. brasiliensis
*Candida spp.
*Malassezia spp.
*Saccharomyces spp.
*Candida spp.
*A. fumigatus
*Chrysosporium tropicum
PAMP
β-glucan
α-mannan
β-mannan
?
N-mannan
CLR
Dectin-1
Dectin-2
Galectin-3
MINCLE
DC-SIGN
Plasma membrane
Gene expression output
TNFα
IL-1β
IL-23
IL-2
IL-10
IL-6
IL-12
TNFα
IL-1β
IL-23
IL-6
IL-10
IL-1β
IL-6
IL-10
TNFα
TGFβ
TNFα
IL-6
IL-10
IL-1β
IL-23
IL-6
IL-10
IL-12

FIG 29.3 **C-Type Lectin Receptors (CLRs) and Fungal Immunity.** CLRs recognize fungal PAMPs and trigger downstream signaling that leads to cytokine production.

promotes acidification and maturation of phagosomes.[22] Patients with mutations in dectin-1 are more susceptible to IFIs, which demonstrates the essential role of dectin-1 in mediating fungal immunity.[23]

Upon dectin-1 binding to β-1,3 glucan, a tyrosine residue within the cytoplasmic hemi–immunoreceptor tyrosine-based activation motif (ITAM) domain of dectin-1 is phosphorylated by Src family kinases, which leads to the recruitment and activation of Syk. Syk triggers the activation of the caspase recruitment domain family member 9–Bcl-10– mucosa-associated lymphoreticular tissue 1 (MALT1) complex, inducing nuclear factor (NF)-κB and production of cytokines, including IL-1β, IL-6, IL-10, IL-112, IL-23, and TNF-α. Dectin-1 can signal through a Syk-independent pathway via activation of the serine/threonine kinase, Raf-1, that phosphorylates p65, a subunit of NF-κB.[24] Cytokines produced in response to canonical (via Syk) and noncanonical (via Raf-1) NF-κB signaling are essential for the activation of Th1 and Th17 cells (Fig. 29.4).

Syk also triggers recruitment and activation of NADPH oxidase, which stimulates production of ROS, elaboration of proinflammatory cytokines, and ultimately fungal killing. Syk appears to be a master controller for phagosomal maturation and recruitment of light chain 3 (LC3), a protein associated with autophagy.

Collaboration Between TLRs and CLRs

Although most cognate ligands found on the fungal cell wall are carbohydrates that preferentially trigger CLRs, experimental evidence indicates that CLRs can collaborate with TLRs on the cell surface to coordinate cytokine production. Dectin-1 can collaborate with TLR2 and TLR4 to induce synergistically cytokines, including TNF-α, IL-10, and IL-23. This collaboration has been used clinically for management of clinical disease. *Fonsecaea pedrosoi* is the fungus that causes chromoblastomycosis, a skin infection that is difficult to treat. The CLR Mincle recognizes this pathogen. If costimulation of both CLR and TLR pathways are not engaged, the inflammatory response is defective. Exogenous administration of imiquimod, purified TLR7 ligand, facilitated pathogen clearance in mouse models and humans. This signaling requires both Syk/CARD9 and MyD88. Imiquimod is a synthetic compound with potent antiviral and antitumor activity that stimulates the innate immune system through TLR7 activation. When applied to the lesions of four patients with chromoblastomycosis, there was a rapid resolution of the infection.[25] This provides proof of concept that multiple signaling pathways are necessary for optimal activation of the innate immune system against fungal pathogens.

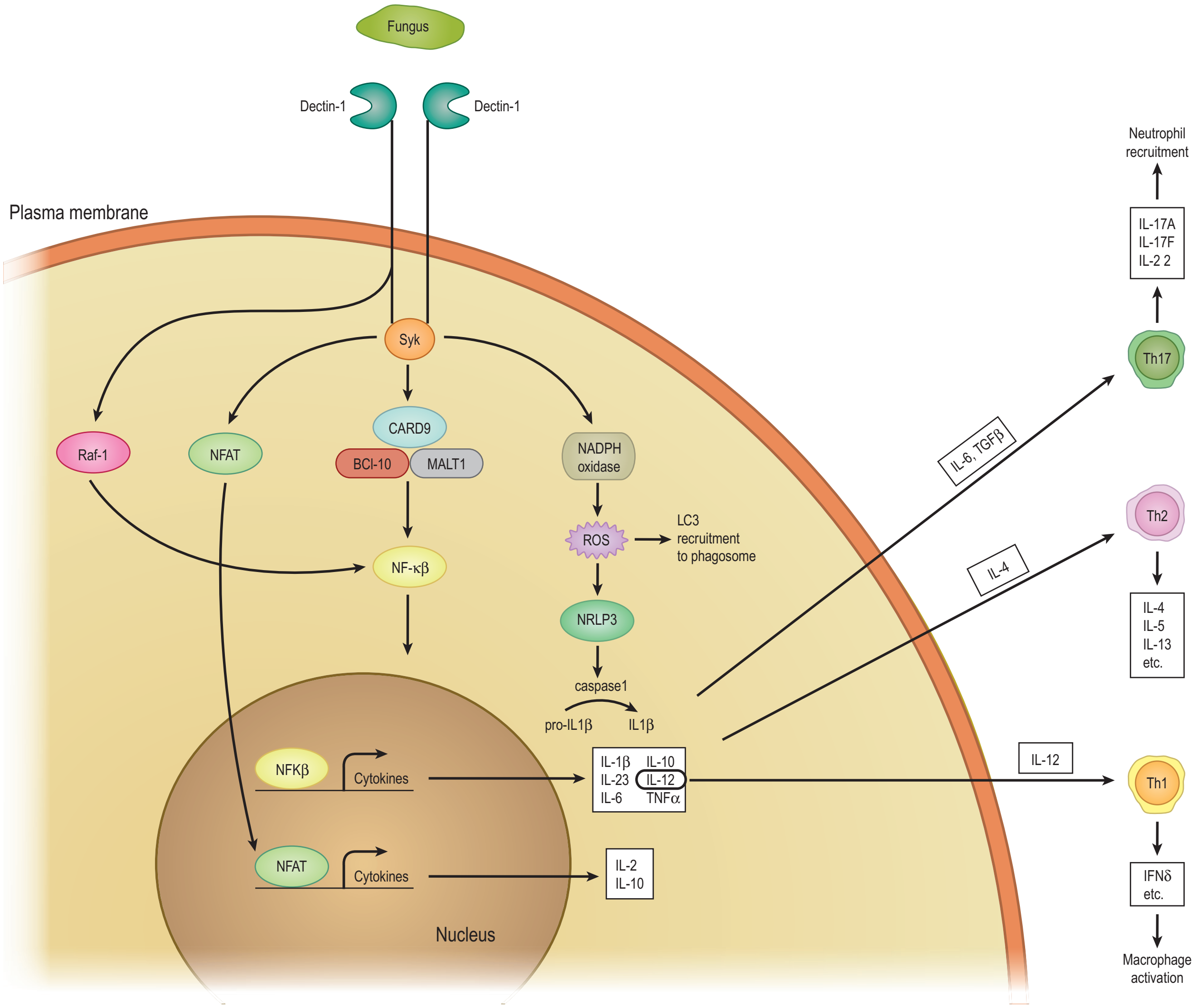

Fungus
Dectin-1
Dectin-1
Plasma membrane
Syk
Raf-1
NFAT
CARD9
BCl-10
MALT1
NADPH oxidase
ROS
LC3 recruitment to phagosome
NF-κβ
NRLP3
caspase1
pro-IL1β
IL1β
NFKβ
Cytokines
IL-1β IL-10
IL-23 IL-12
IL-6 TNFα
NFAT
Cytokines
IL-2
IL-10
Nucleus
IL-6, TGFβ
IL-4
IL-12
Neutrophil recruitment
IL-17A
IL-17F
IL-2 2
Th17
Th2
IL-4
IL-5
IL-13
etc.
Th1
IFNδ
etc.
Macrophage activation

FIG 29.4 Dectin-1 Signaling in Response to Fungal Pathogens. Upon recognition of fungal pathogens, dectin-1 signaling triggers secretion of cytokines, production of reactive oxygen species (ROS), and activation of the adaptive immune system to facilitate elimination of the pathogen.

CARD9

Mutations in the CARD9 adaptor molecule, downstream of dectin-1, are associated with increased susceptibility to invasive *Candida* infections. Patients with CARD9 deficiencies have decreased Th17 cells, reduced chemokine/cytokine production, and impaired neutrophil activation.[26] Numerous patients have been identified to have mutations in CARD9 that lead to fungal infections. The vast majority of inborn errors in immunity that predispose to fungal disease typically also lead to development of nonfungal infections. Patients with gain-of-function (GOF) signal transducer and activator of transcription 1 (STAT1) or autosomal dominant STAT3 also suffer from mycobacterial and other bacterial infections. In contrast, mutations of CARD9 appear to affect susceptibility to fungal infections exclusively. This observation makes CARD9 a key regulator in antifungal immunity. Other genes associated with increased risk of fungal infections typically predispose patients to increased mucocutaneous or invasive disease. In contrast, CARD9 deficiency predisposes to both mucosal and systemic fungal diseases.

THE INFLAMMASOME

IL-1β activation occurs when its precursor pro–IL-1β is cleaved by caspase-1. IL-1β is critical for neutrophil recruitment and for induction of the Th17 response. Activation of the inflammasome regulates IL-1β, IL-18, IL-17, and IFN-γ. NLR family domain containing protein 3 (NLRP3) is required for activation of caspase 1. Caspase 8 may also play a role in regulating activation of IL-1β.[26] Syk-dependent production of ROS activates the NLRP3 inflammasome, which coordinates with NF-κB to promote production of IL-1β.[27] Mice lacking the inflammasome or IL-1β are highly susceptible to disseminated candidiasis.[28] In addition to ROS, lysosomal rupture, release of cathepsins, and efflux of potassium can lead to activation of the inflammasome. Many components of the fungal cell wall can trigger inflammasome activation *in vitro*. However, in these experiments, cells were pretreated with adenosine triphosphate (ATP) or lipopolysaccharide (LPS) to provide an extra signal. When whole live fungal organisms are used, pretreatment with ATP or LPS is not required.

Mice lacking NLRP3, apoptosis-associated speck-like protein containing a C-terminal caspase recruitment domain (ASC), or caspase 1 demonstrate increased susceptibility to *C. albicans* infection. The mechanism appears to be a reduction in neutrophil influx. However, in these mouse models, early neutrophil influx does not appear to be dependent on an intact inflammasome; rather, it is required for sustained neutrophil influx.

MEMORY OF INNATE IMMUNE CELLS

Historically, the innate immune response was considered a rapid and nonspecific attack against invading pathogens. The widely accepted dogma was that innate immune cells confer no immunological memory, unlike adaptive immune cells that respond more slowly but execute a specific attack based on memory. An increasing body of evidence suggests that the innate immune system can, in fact, form memory, challenging the historical dogma. For example, innate immune memory has been demonstrated in organisms that lack the adaptive immune system, such as plants and invertebrates.[29] Several studies have demonstrated increased responsiveness of innate immune cells to secondary encounters with pathogens.[30] Immunological memory of innate immune cells has been described in the context of *C. albicans* infection; in mice, preexposure to a nonlethal infection or purified β-1,3 glucan confers protection from a secondary lethal infection in the absence of functional B cells and T cells. Epigenetic modifications serve as the mechanism by which innate immune cells generate immunological memory.[31] There are numerous lingering questions about innate immune memory, including the duration of this response, its ability to provide protection from other microbes, and whether this type of response can be elicited by vaccination.

LINKING THE INNATE IMMUNE RESPONSE TO ADAPTIVE IMMUNE RESPONSE

A secondary, but critical, role of the innate immune system is to amplify the immune response by activating and recruiting cells of the adaptive immune system, including T and B lymphocytes. The innate immune signaling events that trigger activation of adaptive immunity are complex and may be pathogen dependent. This amplification of the immune response is partially achieved by secretion of cytokines and ROS that recruit immune cells to the site of infection. Additionally, antigen-presenting cells (APCs) of the innate immune system use class II MHC molecules on their surface to present pathogen-derived antigens to I CD4 T cells, resulting in differentiation into effector or regulatory T cells (Tregs).

DCs provide the critical link between the innate and adaptive arms of the immune system. DCs are capable of recognizing fungal pathogens similar to macrophages. Endowed with both TLRs and CLRs, DCs are activated by components of the fungal cell wall. Subsets of DCs provide specialized function. pDCs produce IFN-α in response to ligands that engage endosomal TLRs. Although pDCs play a key role in viral infections, their role in fungal infections is not fully understood. pDCs are able to recognize *A. fumigatus* through engagement of TLR9. When *A. fumigatus* and pDCs are coincubated *in vitro*, growth of the fungus is inhibited, indicating that these cells are capable of controlling growth of this mold. A subset of pDCs secretes IL-6 and IL-23, the latter of which can serve to prime antigen-specific Th17 lymphocytes.

$CD8^+$ DCs possess the unique capacity to cross-present antigens. Cross-presentation of *Histoplasma capsulatum* antigens has been demonstrated. Extracellular *H. capsulatum* are taken up by DCs and are then capable of triggering CD8 T cells through antigen-specific peptide-loaded class I MHC molecules.

Monocyte-derived DCs have a nonredundant role in antifungal immunity, through the induction of Th1 cells. $CCR2^{-/-}$ mice show skewed Th2 responses and poorly controlled *H. capsulatum* infection compared with wild-type control mice. Priming of Th1 lymphocytes require $CCR2^+$ monocyte-derived inflammatory DCs in mouse models of *A. fumigatus* or *C. neoformans* infection. These data indicate a critical role for DCs to connect the innate immune response to drive T-cell responses and coordinate the adaptive immune response.

THE ADAPTIVE IMMUNE RESPONSE TO FUNGAL PATHOGENS

The primary function of the adaptive immune system is to launch a highly specific attack to destroy invading pathogens. The adaptive immune system also provides long-term protection by creating “immunological memory” following an initial response to a pathogen. The two major branches of the adaptive immune system are cell-mediated immunity (CMI) and humoral immunity. B cells and T cells carry out this adaptive immune response. CMI by T cells is critical for the host response to fungal pathogens. There are two major classes of T cells that orchestrate CMI in response to invading pathogens: CD4 T cells and CD8 T cells. Antigen presentation by DCs activates T cells. T-cell activation is also driven by cytokines secreted from macrophages and DCs. In turn, activated T cells secrete cytokines that elaborate the immune response.

KEY CONCEPTS

Adaptive Immunity to Fungal Pathogens

- T-cell immunity is critical for fungal host defenses. Patient with advanced HIV infection and AIDS are at heightened risk for *C. neoformans* infection.
- T-helper 1 (Th1) and Th17 T cell subsets are the most important T cells for control of fungal infection
- Th2 cells mediate fungal asthma and hyperreactivity.
- Dectin-1 signaling is required for Th17 immunity.
- Humoral immunity may play a role in antifungal defense but appears to play a supportive role.

CD4 T Cells

CD4 T cells are major players in host fungal immunity. The critical role for CD4 T cells became overwhelmingly apparent during the AIDS epidemic. Patients with HIV/AIDS experience loss of CD4 T cells and are extremely susceptible to opportunistic fungal pathogens, especially *C. neoformans* infection. Additionally, subsets of CD4 T cells, including Th1, Th2, and Th17 cells, have specialized roles in fungal defense (Chapter 16).

Th1 Cells

Th1 cells are a subset of CD4 T cells that play a critical role in controlling fungal infections. After exposure to fungal pathogens, APCs secrete IL-12, which is critical to the maintenance and expansion of Th1 lymphocytes. Mutations in the IL-12 signaling pathway predispose mammals to *Candida, Cryptococcus,* and *Coccidiomycosis* infections. Th1 lymphocytes produce copious amounts of IFN-γ, TNF-α, and granulocyte macrophage–colony stimulating factor (GM-CSF), all of which play a critical role in antifungal defense. Although IFN-γ has multiple effects on different cell types, this cytokine induces several genes in macrophages that potently modify the distribution of proteins on the phagosome. The net result of these effects is enhanced intracellular killing. IFN-γ–treated macrophages have enhanced phagocytosis and are more efficient in antigen presentation. TNF-α has similar effects similar to those of IFN-γ and enhances the function of macrophages. Patients with mutations in the TNF-α receptor or those who are treated with TNF-α are at heightened risk for fungal infections, including *A. fumigatus.* GM-CSF appears to increase ROS production in neutrophils and macrophages. Two lines of evidence in humans suggest an important role for Th1 cells in antifungal defense. Patients with fever and neutropenia with documented IFIs, including candidemia, have improved resolution of infection independent of the effect of GM-CSF on neutrophil mobilization. Patients with protein alveolar proteinosis, which is a disease related to anti–GM-CSF autoantibodies, have increased risk of IFIs.

CLINICAL RELEVANCE

Opportunistic Fungal Pathogens Cause Invasive Fungal Infections in Immunocompromised Patients

- Advances in clinical medicine provide treatment modalities that bring new promise to once-deadly diseases. During the course of these treatments, many patients develop significant deficiencies in their immune system that permit the development of opportunistic infections.
- Invasive fungal infections (IFIs) caused by *Aspergillus fumigatus, Candida albicans,* or *Cryptococcus neoformans* are considered the most serious infections that affect immunocompromised patients. Despite effective fungicidal therapies, such as voriconazole and amphotericin, mortality in these patients exceeds 25%.
- These data indicate that a robust immune response is required for clinical efficacy. Patients with defects in innate immunity are at highest risk despite the fact that they typically have functional circulating lymphocytes.
- There is a dire clinical need to develop novel therapeutics for IFIs to address the significant public health burden caused by opportunistic fungal pathogens.

Th2 Cells

In contrast to Th1 cells, Th2 lymphocytes appear to be defense against most fungal infections. Th2 lymphocytes secrete IL-4, IL-5, and IL-13. The effects of these cytokines on macrophages diminish the antifungal response. IL-4 drives macrophage differentiation toward the alternatively activated phenotype. In mice exposed to either *Aspergillus* or *Cryptococcus* in the respiratory tract, Th2 skewing drives an allergic phenotype with airway hyperresponsiveness (AHR) and goblet hyperplasia. There is one notable exception to the deleterious effect of Th2 cells on invasive fungal disease—*P. jiroveci.* Th2 responses seem to be protective in this infection. Mouse studies suggest that alternatively activated macrophages are critical for clearance of this fungal pathogen. Recent studies have identified a nonredundant role for Th17 lymphocytes, suggesting that Th2 responses are necessary, but not sufficient, for clearance of *P. jiroveci.*

Th17 Cells

Th17 cells are a subset of CD4 T cells that have emerged as major contributors to host defense against fungal pathogens. These lymphocytes play a critical and nonredundant role in the control of IFIs. Th17 cells are characterized by production of cytokines, IL-17A, IL-17F, and IL-22. Differentiation of this T-cell lineage requires stage-specific stimulation by cytokines and activation of specific transcription factors. Transforming growth factor-β (TGF-β) and IL-6 prime the initial differentiation of CD4 T cells to Th17 cells. IL-23 is required for both maintenance and expansion of these cells.[32] Activation of STAT3 by IL-6 regulates transcription of retinoid orphan receptor-γt (ROR-γt), the master transcription factor controlling Th17 lineage commitment. Several studies have confirmed the importance of Th17 cells and IL-17/IL-23 signaling for antifungal immunity.[33] Mice deficient in IL-17 receptor fail to generate functional Th17 cells and are highly susceptible to both systemic and mucocutaneous candidiasis. Additionally, IL-23–deficient animals develop progressively worse mucocutaneous disease. Likewise, in mouse models of vulvovaginal *C. albicans* infection, impairment of Th17 cell function led to increased fungal burden. Interestingly, fungal pathogens have devised mechanisms to subvert this component of the immune system. Cell surface receptors found on *Candida* and *Aspergillus* bind IL-17 and cause increased hyphal growth and transcriptional changes within these fungal organisms that are thought to combat the immune response.

CD8 T Cells

Although CD8 T cells are critical for immune defense against viral pathogens and tumors, their contribution to fungal immunity is incompletely understood. They may provide functional redundancy to confer protection in the absence of CD4 T cells. These cells will be important in the development of fungal vaccines, but further studies are required to delineate if they have any nonredundant role in antifungal immunity.

B Cells

Early data dismissed the role of humoral immunity during fungal infection. However, clinical observations have suggested that antibodies contribute to fungal host defense. Patients with B-cell defects, including X-linked hyper-IgM syndrome and hypogammaglobulinemia, demonstrate increased susceptibility to cryptococcosis. Additionally, pediatric patients with selective IgG2 deficiency are at heightened risk for IFIs and other bacterial infections that rely on antipolysaccharide antibody formation.

Antibodies observed in normal individuals are typically against fungal cell wall components and have the capacity to inflict negative changes within the fungal cell. These include inducing transcriptional changes that are deleterious to the microbe. Antibodies can work in conjunction with complement and phagocytic cells to facilitate damage of the cell wall and trigger uptake by cells, respectively. Antibodies also have the capacity to trigger transcriptional changes that increase pathogen virulence, thereby exacerbating disease. Pleiotropic effects have made it difficult to resolve the net value of humoral immunity to IFIs.

Natural Killer Cells

Natural killer (NK) cells exert an antifungal response, but the rules that govern this response have only been elucidated recently. NK cell–mediated killing requires direct contact with the fungus, which leads to release of perforin. The NK cell receptor p30 is required for NK cell-fungal conjugate formation, pI3K signaling, and perforin release. p30 was previously identified as an activating receptor against tumor cells. In patients with

AIDS, who are highly susceptible to *C. neoformans* infection, p30 expression is decreased on NK cells leading to defective perforin release and reduced microbicidal activity. IL-12 restored p30 expression and fungal killing in a mouse model of *C. neoformans* infection by NK cells. These data indicate a direct role for NK cells against *Cryptococcus* through direct recognition by a cell surface receptor. Evidence for the role of NK cells against invasive fungal pathogens has been expanded to show that they play a nonredundant role in clearing *C. glabrata* infections. The human NK receptor p46 and its mouse orthologue natural cytotoxicity triggering receptor 1 (NCR1) bind *C. glabrata. In vitro,* NK cells lacking NCR1 show reduced capacity to kill *C. glabrata*. Indeed, NCR1$^{-/-}$ mice have increased susceptibility to *C. glabrata* infection. The ligands for p46/NCR1 are discrete members of a family of proteins called EPA (epithelial adhesion), which are glycan-binding lectins and permit attachment of fungal cells to host cells.

GENETIC SUSCEPTIBILITIES TO INVASIVE FUNGAL INFECTIONS

Although certain clinical diagnoses and medical interventions are associated with susceptibility to fungal infections, not all patients in these high-risk groups become infected. Moreover, individuals lacking traditional risk factors (*e.g.,* immunosuppression, HIV, etc.) can develop chronic or invasive fungal infections, suggesting that additional factors confer susceptibility. Primary immunodeficiencies (PIDs) include a group of hereditary immune disorders that render patients more susceptible to infection. Recent studies of PIDs have provided the opportunity to define some of the genetic and molecular mechanisms that contribute to infection. A list of genes associated with susceptibility to fungal infection are summarized in Table 29.1.

Chronic granulomatous disease (CGD) is a well-known example of a genetic disorder associated with increased susceptibility to invasive bacterial and fungal infections, especially invasive aspergillosis. CGD is caused by mutations in genes encoding subunits of the NADPH complex (*e.g., CYBA, CYBB NCF1, NCF2, NCF4*), which results in defective ROS production, a critical step in the microbial killing process.[34,35] Recent work has shown that deficiencies in ROS production resulting from a mutation in a subunit of the NADPH complex also result in defective autophagy. These defects prevent autophagy-dependent inhibition of IL-1β and increased inflammasome activation. Blocking of IL-1 protects mice with CGD from invasive aspergillosis and colitis and restores autophagic function in monocytes derived from patients with CGD.[36]

Chronic mucocutaneous candidiasis (CMC) is an immune disorder characterized by persistent or recurrent candidal infections of the skin, nails, and mucous membranes. Primarily, CMC appears to be associated with defects in IL-17 signaling. A direct role for IL-17 signaling in the protection against CMC was demonstrated in two families with deficiencies in the IL-17 receptor A (IL-17RA) and the cytokine IL-17F.[37] As mentioned previously, mutations in dectin-1 cause increased susceptibility to CMC as a result of impaired IL-1β production and a reduced frequency of Th17 cells.[38] Similarly, patients with CARD9 deficiencies have increased susceptibility to invasive fungal infections resulting from decreased Th17 cells, reduced chemokine/cytokine production, and impaired neutrophil killing.[39]

CMC is also common in patients with autoimmune polyendocrinopathy syndrome type 1 (APS-1), an autosomal recessive disorder caused by defects in the autoimmune regulator (*AIRE*) gene. In these patients, CMC results from the generation of autoantibodies against IL-17 and IL-22.[40] Hyper-IgE syndrome (HIES) is an autosomal dominant or recessive immune disorder

TABLE 29.1 Genes Associated With Immunity to Fungi in Humans

Gene Product	Disease	Fungal Pathogen	Immunological Phenotype
ACT1	CMC[a]	*Candida*	Impaired interleukin-17 (IL-17) signaling.
AIRE	CMC; Autoimmune Polyendocrinopathy Syndrome-1	*Candida*	Autoantibodies against IL-17 and IL-22
CARD9	CMC; disseminated candidiasis; *Candida* meningoencephalitis; deep dermatophytosis; subcutaneous phaeohyphomycosis; invasive *Exophiala* infection	*Candida, Trichophyton, Phialophora, Exophiala*	Defective dectin-1 signaling and reduced frequency of T-helper 17 (Th17) cells Impaired neutrophil killing
CYBA *CYBB* *NCF1* *NCF2* *NCF4*	CGD[b]	*Candida, Aspergillus*	Nicotinamide adenine dinucleotide phosphate (NADPH) complex deficiency causing impaired reactive oxygen species (ROS) production
Dectin-1	CMC	*Candida*	Impaired IL-1β and IL-17 production and reduced frequency of Th17 cells
DOCK8	CMC; HIES[c]	*Candida*	Defective activation of Th17 cells
IL-17F	CMC	*Candida*	Defective IL-17 signaling
IL-17RA	CMC	*Candida*	Defective IL-17 signaling
RORC	Candidiasis	*Candida*	Absent IL-17A/-17F–producing T cells
STAT1	CMC, cutaneous fusariosis, disseminated coccidioidomycosis, histoplasmosis, mucormycosis, and *Penicillium marneffei* infection	*Candida, Fusarium, Coccidioides, Histoplasma, Penicillium marneffei, Apophysomyces*	Defective production of IL-17, IL-22, and interferon-γ (IFN-γ) and reduced frequency of Th17 cells
STAT3	CMC; HIES	*Candida*	Impaired Th17 differentiation and reduced frequency of Th17 cells
TYK2	CMC; HIES	*Candida*	Defective IL-23 signaling and reduced frequency of Th17 cells

[a] Chronic mucocutaneous candidiasis.
[b] Chronic granulomatous disease.
[c] Hyper–immunoglobulin E syndrome.

characterized by high serum IgE, eczema, and CMC. HIES is caused by mutations in multiple genes, including *STAT3, Tyk2,* and *DOCK8.*[41] Patients with *STAT3* mutations have impaired Th17 differentiation, which results in reduced number of Th17 cells and IL-17 production.[42] Autosomal dominant CMC is caused by GOF mutations in *STAT1,* which lead to decreased IL-17, IL-22, and IFN-γ production and defective Th1 and Th17 responses.[43-45] In contrast, STAT1 deficiencies confer susceptibility to viral and mycobacterial infections. *RORC* loss-of-function mutations have also been reported in humans with candidiasis; RORC isoforms ROR-γ and ROR-γT are critical for T-cell development and function.[46] Thus the absence of IL-17A/-17F–producing T cells likely accounts for the candidiasis in these patients.

The identification of genes associated with PIDs revealed several important pathways in host immunity to fungal infections. Collectively, genes associated with PIDs highlight the protective and critical role of IL-17 signaling against invasive fungal infections.

ON THE HORIZON

- The study of fungal immunology has led to a fundamental understanding of the role of C-type lectin receptors (CLRs) in immunology. Innate immunology plays a critical and nonredundant role in the defense against invasive fungal infections.
- When patients lack neutrophils or develop defects in neutrophil function, only a subset of patients develop invasive fungal infection (*e.g., A. fumigatus* infection).
- These observations suggest that genetic defects in the immune system likely contribute to increased risk of infection.
- Identification of genetic mutations that are associated with increased risk of fungal disease may permit the development of genomic risk stratification for high-risk patients.
- To this end, rational use of antifungal prophylaxis may decrease the burden of fungal disease.

SUMMARY

To date, more than 5 million distinct species of fungi have been identified, but only a few are relevant to human health. *Aspergillus* and *Cryptococcus* are ubiquitous in the environment, and their spores are routinely inhaled by humans; *Candida* spp. colonize human skin, mucosae, and GI tract and, under normal conditions, do not cause disease. Normally, the immune system is able to clear and/or limit these pathogens, but in immunocompromised individuals, these fungi can cause serious and often life-threatening infections. Using both TLRs and CLRs, the innate immune system plays a critical role in clearing the invading pathogens. Indeed, patients with defects in either innate immune cells (*e.g.,* neutropenia) or signaling through TLRs and CLRs have an increased risk for diseases caused by invasive fungal organisms. Adaptive immunity is also critical, as evidenced by patients with advanced HIV disease who are at heightened risk for *C. neoformans* disease. In general, Th1 and Th17 responses are protective, whereas Th2 responses skew the immune response to allergic and hyperactive responses.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Ruping MJ, Vehreschild JJ, Cornely OA. Patients at high risk of invasive fungal infections: when and how to treat. Drugs 2008;68(14):1941–62.
2. Azie N, Neofytos D, Pfaller M, et al. The PATH (Prospective Antifungal Therapy) Alliance(R) registry and invasive fungal infections: update 2012. Diagn Microbiol Infect Dis 2012;73(4):293–300.
3. Sobel JD. Vulvovaginal candidosis. Lancet 2007;369(9577):1961–71.
4. Brown GD, Denning DW, Gow NA, et al. Hidden killers: human fungal infections. Sci Transl Med 2012;4(165):165rv13.
5. Herbrecht R, Denning DW, Patterson TF, et al. Voriconazole versus amphotericin B for primary therapy of invasive aspergillosis. N Engl J Med 2002;347(6):408–15.
6. Garcia-Vidal C, Viasus D, Carratala J. Pathogenesis of invasive fungal infections. Curr Opin Infect Dis 2013;26(3):270–6.
7. Vazquez JA, Miceli MH, Alangaden G. Invasive fungal infections in transplant recipients. Ther Adv Infect Dis 2013;1(3):85–105.
8. Cornely OA, Gachot B, Akan H, et al. Epidemiology and outcome of fungemia in a cancer Cohort of the Infectious Diseases Group (IDG) of the European Organization for Research and Treatment of Cancer (EORTC 65031). Clin Infect Dis 2015;61(3):324–31.
9. Kullberg BJ, Arendrup MC. Invasive Candidiasis. N Engl J Med 2015;373(15):1445–56.
10. van Woerden HC, Gregory C, Brown R, et al. Differences in fungi present in induced sputum samples from asthma patients and non-atopic controls: a community based case control study. BMC Infect Dis 2013;13:69.
11. Weindl G, Wagener J, Schaller M. Epithelial cells and innate antifungal defense. J Dent Res 2010;89(7):666–75.
12. Naglik JR, Moyes DL, Wachtler B, et al. *Candida albicans* interactions with epithelial cells and mucosal immunity. Microbes Infect 2011;13(12-13):963–76.
13. Liu Y, Zhao G, Lin J, et al. The role of Syk signaling in antifungal innate immunity of human corneal epithelial cells. BMC Ophthalmol 2015;15:55.
14. Moyes DL, Runglall M, Murciano C, et al. A biphasic innate immune MAPK response discriminates between the yeast and hyphal forms of *Candida albicans* in epithelial cells. Cell Host Microbe 2010;8(3):225–35.
15. Kim YG, Udayanga KG, Totsuka N, et al. Gut dysbiosis promotes M2 macrophage polarization and allergic airway inflammation via fungi-induced PGE(2). Cell Host Microbe 2014;15(1):95–102.
16. Nosanchuk JD, Stark RE, Casadevall A. Fungal Melanin: What do We Know About Structure? Front Microbiol 2015;6:1463.
17. Urban CF, Reichard U, Brinkmann V, et al. Neutrophil extracellular traps capture and kill *Candida albicans* yeast and hyphal forms. Cell Microbiol 2006;8(4):668–76.
18. Plato A, Hardison SE, Brown GD. Pattern recognition receptors in antifungal immunity. Semin Immunopathol 2015;37(2):97–106.
19. Bourgeois C, Kuchler K. Fungal pathogens-a sweet and sour treat for toll-like receptors. Front Cell Infect Microbiol 2012;2:142.
20. Khan NS, Kasperkovitz PV, Timmons AK, et al. Dectin-1 Controls TLR9 Trafficking to Phagosomes Containing beta-1,3 Glucan. J Immunol 2016;196(5):2249–61.
21. Kasperkovitz PV, Khan NS, Tam JM, et al. Toll-like receptor 9 modulates macrophage antifungal effector function during innate recognition of *Candida albicans* and *Saccharomyces cerevisiae*. Infect Immun 2011;79(12):4858–67.
22. Mansour MK, Tam JM, Khan NS, et al. Dectin-1 activation controls maturation of beta-1,3-glucan-containing phagosomes. J Biol Chem 2013;288(22):16043–54.
23. Ferwerda G, Netea MG, Joosten LA, et al. The role of Toll-like receptors and C-type lectins for vaccination against *Candida albicans*. Vaccine 2010;28(3):614–22.
24. Gringhuis SI, den Dunnen J, Litjens M, et al. Dectin-1 directs T helper cell differentiation by controlling noncanonical NF-κB activation through Raf-1 and Syk. Nat Immunol 2009;10(2):203–13.

25. de Sousa Mda G, Belda W Jr, Spina R, et al. Topical application of imiquimod as a treatment for chromoblastomycosis. Clin Infect Dis 2014;58(12):1734–7.
26. Underhill DM, Pearlman E. Immune interactions with pathogenic and commensal fungi: a two-way street. Immunity 2015;43(5):845–58.
27. van de Veerdonk FL, Joosten LA, Netea MG. The interplay between inflammasome activation and antifungal host defense. Immunol Rev 2015;265(1):172–80.
28. Caffrey AK, Obar JJ. Alarmin(g) the innate immune system to invasive fungal infections. Curr Opin Microbiol 2016;32:135–43.
29. Netea MG, Latz E, Mills KH, et al. Innate immune memory: a paradigm shift in understanding host defense. Nat Immunol 2015;16(7):675–9.
30. Netea MG, Joosten LA, Latz E, et al. Trained immunity: A program of innate immune memory in health and disease. Science 2016;352(6284):aaf1098.
31. Quintin J, Saeed S, Martens JH, et al. *Candida albicans* infection affords protection against reinfection via functional reprogramming of monocytes. Cell Host Microbe 2012;12(2):223–32.
32. Zuniga LA, Jain R, Haines C, et al. Th17 cell development: from the cradle to the grave. Immunol Rev 2013;252(1):78–88.
33. Hamad M. Innate and adaptive antifungal immune responses: partners on an equal footing. Mycoses 2012;55(3):205–17.
34. Clark RA, Malech HL, Gallin JI, et al. Genetic variants of chronic granulomatous disease: prevalence of deficiencies of two cytosolic components of the NADPH oxidase system. N Engl J Med 1989; 321(10):647–52.
35. Wang X, van de Veerdonk FL. When the Fight against Fungi Goes Wrong. PLoS Pathog 2016;12(2):e1005400.
36. de Luca A, Smeekens SP, Casagrande A, et al. IL-1 receptor blockade restores autophagy and reduces inflammation in chronic granulomatous disease in mice and in humans. Proc Natl Acad Sci USA 2014;111(9): 3526–31.
37. Puel A, Cypowyj S, Bustamante J, et al. Chronic mucocutaneous candidiasis in humans with inborn errors of interleukin-17 immunity. Science 2011;332(6025):65–8.
38. Ferwerda B, Ferwerda G, Plantinga TS, et al. Human dectin-1 deficiency and mucocutaneous fungal infections. N Engl J Med 2009;361(18): 1760–7.
39. Drummond RA, Collar AL, Swamydas M, et al. CARD9-dependent neutrophil recruitment protects against fungal invasion of the central nervous system. PLoS Pathog 2015;11(12):e1005293.
40. Puel A, Doffinger R, Natividad A, et al. Autoantibodies against IL-17A, IL-17F, and IL-22 in patients with chronic mucocutaneous candidiasis and autoimmune polyendocrine syndrome type I. J Exp Med 2010; 207(2):291–7.
41. Gow NA, Netea MG. Medical mycology and fungal immunology: new research perspectives addressing a major world health challenge. Philos Trans R Soc Lond B Biol Sci 2016;371(1709):pii: 20150462.
42. Ma CS, Chew GY, Simpson N, et al. Deficiency of Th17 cells in hyper IgE syndrome due to mutations in STAT3. J Exp Med 2008;205(7):1551–7.
43. Liu L, Okada S, Kong XF, et al. Gain-of-function human STAT1 mutations impair IL-17 immunity and underlie chronic mucocutaneous candidiasis. J Exp Med 2011;208(8):1635–48.
44. van de Veerdonk FL, Plantinga TS, Hoischen A, et al. STAT1 mutations in autosomal dominant chronic mucocutaneous candidiasis. N Engl J Med 2011;365(1):54–61.
45. Depner M, Fuchs S, Raabe J, et al. The extended clinical phenotype of 26 patients with chronic mucocutaneous candidiasis due to gain-of-function mutations in STAT1. J Clin Immunol 2016;36(1):73–84.
46. Okada S, Markle JG, Deenick EK, et al. Immunodeficiencies. Impairment of immunity to *Candida* and *Mycobacterium* in humans with bi-allelic RORC mutations. Science 2015;349(6248):606–13.

MULTIPLE-CHOICE QUESTIONS

1. Loss of dectin-1 in patients leads to:
   A. Chronic mucocutaneous candidiasis
   B. Candidemia
   C. No clinical disease unless patient has neutropenia
   D. Increased risk of *Cryptococcus neoformans* disease

2. Which of the following carbohydrates is NOT found on the fungal cell wall?
   A. β-1,3 glucan
   B. Galactomannan
   C. Glucuronoxylomannan
   D. Peptidoglycan

3. Which T-cell response typically makes fungal infection worse?
   A. T-helper 1 (Th1)
   B. Th2
   C. Th17
   D. CD8

Host Defenses to Protozoa

Peter C. Melby, Robin Stephens, Sara M. Dann

Protozoal infections are an important cause of morbidity and mortality worldwide (Table 30.1). Protozoan pathogens exact their major toll in the tropics, but infection by these parasites remains a significant problem in developed countries because of travel to and emigration from developing countries, the susceptibility of patients with acquired immunodeficiency syndrome (AIDS) to opportunistic protozoans, and episodic transmission within communities.

KEY CONCEPTS

Host Defense Against Protozoa

- Interaction of the parasite with host cells induces an array of cytokines that stimulate the innate and adaptive immune responses to eliminate the pathogen, and/or cytokines that inhibit or downregulate the antiparasitical responses to enable the initiation of tissue parasitism.
- The outcome of infection is determined by the balance between the infection-promoting and the host-protective cytokines and effector cells. Often there is a mixed response, resulting in a persistent infection.
- A persistently infected host may develop clinical disease if there is a waning of the immune mechanisms (*e.g.,* in acquired immunodeficiency syndrome [AIDS]) that are critical to the control of infection.

Protozoan pathogens make up a group of highly diverse organisms that utilize a wide array of mechanisms for pathogenesis and immune evasion. There are numerous host targets for the intracellular protozoan parasites, including erythrocytes (*Plasmodium* and *Babesia*), macrophages (*Leishmania* and *Toxoplasma gondii*), or multiple cell types (*Trypanosoma cruzi*). The luminal parasitical protozoan may be extracellular, such as amebae and the flagellates (*Giardia* and *Trichomonas*), or primarily intracellular, such as the coccidian parasite *Cryptosporidium.*

The innate and adaptive immune systems respond in diverse ways to the blood and tissue and intestinal protozoan pathogens. Neutrophils, macrophages, and natural killer (NK) cells are the effector cells that mediate the innate response against the extracellular protozoan parasites. The NK cell–activated macrophage system is central to the innate response to intracellular parasites (Fig. 30.1) (Chapters 3, 17). The innate cytokine response activates phagocytes and is critical to the induction of the adaptive immune response via antigen presentation by dendritic cells (DCs). For the intracellular pathogens (*e.g., Leishmania* spp., *T. cruzi*, *T. gondii*), the early production of interleukin-12 (IL-12) and interferon-γ (IFN-γ) drives the differentiation of T cells to a protective T-helper 1 (Th1) phenotype. In most cases CD4 T cells play a primary role in adaptive cellular immunity, but CD8 T cells can be critically important through cytokine production (*e.g., Plasmodium* spp., *T. cruzi*, *T. gondii*) or direct cytotoxic activity (*e.g., Cryptosporidium*). For the parasites that have an extracellular stage (*e.g., Plasmodium* spp., *Trypanosoma* spp., *Giardia*, and *Trichomonas*), specific antibodies mediate acquired immunity.

Intensive effort has been dedicated to the development of effective vaccines for protozoal diseases, but as of 2016, only the malaria circumsporozoite vaccine (RTS,S) has reached the stage of clinical use. The reader is referred to a number of excellent reviews of the potential vaccine candidates.[1-4] A discussion of the immune responses to some of the individual protozoal pathogens follows.

Plasmodium spp.

Pathogenesis

Soon after *Plasmodium* spp. sporozoites are injected into the bloodstream by the *Anopheles* mosquito, they invade hepatocytes and undergo schizogony (asexual reproduction). A dormant form of *P. vivax* and *P. ovale* (hypnozoites) can reside within hepatocytes for months and then cause clinical bloodstream infection. Following schizogony, merozoites are released from hepatocytes into the bloodstream in a membrane-bound structure, known as *merosomes*. The merosomes rupture in blood, and free merozoites invade red blood cells (RBCs) to produce ring-stage parasites. These parasites mature into trophozoites, which again undergo schizogony, leading to rupture of the erythrocyte and the release of new invasive merozoites. Merozoites can also develop into sexual-stage gametocytes, which can be ingested by a feeding mosquito to continue the transmission cycle.

The clinicopathological features of malaria are caused by intraerythrocytic infection and the associated immune response. The cyclical rupture of erythrocytes is associated with fever. The induction of a proinflammatory cytokine cascade plays a central role in the pathogenesis of *P. falciparum* malaria and its complications. Parasite antigens, particularly those having glycophosphatidyl inositol (GPI) membrane anchors, released during the rupture and reinvasion of RBCs, activate the innate immune response. The production of proinflammatory cytokines (IL-1, TNF, lymphotoxin, IL-12, and IFN-γ) leads to fever, expression of endothelial adhesion molecules, and cytoadherence. It is mediated, in part, by TLR2 and is MyD88-dependent.[5] NK cells and memory T cells produce early IFN-γ, which contributes to the production of a pathologically high level of TNF. IL-10- and transforming growth factor-β (TGF-β)–mediated downregulation of the Th1 immune response and leukotriene (LT)/tumor necrosis

TABLE 30.1 Worldwide Significance of the Major Protozoal Infections

Parasite	Estimated Worldwide Cases (Annual Mortality)	Clinical Manifestations
Plasmodium spp.	400–490 million (*P. falciparum*: >2 million deaths/year, primarily children)	Fever with potential complications of severe hemolysis, renal failure, pulmonary edema, cerebral involvement
Leishmania spp.	10–50 million people infected, 1.2 million new cases per year	Asymptomatic infection; skin ulcers or nodules; destructive oropharyngeal lesions; visceral disease with fever, hepatosplenomegaly, cachexia, pancytopenia
Trypanosoma cruzi	24 million (60 000 deaths)	Asymptomatic infection; dysrhythmias or chronic heart failure; hypertrophy and dilation of the esophagus, colon
Toxoplasma gondii	Several hundred million people infected worldwide. 5–9% of healthy US adults are seropositive	Self-limited fever, hepatosplenomegaly; lymphadenopathy and encephalitis (reactivation in patients with acquired immunodeficiency syndrome [AIDS]); congenital infection, with fetal death, chorioretinitis, meningoencephalitis
Entamoeba histolytica	50 million (100 000 deaths)	Asymptomatic infection, diarrhea, dysentery, or liver abscess
Giardia lamblia	200 million (most common in young children and immunocompromised persons)	Asymptomatic infection, chronic diarrhea
Cryptosporidium parvum and *C. hominis*	Prevalence 3–10% in patients with diarrhea in developing countries	Self-limited diarrhea in immunocompetent persons, severe intestinal and biliary disease in patients with AIDS
Trichomonas vaginalis	170 million/year	Asymptomatic infection, vaginal discharge, urethritis

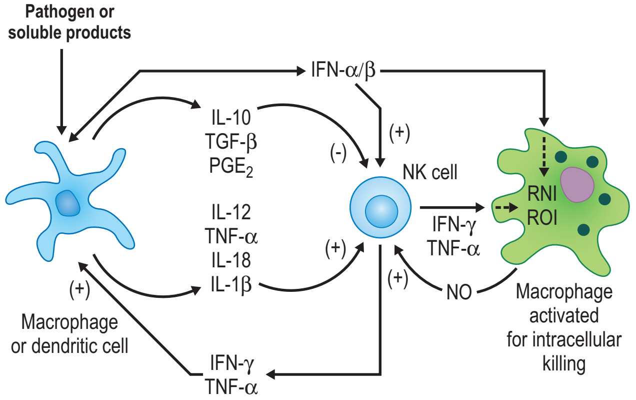

Pathogen or soluble products
IFN-α/β
IL-10
TGF-β
PGE$_2$
(-)
(+)
NK cell
RNI
ROI
IL-12
TNF-α
IL-18
IL-1β
(+)
IFN-γ
TNF-α
(+)
NO
Macrophage activated for intracellular killing
(+)
Macrophage or dendritic cell
IFN-γ
TNF-α

FIG 30.1 **Macrophage, Natural Killer (NK) Cell, and Cytokine Interactions in the Innate Immune Response to Intracellular Protozoa.** Exposure of macrophages or dendritic cells to a pathogen or microbial product can result in the release of cytokines and inflammatory mediators that may stimulate (+) or suppress (–) NK cell activation. Activated NK cells produce cytokines that can then activate macrophages for intracellular killing. It must be recognized that this diagram is oversimplified and that these cytokines, most notably interferon (IFN)-α/β, interleukin (IL)-10, transforming growth factor (TGF)-β, and IL-12, may be produced by other types of cells, such as epithelial cells or enterocytes. NO, nitric oxide; RNI, reactive nitrogen intermediates; ROI, reactive oxygen intermediates.

factor (TNF) production, which have a role in severe and cerebral malaria, act to limit the inflammatory response.

Severe malaria includes severe anemia, respiratory distress, placental malaria, and cerebral malaria, with the latter causing up to 90% of deaths. In cerebral malaria, a combination of inflammation, cytoadhesion of parasites, and leukocytes, and vascular pathology leads to coma. Cytokines and endothelial cell production of nitric oxide (NO) contribute to inflammatory lesions in the brain. Patients may die as a result of severe edema and swelling of the brainstem. Severe anemia is the result of a combination of destruction of infected and uninfected RBCs and dyserythropoiesis. Both destruction of uninfected RBCs and malformation of RBCs in bone marrow are increased by inflammatory cytokines, as is the damage in placental malaria.

KEY CONCEPTS

Immunopathogenesis of Severe Plasmodium falciparum *Malaria*

- Release of malarial antigens stimulates tumor necrosis factor (TNF), interleukin (IL)-1, and lymphotoxin production from innate immune cells.
- TNF/leukotriene (LT) induce vascular leakage, hemorrhage, endothelial cell activation with expression of endothelial adhesion molecules, platelet activation and adhesion, and coagulation.
- Inflammatory cytokines amplify severe anemia caused by loss of infected red blood cells (RBCs) by inducing dyserythropoiesis and phagocytosis of uninfected RBCs, which can be coated with parasite antigens.
- Lethal outcomes for the fetus of infected women during pregnancy are caused by the local immune response induced by placenta-specific parasites.

Innate Immunity

Complement-mediated lysis can occur at the sporozoite and merozoite stages, though parasites have evolved evasive mechanisms. While sporozoites rapidly transit from the skin to the liver, they can activate DCs at the site of inoculation or in the regional lymph node. Early activation of NK cells and IFN-γ are associated with better outcomes of infection. Although high levels of TNF are associated with severe malaria, physiological levels are protective through the activation of macrophages. γδ T-cell receptor (TCR)$^+$ T cells that respond to phosphorylated nonpeptide antigens on live parasites have been demonstrated. DCs initially induce a strong adaptive immune response to the erythrocytic stage, which is tempered during prolonged infection.

Adaptive Immunity

Partial immunity to *Plasmodium* spp. infection is acquired slowly following repeated exposure in endemic areas.[6] In areas of intense perennial *P. falciparum* transmission, the density of parasitemia, morbidity, and the incidence of cerebral malaria and malaria-related deaths are highest in the early childhood years, declining thereafter. Naturally acquired immunity to repeated infections develops in young adults, with the exception of the morbidity associated with placental infections.

Adaptive immunity to the preerythrocytic stage is primarily mediated through major histocompatibility complex (MHC) class I-restricted parasite-specific CD8 T cells and, to a lesser extent, CD4 T cells, via IFN-γ-induced NO-dependent killing of intrahepatocyte parasites.[7] Potentially protective immune mechanisms against the preerythrocytic stage have largely been identified by study of mice vaccinated with irradiated sporozoites and challenged with murine *Plasmodium* spp. Irradiated sporozoite challenge also protects humans and is being scaled up for vaccine trials. Antisporozoite immunity requires the presence of high antibody titers and high numbers of T cells to block sporozoite invasion of the hepatocyte, which occurs in just a few cells within minutes of inoculation.

Both antibody-dependent and cell-mediated immune mechanisms are active against the erythrocytic stage of infection. Both B cells and CD4 T cells are required for complete clearance of parasites.[8] Adoptive transfer of human immune serum is protective for naïve individuals. Antibodies directed against merozoite surface proteins can inhibit invasion. A significant proportion of antibodies to infected erythrocytes are directed toward variant antigens, such as PfEMP-1 on the surface of the RBC. Immunoglobulin G1 (IgG1) and IgG3 isotype antibodies specific for parasite antigens exported to the surface of the RBC play a role in naturally acquired immunity by opsonizing infected cells for phagocytosis in the spleen.

CD4 T cells are implicated in protection by experiments showing that MHC class II–restricted antigen presentation is required for reduction of parasitemia and pathology. Protective cellular immune responses (CD4 and CD8 T-cell proliferation, IFN-γ production, and NO synthesis) in the absence of detectable antibody responses were identified in naïve volunteers, who were protected by repeated exposure to low doses of blood-stage parasites. In addition to promoting phagocytosis, CD4 T cells are required to help B cells, especially by the production of IL-21. One of the most important roles for CD4 T cells is the regulation of the intense inflammatory response through production of antiinflammatory cytokines IL-10 and TGF-β.

Evasion of Host Immunity

The malaria parasite uses several different mechanisms to evade the host immune response.[9,10] Sporozoites and merozoites evade circulating antibody by rapidly entering hepatocytes or RBCs, respectively. Some sporozoite proteins enter the hepatocyte nucleus and influence the expression of a number of host genes, thereby favoring parasite survival. Mature RBCs do not express MHC molecules on their surface and so avoid recognition by T cells. The few parasite proteins that are expressed on the erythrocyte surface exist in multiple allelic forms to avoid quick recognition by the adaptive immune system. Many of the immunodominant antigens in *Plasmodium* spp. are proteins having extensive repeat sequences that vary over time. The large number of bloodborne antigens induces a plasmablast response generating a short-lived burst of low-affinity antibodies. Although there is an increase in atypical B cells, memory B cells and long-lived plasma cells are generated. It is unclear if poor immunity following a single infection is caused by ineffective, atypical B cells or misdirected antibody specificity.

Leishmania spp.

Pathogenesis

The intracellular *Leishmania* amastigote replicates within macrophages in the vertebrate host, and the extracellular promastigote develops within the insect vector. The female phlebotomine sand fly becomes infected by ingesting amastigotes during a blood meal. In the sand fly gut, the amastigotes differentiate into infectious metacyclic promastigotes that infect the vertebrate host during the next blood meal. The surface lipophosphoglycan (LPG) plays a central role in the parasite's entry and survival in host cells. Immunomodulatory factors present in the sand fly saliva may enhance the infectivity of the parasite. Once introduced into skin, the promastigotes are phagocytosed (through complement–complement receptor–mediated coiling phagocytosis) by neutrophils, DCs, and macrophages, where they transform to amastigotes and replicate within the acidic and hostile environment of the phagolysosome. Eventually, the phagocytes rupture and release amastigotes to infect other macrophages or a sand fly.

Innate Immunity

Much of what we know of immunity in leishmaniasis comes from studies of inbred mouse strains, which demonstrate a genetically determined spectrum of innate and adaptive immune responses that shape the outcome of infection. The innate immune response to *Leishmania* is mediated by complement, NK cells, cytokines, and phagocytes.[11] The production of IL-12 early in the course of infection by DCs leads to the early activation of NK cells and the production of IFN-γ. Chemokines (IP-10, MCP-1, and lymphotactin), as well as LPG–TLR interaction, can also promote the early NK-cell activation. Activated NK cells have been shown to be cytolytic for *Leishmania*-infected macrophages, but NK cell–derived IFN-γ plays a more prominent role in host defense by activating macrophages to kill the intracellular parasite through the generation of reactive oxygen intermediates (ROIs) or reactive nitrogen intermediates (RNIs). Parasite-induced, MyD88-dependent signaling through TLR2, TLR3, and TLR4 contributes to macrophage activation and NO production. Activated polymorphonuclear leukocytes (PMNs) can kill parasites through oxidative mechanisms, but the role of neutrophils *in vivo* depends on the timing of their recruitment and their interaction with other immune cells.[12] An important study demonstrated that infiltrating neutrophils promote sand fly–transmitted infection, probably through modulation of macrophage function following engulfment of apoptotic parasitized neutrophils.[13] Type 1 IFNs participate in the early induction of NO and control of parasite replication early in infection.

Adaptive Immunity

Within an endemic area there is acquisition of immunity in the population over time. Retrospective epidemiological studies indicate that most individuals with prior (primary) infection (subclinical or healed) are immune to a subsequent clinical infection. Following primary infection, parasites persist for the life of the host and maintain long-term immunity.

There is extensive evidence from experimental models that cellular immune mechanisms mediate adaptive resistance to *Leishmania* infection, and human studies have generally confirmed this. Antileishmanial antibodies, which are produced at a low level in localized cutaneous leishmaniasis (LCL) and at a very high level in visceral leishmaniasis (VL), play no role in protection. The general mechanisms of cellular immunity in leishmaniasis can be summarized (Fig. 30.2). Following infection in the skin, migratory dermal DCs phagocytose *Leishmania* and presumably transport the intracellular parasite to the regional lymph node, where they induce a T-cell response. Adaptive immunity is primarily mediated by parasite-induced production of IFN-γ by CD4

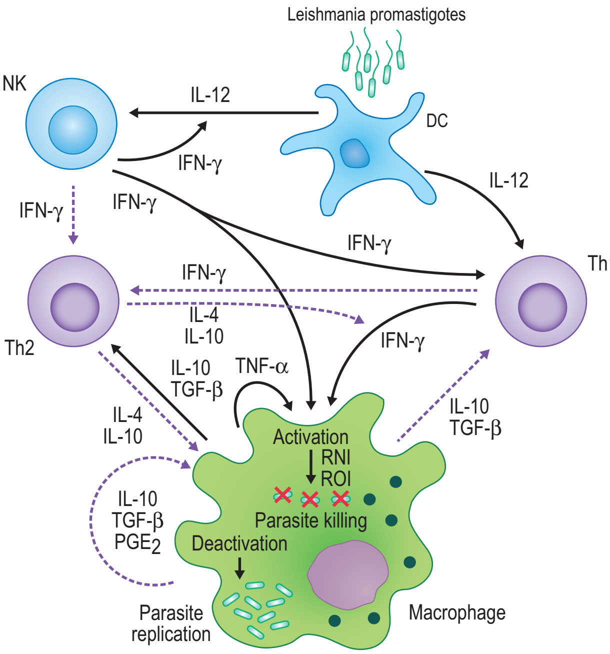


FIG 30.2 Immunity in Leishmaniasis. Exposure of dendritic cells to parasites or parasite antigens leads to the release of interleukin (IL)-12, which induces natural killer (NK) cells to produce interferon (IFN)-γ and drives the adaptive immune response toward a protective T-helper 1 (Th1) phenotype. IL-12 production by dendritic cells and IFN-γ production by NK and Th1 cells negatively regulates the Th2 response. IFN-γ activates macrophages to kill the intracellular pathogen. In genetically susceptible individuals, a counterregulatory Th2 cytokine response can suppress the Th1 response and impair classic macrophage activation, leading to parasite replication and uncontrolled infection. Counterprotective macrophage-derived cytokines can also inhibit the Th1 response, stimulate the Th2 response, and impair classical activation through an autocrine loop. Activating stimuli are shown by solid arrows, and deactivating stimuli are shown by dashed arrows.

T cells (Th1 subset). CD4 T cells are absolutely required, but immunity to cutaneous disease is also mediated by CD8 T cells via production of IFN-γ.[14] Both CD4 and CD8 T cells are required for an effective defense against murine visceral *L. donovani* infection, but the precise role of CD8 T cells is unclear. The generation of the Th1 response is critically dependent on CD40-CD40L-mediated IL-12 production and driven by NK cell–derived IFN-γ. IL-12 and STAT4 are required for the maintenance of immunity. Tumor necrosis factor-α (TNF-α) contributes to protective immunity by synergizing with IFN-γ to activate macrophages. Recently nuclear factor-kappa B (NF-κB) family members have been shown to regulate T-cell responses and immunity to *L. major* infection in mice.

Two subpopulations of CD4 T cells mediate immunity induced by primary infection. Effector memory cells, which are short-lived and dependent on the persistence of antigen, rapidly respond to secondary infection by migrating to the infected tissue and generating effector cytokines. Central memory T (T_{cm}) cells, which can be maintained in the absence of persistent antigen, circulate throughout the lymphatic system and upon secondary challenge migrate to and proliferate in the draining lymph node, gain the capacity to produce IFN-γ, and then migrate to the site of infection. Thus T_{cm} cells act as a reserve of antigen-reactive T cells that can expand and become effector T cells in response to secondary antigenic challenge.

The generation of RNI by activated macrophages is the primary mechanism of parasite killing in the murine model. Although IFN-γ-induced production of NO may not be detectable in human macrophages, inhibition of nitric oxide synthase 2 (NOS2) was shown to impair killing of intracellular *Leishmania*.

Several adaptive immune mechanisms promote parasite replication and disease.[15] The progression of murine *L. major* infection has been correlated with the expansion of Th2 cells and the production of IL-4, IL-5, and IL-10. In susceptible mice, IL-4 production within the first day of infection was shown to downregulate IL-12 receptor β-chain expression and drive the response to a Th2 phenotype. However, other nonsusceptible mouse strains appear to be able to overcome an early IL-4 response and develop a resistant phenotype, and susceptibility to some *L. major* strains is not strictly mediated by IL-4 (IL-13 and/or IL-10 may have a prominent role). The cytotoxic activity of CD8 T cells may promote cutaneous inflammation and lesion pathology.[14] The macrophage production of immune suppressive molecules, such as TGF-β or prostaglandin E_2 (PGE_2), may also contribute to susceptibility.

Peripheral blood mononuclear cells (PBMCs) isolated from patients with localized cutaneous leishmaniasis demonstrate a Th1 response to *Leishmania* antigens, and in the cutaneous lesion, there is an exuberant Th1 and granulomatous response that mediates parasite killing and localized tissue damage, which usually leads to a scar. Patients with mucosal leishmaniasis (ML) exhibit vigorous cellular immune responses characterized by high levels of TNF-α and Th1 and Th17 cytokines; it is postulated that this hyperresponsive state contributes to the prominent tissue destruction of ML. Patients with diffuse cutaneous leishmaniasis (DCL) resemble the progressive infection caused by *L. major* in BALB/c mice in that there are minimal or absent *Leishmania*-specific lymphoproliferative responses, and predominant Th2 cytokine expression. During active VL in humans, there is a marked depression of *Leishmania*-specific lymphoproliferative and IFN-γ responses, contraction of circulating memory T cells, and an absence of delayed-type hypersensitivity (DTH) response to parasite antigens. This T-cell unresponsiveness appears to be mediated, at least in part, by a suppressive effect of IL-10 and low levels of IL-12. Successful treatment of active disease restores an antigen-specific Th1 response.

Evasion of Host Immunity

The *Leishmania* parasite has numerous ways in which it adapts to and survives within the vertebrate host.[16] In skin, the promastigotes may be phagocytosed by neutrophils and macrophages, which, unlike DCs, do not actively participate in T-cell priming. Furthermore, the clearance of apoptotic neutrophils is likely to make macrophages more permissive to infection.[13] The parasite's surface LPG (and, to a lesser extent, the surface protein gp63) plays an important role in the entry and survival of *Leishmania* in the mammalian host by conferring complement resistance and by facilitating the entry of complement-opsonized parasites into the macrophage without triggering a respiratory burst. Macrophage phagosome–endosome fusion and phagolysosomal biogenesis are also inhibited by parasite LPG.

Leishmania-infected macrophages have diminished capacity to initiate and respond to a T-cell response, and the impaired

TABLE 30.2 Evidence for Autoimmune and Parasite-Induced Inflammatory Mechanisms in Chronic Chagas Disease

Evidence for Autoimmune-Mediated Disease	Evidence for Parasite-Induced Inflammatory Disease
Inflammatory disease presents in tissues with few or no parasites on routine histopathology studies	Sensitive parasite detection techniques (polymerase chain reaction [PCR], immunohistochemistry) correlate with the presence of parasites (or parasite material) and severity of inflammatory disease
Peculiar pattern of organ involvement (heart and gastrointestinal tract) in patients with chronic disease	Organs free of parasites (by sensitive parasite detection techniques) are also free of disease
Long delay in the onset of chronic disease following infection; only a minority of infected persons develops disease	Absence of effective cellular immune response (in mice or humans) almost invariably exacerbates rather than reduces the parasite burden and disease
Wide variability in the expression of disease among infected people	In chronically infected mice, the destruction of a transplanted heart is dependent on parasite infiltrating the transplanted tissue
Self-reactive antibodies and T cells demonstrable in infected people and in experimental animals. Level of antibodies to the ribosomal P protein (R13 peptide) and cardiac myosin (B13 antigen) correlate with cardiac disease	Degree of disease in hearts transplanted into chronically infected mice correlates with level of parasite burden in transplanted tissue
Transient or limited disease reported in experimental models following lymphocyte transfer	Reduction of parasite burden by chemotherapy usually leads to decreased tissue inflammation and disease

antimicrobial effector activity provides a safe haven for the intracellular parasite.[16] Infected macrophages have decreased synthesis of IL-1 and IL-12, and blunted IFN-γ-mediated activation through the disruption of signal transduction pathways involving JAK/STAT, protein kinase C, p38 MAPK, ERK, AP-1, and NF-κB. Signaling mediated by tyrosine phosphorylation is decreased by the rapid induction of the host phosphotyrosine phosphatase SHP-1. Conversely, there is increased synthesis of the immunosuppressive molecules IL-10, TGF-β, and PGE_2. Parasite factors and IL-4/IL-13 enhance expression of arginase by infected macrophages. Arginase promotes infection through depletion of L-arginine, enhanced production of polyamines, and reduced NO production. IL-10 produced by $CD4^+CD25^+$ T regulatory cells (Tregs) has an essential role in parasite persistence. Recently, it was shown that metastasizing parasites that cause mucosal disease harbor a high burden of *Leishmania* RNA virus, which subverts the host immune response and promotes parasite persistence through activation of Toll-like receptor 3 (TLR3).[17]

Trypanosoma cruzi

Pathogenesis

T. cruzi is transmitted to the mammalian host when the infectious metacyclic trypomastigote, which is deposited on skin in the feces of the reduviid insect vector while it takes a blood meal, is scratched into the wound or transferred to a mucous membrane (*e.g.*, the eyes). The trypomastigotes can infect almost any cell type and replicate as amastigotes in the cytoplasm. Eventually, the amastigotes transform back into trypomastigotes and rupture the cell to enter the bloodstream, from where they invade other cells or are picked up by another insect vector.

Following primary infection, the parasites replicate locally and then disseminate through the bloodstream to a variety of tissues. Muscle and glial cells are the most frequently infected cells, and acute myocarditis or meningoencephalitis can develop. In most cases, however, primary infection occurs without clinical symptoms, and the infected individual may enter an indeterminate phase of asymptomatic seropositivity. Only 10–30% of chronically infected individuals will ultimately develop symptomatic Chagas disease, usually involving the heart or the gastrointestinal (GI) tract. Pathologically, there are few parasites observed in cardiac tissue, but an intense chronic inflammatory infiltrate with fibrosis and loss of muscle fibers is evident. In the digestive tract, there is lymphohistiocytic infiltration of the myenteric plexuses, with reduction in the number of ganglion cells.

The tissue damage of acute *T. cruzi* infection is the result of a direct effect of the parasite and an indirect effect of the acute inflammatory response. In chronic infection, the balance between immune-mediated parasite containment and host-damaging inflammation determines the course of disease. The pathological mechanisms related to chronic Chagas disease are controversial and certain to be multifactorial. Trypanosomal clonal variation and host genetic polymorphisms both contribute to tissue tropism, parasite persistence, and disease severity. Whether tissue damage is caused directly by parasites or indirectly through parasite-driven inflammatory or autoimmune mechanisms, parasite persistence is a significant driver of disease (Table 30.2).[18,19] Autoimmunity could arise from molecular mimicry of self by parasite antigens or by the release of self molecules from damaged or dying host cells within the environment of an activated immune response. There is evidence for both of these autoimmune mechanisms. The production of IL-10 by *T. cruzi*–infected cells may downregulate the pathological cellular immune response.

Innate Immunity

The early innate immune response to *T. cruzi* infection is mediated by NK cells, DCs, and macrophages.[20] Macrophages and DCs exposed to *T. cruzi* trypomastigote antigens produce IL-12 and TNF, through a MyD88-dependent mechanism. MyD88-deficient mice had impaired inflammatory responses and host defense against *T. cruzi*. Immune activation through another protein involved in TLR2 signaling, the Toll/IL1R domain-containing adapter protein– inducing IFN-β (TRIF), promotes resistance through production of IFN-β and downstream expression of IFN-β inducible genes, such as the p47 guanosine triphosphatases (GTPase) *IRG47*. IL-12 activates NK cells to secrete IFN-γ, which synergizes with TNF to activate macrophages to control parasite replication. The generation of NO is the primary trypanocidal mechanism in murine macrophages. A number of trypomastigote antigens, including free glycosyl-phosphatidylinositol (GPI) anchors, glycoinositol phospholipids (GIPLs), GPI-linked glycoproteins, and GPI-mucins activate the innate immune response, in part through TLR2 and possibly TLR4. CpG motifs in the *T. cruzi* genome have also been shown to activate TLR9. Several other TLR-independent mechanisms of innate immunity, including the activation of nucleotide-binding oligomerization domain (NOD)–like receptors, have been identified.[20]

Adaptive Immunity

The innate immune response, through production of IL-12, type I IFNS, and other proinflammatory mediators, is critically linked to the generation of an effective adaptive immune response. Antibodies contribute to immunity within the bloodstream through opsonization, complement activation, and antibody-dependent cellular cytotoxicity.

Several lines of evidence establish the importance of T cells in adaptive immunity to *T. cruzi* infection. Parasite-specific CD4 and CD8 T cells are activated during infection, and mice lacking CD4 or CD8 T cells have impaired parasite control. CD8 T cells with cytotoxic activity against *T. cruzi*–infected cells have been identified in infected mice, and these cells confer protection when passively transferred to naïve mice.[21] In the early stage of infection, CD4 T cells are the predominant subset recruited to the myocardium, but activated CD8 T cells soon dominate the inflammatory process in cardiac tissue. IFN-γ and TNF production by parasite-specific CD8 T cells is more important than cytolytic activity in the control of infection.[21]

T. cruzi infection leads to a mixed Th1/Th2 cytokine response, and in general, the Th1/Th2 balance determines resistance or susceptibility. As noted earlier, IL-12/STAT4-dependent IFN-γ production by NK cells in early infection, and later by T cells, is critically important to protection. IL-4 does not appear to play a major role in susceptibility to *T. cruzi* infection, but IL-10 promotes parasite replication by inhibiting macrophage trypanocidal activity. IL-10 also plays a critical role in minimizing inflammation-mediated tissue pathology by regulating the Th1 and TNF responses. Polymorphisms in the IL-10 gene that lead to reduced IL-10 production are associated with increased severity of cardiomyopathy in patients with Chagas disease. Similarly, TGF-β has been shown to inhibit macrophage trypanocidal activity and increase parasitemia and mortality. In addition to these regulatory cytokines, the secretion of prostaglandins and NO, the induction of apoptosis of T and B cells, and the expansion of a myeloid suppressor cell population serve to control the intensity of the immune response.

Evasion of Host Immunity

A significant part of the pathogenesis following *T. cruzi* infection is its dissemination through the bloodstream to many tissues. *T. cruzi* bloodstream trypomastigotes resist complement lysis via a complement regulatory protein (GP-160). This parasite protein is functionally similar to mammalian decay-accelerating factor in that it inhibits C3 convertase formation and activation of the alternate complement pathway. *T. cruzi* invades host cells, in particular cardiomyocytes, by subverting a host plasma membrane repair pathway that promotes parasite persistence and tissue tropism.[22]

The establishment of chronic infection by *T. cruzi* is favored by a generalized depression of T-cell responses. A number of different mechanisms may contribute to this, including low IL-2 production or IL-2 receptor expression; downregulation of components of the T cell–receptor complex; T-cell receptor dysfunction; apoptosis of T cells; defects in the processing and presenting of antigens in the MHC class II (but not the class I) pathway, T-cell or macrophage suppressor activity, and PGE_2 production. Within the foci of myocarditis, apoptosis of both parasites and host cells occurs. The phagocytosis of these apoptotic cells by macrophages leads to their acquisition of an M2 phenotype, which enables parasite replication and persistence. The rapid escape of the parasite from the phagosome into the cytoplasm through the action of acid-activated porins enables the organism to avoid phagolysosomal enzymatic destruction.

Toxoplasma gondii

Pathogenesis

Transmission occurs via the ingestion of oocysts, which are shed in the feces of felines or via tissue cysts present in undercooked meat. Following the oral ingestion of cysts, phagocytes recruited to the gut lumen facilitate transepithelial migration into the lamina propria. These phagocytes have been termed "Trojan Horses," as they carry the parasite to new, unsuspecting host cells. Host-cell invasion starts with loose attachment facilitated by laminin. Intracellular tachyzoites replicate within a parasitophorous vacuole (PV) and ultimately leave the cell by using an active egress pathway. The released tachyzoites can disseminate to invade virtually any nucleated cell type, but mononuclear phagocytes are the preferred host cells. The toxoplasma SAG proteins mediate parasite attachment, and the transmembrane adhesins, MIC and AMA1, and RON family proteases play major roles in the myosin motor-driven process of host-cell invasion. Under pressure from the host immune response, tachyzoite replication is controlled and tissue cysts, containing slowly replicating bradyzoites, are formed. The tissue cysts persist as a chronic latent infection as long as the host immune function is intact. If the latently infected person is immunosuppressed, reactivation occurs, and tachyzoites are released to infect more cells. Because tissue cysts (bradyzoites) are found in proportionately larger numbers in the brain, reactivation of latent infection in the immunocompromised host most commonly manifests as encephalitis.

Innate Immunity

Similar to immunity to *Plasmodium* and *Leishmania,* IL-12, IFN-γ, TNF-α, and NK cells contribute to the control of the early stages of *T. gondii* infection.[23] $CD8\alpha^+$ DCs are the primary producers of IL-12, which is essential to host resistance. IL-12 is produced through a MyD88-dependent mechanism, at least in part, through the interaction of parasite-derived proteins and TLR2 and TLR11, and activation of CCR5.[24] IL-12-dependent activation of NK cells leads to IFN-γ production, which, in turn, activates macrophages to limit parasite replication. T cells expressing the γδ receptor and neutrophils are also an early source of IFN-γ and TNF-α. Recruitment of inflammatory macrophages to the site of infection is crucial to control of parasite replication and dissemination. IFN-γ-induced macrophage activation is the key effector mechanism leading to parasite killing. In mice, the induction of immunity-related GTPases (IRG), which damage the parasitophorous vacuole membrane and kill *T. gondii* within the cytosol, is the primary macrophage effector mechanism. The generation of reactive nitrogen and oxygen intermediates, degradation of tryptophan, and the production of LTs have also been implicated in the control of *T. gondii* in human macrophages. However, humans lack the entire IRG family, as well as TLR11, and the mechanism(s) of early IFN-γ–mediated macrophage activation and parasite control in humans is unclear.

Adaptive Immunity

Although serum antibodies can be used in the diagnosis of *T. gondii* infection, the systemic antibody response does not play a role in adaptive immunity. Mucosal IgA, however, does provide resistance to oral infection with *T. gondii* cysts. CD4 and CD8 T cells are highly activated during infection and are essential for adaptive immunity. As such, patients with defects in T cell–mediated immune responses (*e.g.*, patients with AIDS) are at risk for reactivation of latent infection.

IL-12 from DCs drives the differentiation of T cells into type-1 CD4 and CD8 T cells that are essential to adaptive immunity. CD4 and CD8 T cells act synergistically to prevent cyst reactivation during chronic latent infection. Parasite-specific cytolytic T cells have been demonstrated; however, CD8 T cells mediate protection primarily through the generation of IFN-γ. IL-12 drives the generation of terminally differentiated CD8 effector T cells, which express the killer cell lectin-like receptor G1 (KLRG1) and high levels of granzyme B and IFN-γ. The combination of IL-7 and IL-15 is required for T_{cm}-cell differentiation.

Antiinflammatory molecules, particularly IL-10 and IL-27 made by Th1 cells, play an important role in modulating the adaptive immune response and restricting host tissue damage.[25]

Evasion of Host Immunity

T. gondii escapes early macrophage killing in a number of ways.[25] Virulent parasites are protected by localization to the parasitophorous vacuole, which does not fuse with host cell lysosomes (probably because the PV membrane proteins are of parasite rather than host origin), and so the vacuole is not acidified to kill the parasite. The infected macrophage is also a suboptimal target for T cell–induced immunity because of reduced expression of MHC class II and costimulatory molecules. Infection also induces the production of counterregulatory molecules, such as IL-10, TGF-β, lipoxin A4. These not only downregulate a potentially pathological host inflammatory response but also inhibit Th1 induction and macrophage antimicrobial activity. In addition, *T. gondii* interferes with normal macrophage signaling. For example, infection inhibits DNA binding of signal transducer and activator of transcription 1 (STAT1) and NF-κB by interfering with chromatin structure, and promotes antiinflammatory pathways downstream of the suppressor of cytokine synthesis proteins. The parasite also has several virulence proteins, including ROP5 and ROP16, which bind to the parasitophorous vacuole and reduce accumulation of the IRGs. ROP proteins also activate host STAT3 and STAT6 driving an M2 macrophage phenotype that is permissive to infection.[26]

Entamoeba histolytica

Pathogenesis

E. histolytica causes asymptomatic intestinal colonization, acute diarrhea, dysentery, colitis, liver abscess, and, rarely, disseminated disease. Susceptibility to amebiasis is determined by the host's nutritional status, intestinal microflora, genetics, and gender.[27] *E. histolytica* cysts are ingested through consumption of food or water contaminated by feces. After excystation, the trophozoites degrade the colonic mucus barrier through secretion of proteases and glycosidases. The trophozoites adhere to the colonic epithelial cells by a galactose/*N*-acetylgalactosamine-inhibitable lectin (Gal/GalNAc), which leads to NF-κB activation and proinflammatory cytokine release. Epithelial cells with adherent amebae undergo microvilli shortening and apical separation that allows the parasite to penetrate between the epithelial cells and degrade connective tissue, resulting in ulceration of the mucosa and submucosa (Fig. 30.3). The trophozoites can lyse multiple cell types, including neutrophils, which release enzymes that further damage the tissue. The cytotoxic effects of amebae are mediated by a secreted cysteine protease, the Gal/GalNAc lectin, phospholipase A, and contact-dependent cytolysis, in which an ion channel (amebapore) is inserted into the membrane of target cells. *E. histolytica* can induce apoptosis in mammalian cells by a caspase-dependent, TNF-α/Fas–independent process. Amebic liver abscesses develop when trophozoites erode through the intestinal submucosa, enter the portal circulation, and disseminate to the liver.[28] Comparison of the transcriptional profiles of virulent and nonvirulent *Entamoeba* species has identified novel virulence factors in *E. histolytica.*[29]

Innate Immunity

Adherence of trophozoites to intestinal epithelial cells stimulates the release of a variety of proinflammatory mediators, including IL-1α, IL-1β, IL 6, IL-8, and TNF-α, which triggers the recruitment of neutrophils and macrophages to the site of invasion. Following activation by IFN-γ and TNF-α, neutrophils and macrophages become amebicidal through the release of reactive oxygen species and production of nitric oxide, respectively.[28] Invariant natural killer T (iNKT) cells, which comprise about 30% of hepatic lymphocytes, are stimulated by amebic lipopeptidephosphoglycan (LPPG) to release IFN-γ. Production of IFN-γ reduces parasite burden and controls abscess formation.[28]

Adaptive Immunity

Secretory IgA responses against *E. histolytica* have been well characterized and shown to correlate with protection against infection and disease. For this reason, vaccines designed to generate IgA antibodies against the Gal/Gal/NAc lectin have been highly efficacious in preventing experimental *E. histolytica* infection in mice and baboons.[28] However, the role of antilectin IgG is unclear, and protection may be subclass-dependent. Cell-mediated immunity is also critical for host defense against *E. histolytica.* IFN-γ–producing CD4 T cells provide protection through their ability to activate macrophages and neutrophils to the parasite. CD8 T cells mediate protection through the secretion of IL-17, a key player in the secretion of mucin and antimicrobial peptides, recruitment of neutrophils, and IgA transport across the epithelial barrier. Impairment of cell-mediated immunity results in parasite dissemination. Indeed, patients with human immunodeficiency virus infection and coinfected with *E. histolytica* have high rates of invasive amebiasis, liver abscesses, and seroconversion.[29]

Evasion of Host Immunity

E. histolytica utilizes a number of strategies to circumvent the immune defenses of the host. It resists complement-mediated lysis during hematogenous spread by proteolytic degradation of C3a and C3b. In addition, the Gal/GalNAc lectin binds to C8 and C9, preventing assembly of the C5b-9 membrane attack complex. The cytolytic capability of *E. histolytica* affords protection from neutrophils, macrophages, and eosinophils, unless these cells are activated. Cytolysis by *E. histolytica* can occur via necrosis and apoptosis. Trophozoites also inhibit the macrophage respiratory burst and the production of IL-1 and TNF-α. A protective antibody response is subverted by the degradation of IgA and IgG by amebic cysteine proteases and by capping, ingesting, or

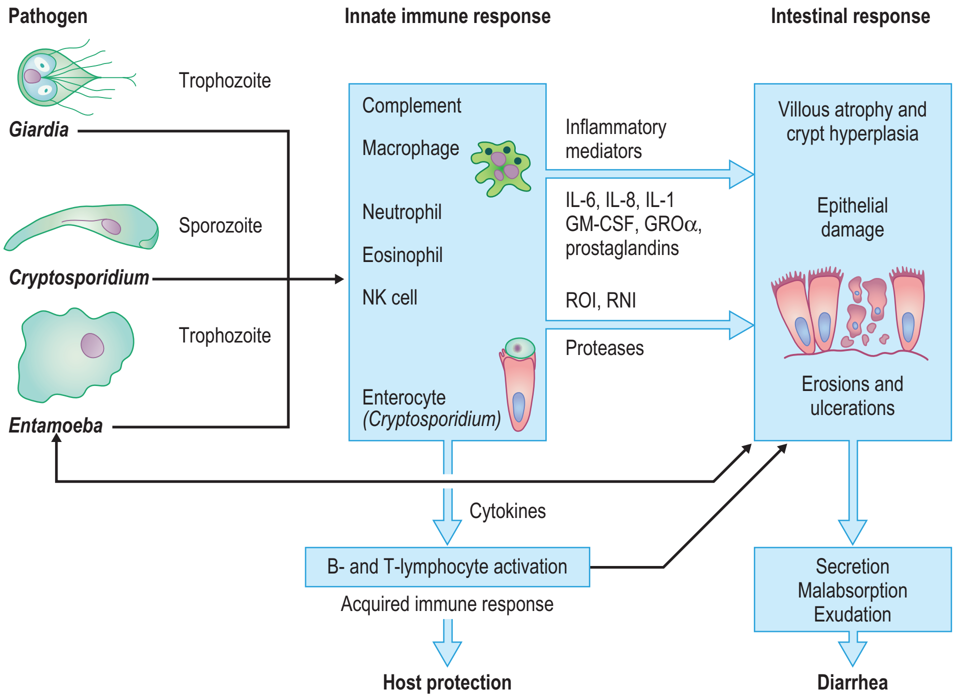

Pathogen
Innate immune response
Intestinal response
Trophozoite
Giardia
Sporozoite
Cryptosporidium
Trophozoite
Entamoeba
Complement
Macrophage
Neutrophil
Eosinophil
NK cell
Enterocyte
(Cryptosporidium)
Inflammatory
mediators
IL-6, IL-8, IL-1
GM-CSF, GROα,
prostaglandins
ROI, RNI
Proteases
Villous atrophy and
crypt hyperplasia
Epithelial
damage
Erosions and
ulcerations
Cytokines
B- and T-lymphocyte activation
Acquired immune response
Secretion
Malabsorption
Exudation
Host protection
Diarrhea

FIG 30.3 **Immunopathogenesis of Intestinal Protozoal Pathogens.** After adherence (*Giardia* and *Entamoeba*) or epithelial invasion (*Entamoeba* and *Cryptosporidium*), there is release of various inflammatory mediators from macrophages and neutrophils. This causes the activation of resident phagocytes and recruitment of phagocytes into the lamina propria. Enterocyte death can be due to direct action of the parasites or to immune-mediated damage from complement, cytotoxic lymphocytes, proteases, and reactive oxygen and nitrogen intermediates (ROI and RNI, respectively). The inflammatory mediators also act on enterocytes and the enteric nervous system, inducing the secretion of water and chloride. In response to enterocyte damage, under the influence of activated T lymphocytes, the crypts undergo hyperplasia, and the villi become shorter (villous atrophy). The immature hyperplastic cells have poor absorptive ability but retain secretory ability. Damage to the epithelium can cause leakage (exudation) from lymphatics and capillaries. Similar mechanisms are probably responsible for the diarrhea that occurs in infection with *Cyclospora* and *Isospora*. *Isospora* is unique in causing an eosinophilic infiltrate.

shedding ameba-specific antibodies.[27] Amebic proteases can also cleave the Fc region so that interaction with host cell surface receptors is avoided. Another secreted product of *E. histolytica*, monocyte locomotion inhibition factor (MLIF), inhibits monocyte locomotion and the monocyte and neutrophil respiratory burst and NO production; enhances antiinflammatory cytokine and chemokine release from host cells; and alters adhesion molecule expression on macrophages. The suppression of host macrophage NO production by an array of trophozoite secretory products, including parasite-derived PGE_2, is a major factor in the persistence of amebic liver abscesses. In chronic infection, *E. histolytica* promotes the development of Tregs that suppress the proliferation of responder T cells by releasing IL-10, TGF-β, and IL-35.[28]

Giardia lamblia

Pathogenesis

Recent studies of *Giardia lamblia* have identified eight genotypes, two of which infect humans (assemblages A and B).[30] The severity of giardiasis ranges from asymptomatic carriage to chronic watery diarrhea, epigastric pain, nausea, vomiting, and weight loss, depending on host factors and the virulence of the *Giardia* strain.[30,31] Recent studies suggest even with treatment the parasite can elicit intestinal complications that persist for years. Younger age, malnutrition, and immunodeficiency increase the risk of severe disease.[31] Infection is initiated following the ingestion of food or water contaminated with *G. lamblia* cysts. Exposure to stomach acids induces the excystation process that releases two trophozoites into the lumen of the proximal small intestine. Colonization occurs when the parasite attaches to the intestinal epithelium and begins to reproduce by binary fission. Trophozoites remain within the lumen and do not invade the epithelial barrier. Parasite migration into the lower intestine triggers encystation, allowing the organism to survive when excreted into the environment.

The *G. lamblia* trophozoite initiates adherence to the intestinal epithelium via a surface mannose-binding lectin. Histopathological changes in symptomatic giardiasis range from a normal appearance to increased crypt-villous ratios, epithelial damage, and chronic inflammatory infiltrate in the lamina propria (see

Fig. 30.3). The factors responsible for the structural changes in the small bowel are not well defined but may include injury from adherence, parasite-induced apoptosis of epithelial cells, and the release of cytotoxins, including proteases. Additional epithelial damage may be mediated by the host cellular immune response. Diarrhea arises from epithelial barrier dysfunction, reduction in microvillous surface area, chloride hypersecretion, and glucose and sodium malabsorption.[31]

Innate Immunity

As *G. lamblia* does not invade the intestinal epithelium, host defense and immune factors present within the lumen are essential for preventing and controlling infection. Increased intestinal motility contributes to *G. lamblia* clearance, by impairing the ability of the parasite to attach to the epithelium and resist the luminal bulk flow. Antimicrobial peptides derived from Paneth cells, including cryptdins, neutrophil defensin, and cathelicidin, effectively kill *G. lamblia* trophozoites *in vitro*.

NO, produced by both epithelial cells and macrophages, inhibits excystation trophozoite division. Despite the protective effect of NO, the parasite can circumvent this defense by competing with host cells for arginine uptake.[30] Mast cells have a significant role in protecting against the parasite. Mice deficient in mast cells fail to clear *G. lamblia* infection, in part because they are unable to mount parasite-specific IgA. Mast cells also contribute to B-cell survival, activation and differentiation into plasma cells, and together with NO induce peristalsis.

Adaptive Immunity

Several lines of evidence suggest the importance of the humoral immune response in the control of giardiasis. Infection with *Giardia* results in the production of anti-*Giardia* antibodies in the serum and mucosal secretions. Patients with severe B-cell defects or selective IgA deficiency have an increased risk of developing chronic infections.[30] Studies in mice have demonstrated key functions of secretory IgA and the polymeric immunoglobulin receptor, which is responsible for IgA transport into the intestinal lumen in controlling parasite burden and eliminating infection.[30]

There is also evidence for a role of T cell–dependent immunity in the control of giardiasis. A reduction or absence of CD4 T cells can lead to chronic infection. IFN-γ- and IL-17A-producing CD4 T cells develop following infection[32] and are important for mediating parasite clearance. Although the mechanisms are still being defined, IL-17A is likely involved in modulating transport of IgA into the intestinal lumen. Epidemiological studies indicate that partial immunity is acquired from *Giardia* infection, which leads to reduced risk and severity of subsequent infections.[30]

Evasion of Host Immunity

Giardia evades the host humoral immune response by undergoing surface antigenic variation by altering a group of variant-specific surface proteins (VSPs). Selection occurs by an immune-mediated process because switching occurs when intestinal anti-VSP IgA responses are first detected.[30] *G. lamblia* also produces a protease that cleaves IgA. Although *Giardia* activates dendritic cells for antigen presentation, it also inhibits IL-12 production, in part by enhancing IL-10 release; the net result is the dampening of a local antiparasitical inflammatory response.[30] The trophozoite also releases arginine deiminase, which degrades arginine, making it less available for host NO production.[31]

Cryptosporidium parvum AND *Cryptosporidium hominis*

In humans, there are four intestinal coccidians that are intracellular parasites of enterocytes: *Isospora belli*, *Cyclospora cayetanensis*, and two species of *Cryptosporidium*, *C. parvum*, and *C. hominis*. Of the four coccidians, *Cryptosporidium* has the greatest epidemiological significance: in 1993, a huge outbreak involving 403 000 persons occurred in Milwaukee, Wisconsin, in the United States. Because of their similarity, only the immunology of cryptosporidiosis will be discussed.

Pathogenesis

Cryptosporidium typically causes self-limited (but often prolonged) diarrhea in the immunocompetent host. However, in the immunocompromised host *Cryptosporidium* can cause severe diarrhea, with malabsorption and wasting, and cholangiopathy. Infection begins with the ingestion of food or water contaminated with oocysts. The acidic environment of the stomach induces excystation and the release of four sporozoites into the small intestines. Glycoproteins on the surface of sporozoites facilitate attachment and invasion of epithelial cells. After entry into epithelial cells, the parasite resides within a unique intracellular but extracytoplasmic vacuole, that protects the pathogen from environmental and host insults. Inside a vacuole the sporozoites develop into trophozoites and undergo schizogony, with a resultant merozoite-containing schizont. The merozoites are extruded and invade neighboring epithelial cells. The merozoites may continue an asexual cycle or develop into macro- or microgametes that fuse to form oocysts. Before being excreted in feces, oocysts undergo sporulation to become infectious. Histologically, infection causes villous atrophy and blunting, and crypt hyperplasia with increased infiltration of lymphocytes, macrophages, and plasma cells.[33] Intraepithelial lymphocytes are uncommon; neutrophils and occasional eosinophils are present between the epithelium and the lamina propria. Disorganized cells undergoing necrosis replace normal enterocyte architecture (see Fig. 30.3). There is an association between the degree of intestinal injury and malabsorption and the intensity of infection, as measured by oocyst excretion.

The neuropeptide substance P, which is produced by endothelial cells, lymphocytes, and monocytes in the lamina propria, contributes to diarrhea by increasing intestinal chloride secretion and glucose malabsorption. Increased expression of substance P is observed in patients with AIDS as well as cryptosporidiosis and severe diarrhea.[34]

Innate Immunity

The type I and type II IFNs play a key role in the innate protective response against cryptosporidium.[35] Because of the parasite's intracellular location near the luminal surface of the enterocyte, the macrophages of the lamina propria are spatially isolated from the parasite. Thus the intestinal epithelium mounts its own assault on the invading microbe through TLR2/TLR4–dependent activation of NF-κB and release of the microbicidal peptide β-defensin-2, TNF-α, and the chemokines IL-8, RANTES, and GRO-α, which act as chemoattractants and activators of neutrophils. In patients with AIDS and cryptosporidiosis, the HIV Tat protein may sabotage host defense against *Cryptosporidium* by inhibiting cholangiocyte TLR4 expression.[33]

IL-15, produced by activated monocytes, stimulates NK-cell proliferation, cytotoxicity, and cytokine production, including IFN-γ. There is significant IL-15 expression in the jejunal mucosa in immunocompetent patients with cryptosporidiosis, and IL-15 levels inversely correlate with parasite burden. However, in patients with AIDS who have chronic uncontrolled cryptosporidiosis, IL-15 is undetectable.[33]

Mannose-binding lectin (MBL) is a serum protein that binds to various pathogens, including *Cryptosporidium.* Upon binding, MBL activates complement, thereby promoting opsonization and phagocytosis. Low serum levels of MBL, which may result from malnutrition or polymorphisms in the *MBL2* gene, increase susceptibility to cryptosporidiosis.[33] Infected intestinal cells also release TGF-β, which decreases necrosis and stimulates the synthesis of extracellular matrix proteins, thereby limiting epithelial damage.

Prostaglandins E_2 and $F_2\alpha$, released by infected enterocytes, not only promote secretory diarrhea but also upregulate mucin production, which may hinder parasite attachment. In addition, these prostaglandins stimulate the release of β-defensin-2, which has direct anticryptosporidial activity and is chemotactic for T cells and DCs.[33]

Adaptive Immunity

Cell-mediated immunity plays an important role the resolution of cryptosporidiosis and protection from reinfection. In immunocompetent, adult mice CD4 intraepithelial lymphocytes (IELs) initiate early control of infection, whereas cytotoxic CD8 IELs appear later and function in parasite elimination. Resolution of infection depends on a balance of Th1 cytokines (IFN-γ, IL-18) needed to control the infection and Th2 cytokines (IL-4, IL-10, and IL-13) that limit immunopathological damage. In mice, γδ T cells are rapidly recruited to control cryptosporidial infection, but their role in human infection is unknown. Severe intestinal disease or biliary involvement is usually seen in patients with AIDS whose CD4 count is <50/μL.[33] Patients with HIV infection who have CD4 counts >200/μL usually experience self-limited disease.

The role of humoral immunity is less clear. Secretory antibodies produced in the intestines are thought to impair parasite attachment to epithelial cells. Indeed, immunoglobulin deficiencies are often associated with persistent or recurrent infections. However, specific anti-*Cryptosporidium* IgA levels have been reported in patients with AIDS who are infected with the parasite, suggesting polyreactive T cell–independent antibodies are not sufficient for parasite eradication.[33] In experimentally infected human volunteers, serum IgM and IgG directed toward sporozoite proteins protect against the development of symptoms, but not infection.

Evasion of Host Immunity

Cryptosporidium evades host defenses primarily by exerting control over infected enterocyte apoptosis. One of the upregulated genes is osteoprotegerin, which inhibits apoptosis by acting as a decoy receptor for TNF-related apoptosis inducing ligand (TRAIL).[34] Control of host apoptosis is complex; early inhibition of apoptosis by NF-κB activation allows the parasite to complete its life cycle, whereas the late promotion of apoptosis facilitates merozoite release. Nevertheless, the infected cells secrete FasL, which promotes apoptosis in uninfected bystander cells. In this way, the host counters the antiapoptotic activity of the parasite by surrounding the parasitized cells by a zone of apoptotic cells.

Trichomonas vaginalis

Pathogenesis

Trichomonas vaginalis is a flagellated protozoan parasite of the human urogenital tract that exists only as a trophozoite. It causes vaginitis, cervicitis, and urethritis. Its adherence to the vaginal squamous epithelium is facilitated by a number of adhesins. *Trichomonas* causes tissue damage by contact-dependent cytolysis caused by pore-forming proteins and proteases, and secretion of a glycoprotein cell-detaching factor that causes sloughing of the vaginal epithelium. Levels of the cell-detaching factor correlate with the severity of the disease, and vaginal antibodies directed against this factor modulate its effects. Inflammation in the genital mucosa and submucosa leads to copious secretions, and the surface epithelium may slough, causing focal erosions and hemorrhage.

The increased risk of HIV transmission in women with trichomoniasis may result from increased recruitment of inflammatory cells, mucosal erosion, or degradation of secretory leukocyte protease inhibitor (SLPI) by trichomonal proteases. Lower levels of SLPI are found in the vaginal fluid of women with trichomoniasis, which can lead to increased tissue damage and HIV transmission.[36] The lipophosphoglycan of *Trichomonas* induces production of the chemokines IL-8 and CCL20, which, which can also facilitate HIV infection by promoting DC recruitment.

Innate Immunity

Although trichomoniasis has recently received increased attention as a risk factor for HIV transmission and obstetrical complications, there is little known about the protective immune response against this organism. *Trichomonas* secretes a factor that promotes neutrophil chemotaxis, causing profuse leukorrhea, but the oxidative microbicidal mechanisms of the neutrophils have decreased efficacy in the anaerobic vaginal environment. Activated macrophages can destroy trichomonads in a T and B cell–independent manner and release IL-lβ and TNF-α, which are chemotactic for neutrophils.[37] *Trichomonas* induces neutrophil apoptosis, and macrophage clearance of these apoptotic cells causes release of IL-10, which may contribute to resolution of the inflammatory response.[38]

Adaptive Immunity

Repeated infections with *T. vaginalis* do not induce immunity; however, the infection is self-limited in most cases, so there are effective mechanisms of host defense. *T. vaginalis* induces the production of antibodies in both the serum and vaginal secretions. The serum antibody response correlates with active infection, and serum, but not vaginal, IgG from infected patients displays complement-mediated lytic activity against trichomonads in culture.[38]

Evasion of Host Immunity

Although *T. vaginalis* activates the alternative pathway of complement, the cervical mucus and menstrual blood are low in complement. Menstrual blood also supplies iron, which upregulates trichomonal adhesins and cysteine proteases, causing the degradation of complement component C3 bound to the surface of the parasite. Parasite virulence is thus enhanced, and the symptoms are exacerbated during menses. Cysteine proteases secreted by *T. vaginalis* also degrade immunoglobulins, sabotaging the antibody response. The parasite also secretes soluble antigens that act as decoys for neutralizing antibodies or cytotoxic T cells and disguises itself by binding to host plasma proteins. Phenotypic variation of surface markers is also a means of antibody evasion.[38]

ON THE HORIZON

Anticipated Approaches to Improved Control of Protozoal Diseases

- Gaining insight into the role of host genetics, gender, intestinal microbiota, and nutritional status in the outcome of protozoal infections
- Defining host immune and inflammatory mechanisms that promote disease caused by protozoal pathogens
- Understanding mechanisms for sustained generation of protective immunity against protozoal pathogens
- Developing vaccines for several blood and tissue protozoa
- Developing new host-directed treatment strategies to overcome immune evasion by protozoal pathogens

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Kumar R, Engwerda C. Vaccines to prevent leishmaniasis. Clin Translat Immunol 2014;3:e13.
2. Bethony JM, Cole RN, Guo X, et al. Vaccines to combat the neglected tropical diseases. Immunol Rev 2011;239:237–70.
3. Crompton PD, Pierce SK, Miller LH. Advances and challenges in malaria vaccine development. J Clin Investig 2010;120:4168–78.
4. Vazquez-Chagoyan JC, Gupta S, Garg NJ. Vaccine development against Trypanosoma cruzi and Chagas disease. Adv Parasitol 2011;75:121–46.
5. Gazzinelli RT, Kalantari P, Fitzgerald KA, et al. Innate sensing of malaria parasites. Nat Rev Immunol 2014;14:744–57.
6. Crompton PD, Moebius J, Portugal S, et al. Malaria immunity in man and mosquito: insights into unsolved mysteries of a deadly infectious disease. Annu Rev Immunol 2014;32:157–87.
7. Schwenk RJ, Richie TL. Protective immunity to pre-erythrocytic stage malaria. Trend Parasitol 2011;27:306–14.
8. Langhorne J, Ndungu FM, Sponaas AM, et al. Immunity to malaria: more questions than answers. Nat Immunol 2008;9:725–32.
9. Stanisic DI, Barry AE, Good MF. Escaping the immune system: how the malaria parasite makes vaccine development a challenge. Trend Parasitol 2013;29:612–22. PubMed PMID: 24176554.
10. Guizetti J, Scherf A. Silence, activate, poise and switch! Mechanisms of antigenic variation in Plasmodium falciparum. Cell Microbiol 2013;15: 718–26.
11. Gurung P, Kanneganti TD. Innate immunity against Leishmania infections. Cell Microbiol 2015;17:1286–94.
12. Ribeiro-Gomes FL, Sacks D. The influence of early neutrophil-Leishmania interactions on the host immune response to infection. Frontiers Cell Infect Microbiol 2012;2:59.
13. Peters NC, Egen JG, Secundino N, et al. In vivo imaging reveals an essential role for neutrophils in leishmaniasis transmitted by sand flies. Science 2008 15;321:970–4.
14. Novais FO, Scott P. CD8+ T cells in cutaneous leishmaniasis: the good, the bad, and the ugly. Semin Immunopathol 2015;37:251–9.
15. Soong L, Henard CA, Melby PC. Immunopathogenesis of non-healing American cutaneous leishmaniasis and progressive visceral leishmaniasis. Semin Immunopathol 2012;34:735–51.
16. Geiger A, Bossard G, Sereno D, et al. Escaping deleterious immune response in their hosts: lessons from trypanosomatids. Frontiers Immunol 2016;7:212.
17. Ives A, Ronet C, Prevel F, et al. Leishmania RNA virus controls the severity of mucocutaneous leishmaniasis. Science 2011;331:775–8.
18. Bonney KM, Engman DM. Autoimmune pathogenesis of Chagas heart disease: looking back, looking ahead. Am J Pathol 2015;185:1537–47.
19. Machado FS, Dutra WO, Esper L, et al. Current understanding of immunity to *Trypanosoma cruzi* infection and pathogenesis of Chagas disease. Semin Immunopathol 2012;34:753–70.
20. Kayama H, Takeda K. The innate immune response to *Trypanosoma cruzi* infection. Microbes Infect 2010;12:511–17.
21. Tarleton RL. CD8+ T cells in *Trypanosoma cruzi* infection. Semin Immunopathol 2015;37:233–8.
22. Fernandes MC, Andrews NW. Host cell invasion by *Trypanosoma cruzi*: a unique strategy that promotes persistence. FEMS Microbiol Rev 2012;36: 734–47.
23. Butler NS, Harris TH, Blader IJ. Regulation of immunopathogenesis during *Plasmodium* and *Toxoplasma* infections: more parallels than distinctions? Trend Parasitol 2013;29:593–602.
24. Yarovinsky F. Innate immunity to Toxoplasma gondii infection. Nat Rev Immunol 2014;14:109–21.
25. Aliberti J. Host persistence: exploitation of anti-inflammatory pathways by *Toxoplasma gondii*. Nat Rev Immunol 2005;5:162–70.
26. Hunter CA, Sibley LD. Modulation of innate immunity by Toxoplasma gondii virulence effectors. Nat Rev Microbiol 2012;10:766–78.
27. Moonah SN, Jiang NM, Petri WA Jr. Host immune response to intestinal amebiasis. PLoS Pathog 2013;9:e1003489.
28. Nakada-Tsukui K, Nozaki T. Immune response of amebiasis and immune evasion by *Entamoeba histolytica*. Frontiers Immunol 2016;7:175.
29. Baxt LA, Singh U. New insights into Entamoeba histolytica pathogenesis. Curr Opin Infect Dis 2008;21:489–94.
30. Lopez-Romero G, Quintero J, Astiazaran-Garcia H, et al. Host defences against *Giardia lamblia*. Parasite Immunol 2015;37:394–406.
31. Cotton JA, Beatty JK, Buret AG. Host parasite interactions and pathophysiology in *Giardia* infections. Int J Parasitol 2011;41:925–33.
32. Saghaug CS, Sornes S, Peirasmaki D, et al. Human memory CD4+ T cell immune responses against *Giardia lamblia*. Clin Vaccine Immunol 2016;23:11–18.
33. Kothavade RJ. Challenges in understanding the immunopathogenesis of *Cryptosporidium* infections in humans. Eur J Clin Microbiol Infect Dis 2011;30:1461–72.
34. Pantenburg B, Dann SM, Wang HC, et al. Intestinal immune response to human *Cryptosporidium* sp. infection. Infect Immun 2008;76:23–9.
35. Barakat FM, McDonald V, Foster GR, et al. *Cryptosporidium parvum* infection rapidly induces a protective innate immune response involving type I interferon. J Infect Dis 2009;200:1548–55.
36. Thurman AR, Doncel GF. Innate immunity and inflammatory response to Trichomonas vaginalis and bacterial vaginosis: relationship to HIV acquisition. Am J Reprod Immunol 2011;65:89–98.
37. Han IH, Goo SY, Park SJ, et al. Proinflammatory cytokine and nitric oxide production by human macrophages stimulated with *Trichomonas vaginalis*. Korean J Parasitol 2009;47:205–12.
38. Petrin D, Delgaty K, Bhatt R, et al. Clinical and microbiological aspects of *Trichomonas vaginalis*. Clin Microbiol Rev 1998;11:300–17.

MULTIPLE-CHOICE QUESTIONS

1. Which one of the following is characteristic of the immunopathogenesis of cerebral malaria?
   - **A.** Absence of *Plasmodium*-specific antibodies
   - **B.** Impaired interferon (IFN)-γ mediated production of nitric oxide
   - **C.** Cytoadhesion of parasites and leukocytes to vascular endothelium
   - **D.** Immune-mediated destruction of infected erythrocytes
   - **E.** Failure of leukocyte production of leukotriene/tumor necrosis factor (LT/TNF)
2. Which one of the following is NOT a mechanism through which *Leishmania* can evade the host immune response?
   - **A.** Complement resistance mediated by the parasite's surface lipophosphoglycan
   - **B.** Blunted interferon (IFN)-γ-mediated macrophage activation through the disruption of signal transduction pathways
   - **C.** Increased synthesis of the immunosuppressive molecules interleukin (IL)-10, transforming growth factor (TGF)-β, and prostaglandin (PG)E_2
   - **D.** Depletion of L-arginine through the activity of host arginase
   - **E.** Suppression of an early neutrophil response at the site of infection
3. Which one of the following is a component of the innate immune response to *Cryptosporidium infection?*
   - **A.** Decreased intestinal mucin production
   - **B.** Increased type I and type II interferon production
   - **C.** Decreased nuclear factor-kappa B (NF-κB) activation
   - **D.** Antigen-specific CD8 intraepithelial T-cell responses
   - **E.** Antigen-specific immunoglobulin A (IgA) responses

Immune Responses to Helminth Infection

Subash Babu, Thomas B. Nutman

Parasitical helminths are complex eukaryotic organisms, characterized by their ability to maintain long-standing, chronic infections in human hosts, sometimes lasting decades. Hence, parasitical helminths are a major health care problem worldwide, infecting more than 2 billion people, mostly in resource-constrained countries (Fig. 31.1). Common helminth infections include those with intestinal helminths, and filarial and schistosome infections are a major medical, social, and economic burden to the countries in which these infections are endemic. Chemotherapy, although highly successful in some areas, still suffers from the disadvantages of the length of treatment, the logistics involved in the distribution of drugs, and, in some cases, the emergence of drug resistance. Vector control measures are at best an adjunct measure in the control of helminth infections but also suffer from the same social, logistic, and economic obstacles as those for mass chemotherapy. Therefore the study of the immune responses to helminth infections attains great importance both in terms of understanding the parasite strategies involved in establishing chronic infection and in the delineation of a successful host immune response to develop protective vaccines against infection.

SPECTRUM OF HOST–PARASITE INTERACTIONS

Helminths have characteristically complex lifecycles with many developmental stages.[1] Thus the host is exposed during the course of a single infection to multiple lifecycle stages of the parasites, each stage with a shared as well as a unique antigenic repertoire. Thus *Schistosoma mansoni* infection begins with penetration of the skin of humans exposed to infested waters by the free-swimming cercariae, which then develop into tissue-dwelling schistosomula. In the liver and mesenteric veins, schistosomula differentiate into sexually dimorphic adult worms, which then mate, and the resultant eggs produced migrate through tissues into the lumen of the intestine or bladder for environmental release. Similarly, in lymphatic filarial infection, the host is exposed to infective-stage larvae in skin, lymph nodes, and lymphatics; to adult worms in lymph nodes and lymphatics; and finally to microfilariae in the peripheral circulation. Hence, the host–helminth interaction is complex not only because of the multiple lifecycle stages of the parasite but also because of the tissue tropism of the different stages.

Antigenic differences among the lifecycle stages can lead to distinct immune responses that evolve differentially over the course of a helminth infection. In addition, depending on the location of the parasite, the responses are compartmentalized (intestinal mucosa and draining lymph nodes in intestinal nematode infection, or skin/subcutaneous tissue and draining lymph nodes in onchocerciasis) or systemic (lymphatic filariasis or schistosomiasis). Moreover, the migration patterns of the parasite might elicit varied cutaneous, pulmonary, and intestinal inflammatory pathologies, as seen, for example, in *Ascaris* or *Strongyloides* infection during their migratory phase. This is further complicated by the fact that human hosts are often exposed to multiple lifecycle stages of the parasite at the same time. Thus a patient with chronic infection with lymphatic filariasis harboring adult worms and microfilariae might be exposed to insect bites, thereby transmitting the infective-stage parasite. The immune response that ensues will not only be a reaction to the invading organism but will also bear an imprint of the previous exposures and the concurrent infection.

KEY CONCEPTS

Helminth Infection

- Divided into nematodes, trematodes, and cestodes
- Produce chronic infections that can persist for decades
- Characteristically cause morbidity rather than mortality
- Multicellular parasites that do not multiply in the definitive host but can reproduce sexually to produce larval stages that ensure continued transmission

Helminth infections can elicit a spectrum of clinical manifestations mirroring diversity in host immune responses. For example, in lymphatic filariasis, most infected individuals remain clinically asymptomatic despite harboring significant worm burdens; this is thought to reflect the induction of parasite-specific tolerance in the immune system. Others exhibit acute manifestations, including fever and lymphadenopathy, and this is thought to reflect inflammatory processes induced by incoming larvae, dying worms, or superadded infections. Individuals who mount a strong but inappropriate immune response end up with lymphatic damage and subsequent immune-mediated pathology—hydrocele and elephantiasis. Finally, a group of infected individuals mount exuberant immune responses that often result in unusual pathology, such as tropical pulmonary eosinophilia. Thus the clinical manifestations of lymphatic filariasis exemplify the spectrum of host–parasite interactions that occur during helminth infections (Fig. 31.2).

Another hallmark of all helminth infections is their chronic nature, with many helminths surviving in the host for decades. For example, adult schistosomes and filariae may survive in host tissues for as long as 30 years, producing eggs and larval stages throughout this time. Similarly, *Strongyloides stercoralis*, with its ability to “autoinfect,” can maintain its lifecycle for decades. Chronic infections certainly reflect an adaptation that leads to “parasitism,” as mortality induced in the host would prevent parasite transmission if the host died before larval release or egg production could occur. In addition to the long-lived nature of the infection, helminths appear to reflect a harmonious host–parasite interface such that relatively asymptomatic carriers are available as reservoirs for ongoing transmission. Of course, failure to establish this harmonious coexistence does occur, leading to pathological conditions exemplified by cirrhosis and portal hypertension in schistosomiasis and elephantiasis associated with lymphatic filariasis.

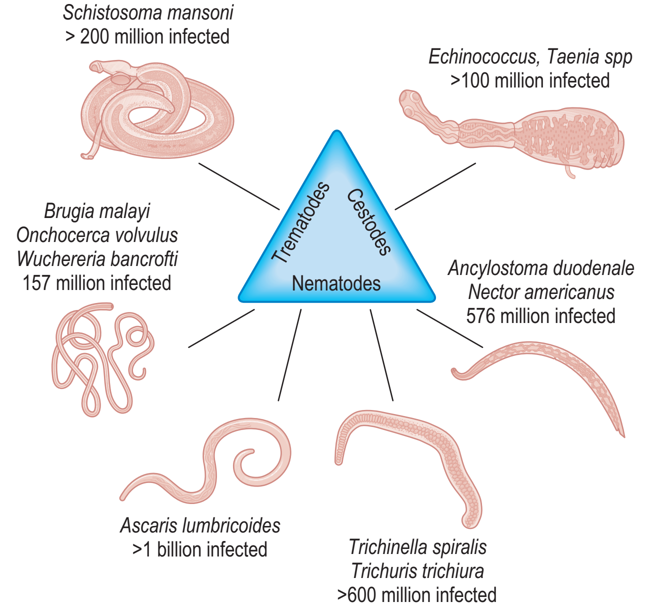

Schistosoma mansoni
> 200 million infected
Echinococcus, Taenia spp
>100 million infected
Trematodes
Cestodes
Nematodes
Brugia malayi
Onchocerca volvulus
Wuchereria bancrofti
157 million infected
Ancylostoma duodenale
Nector americanus
576 million infected
Ascaris lumbricoides
>1 billion infected
Trichinella spiralis
Trichuris trichiura
>600 million infected

FIG 31.1 Common and medically relevant helminth infections and their global prevalence.

PROTOTYPICAL HOST RESPONSES TO HELMINTHS

The canonical host immune response to all helminths is of the T-helper 2 (Th2) type and involves the production of cytokines interleukin (IL)-4, IL-5, IL-9, IL-10, and IL-13; the antibody isotypes immunoglobulin G1 (IgG1), IgG4, and IgE; and expanded populations of eosinophils, basophils, mast cells, type 2 innate lymphoid cells, and alternatively activated macrophages[2] (Chapter 16). However, it is also being increasingly recognized that while the predominant response is Th2 in nature, a large regulatory component involving both regulatory cytokines and cells are also part of this repertoire.[3] The Th2 response induced by helminth parasites is quite stereotypical, but its initiation, progression, and culmination of this response requires interaction with many different cell types, most notably *(i)* epithelial/stromal cells, *(ii)* innate lymphoid cells (ILCs), *(iii)* dendritic cells (DCs) and macrophages; *(iv)* T cells; *(v)* B cells; *(vi)* eosinophils; *(vii)* mast cells/basophils; and *(viii)* neutrophils (Chapter 2). In addition, the host–helminth interactions can lead to a variety of modulated immune responses that are mediated largely by the induction of regulatory T cells (Tregs) and alternatively activated macrophages (AAMs) (Fig. 31.3).

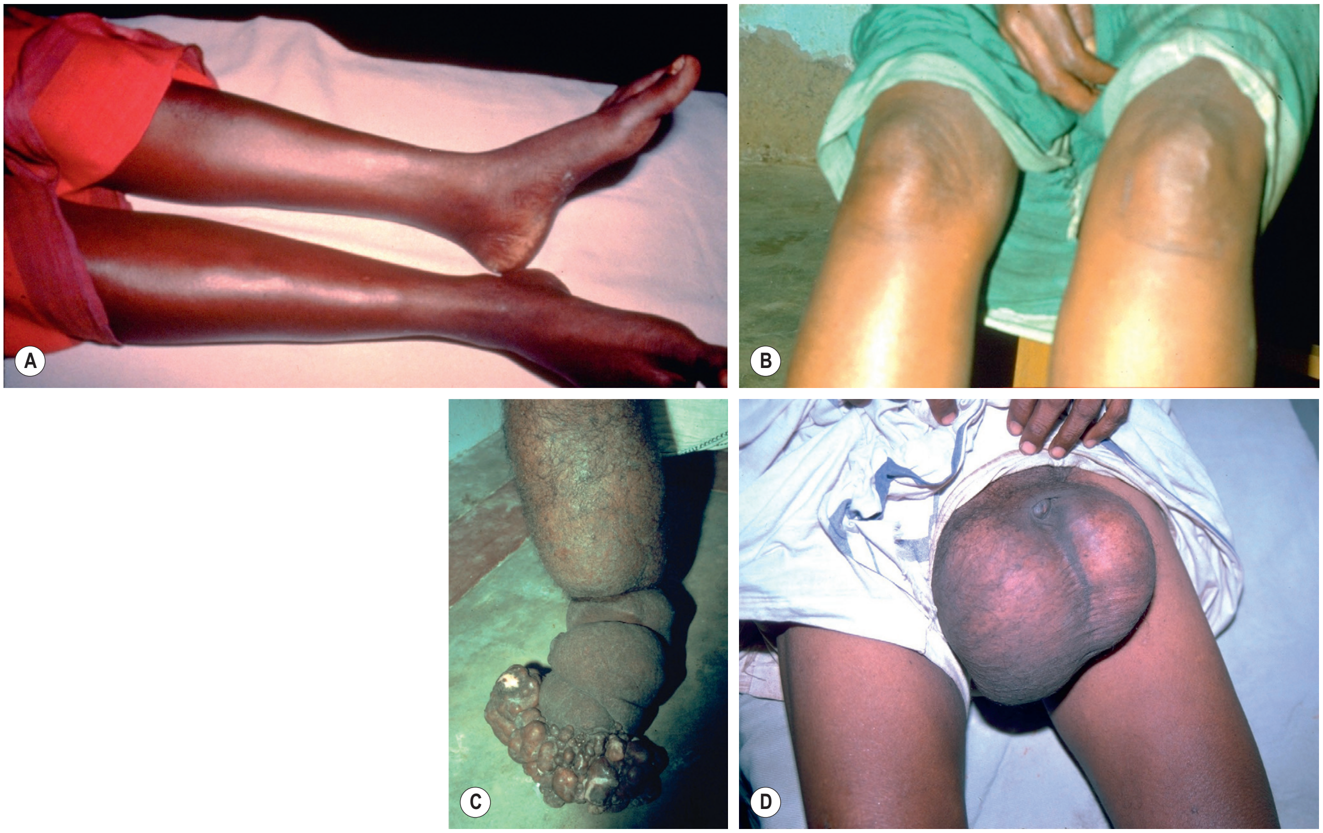


FIG 31.2 The clinical manifestations of lymphatic filariasis, including (A) mild lymphedema, (B) severe lymphedema, (C) elephantiasis, and (D) hydrocele.

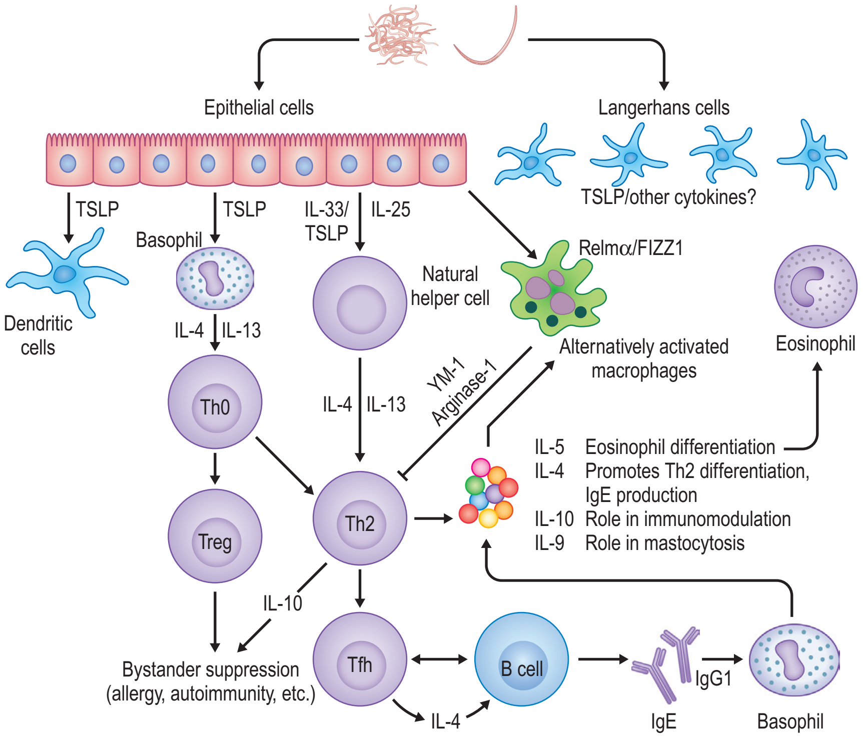

Epithelial cells
Langerhans cells
TSLP
TSLP
IL-33/ TSLP
IL-25
TSLP/other cytokines?
Basophil
Dendritic cells
Natural helper cell
Relmα/FIZZ1
Alternatively activated macrophages
Eosinophil
IL-4
IL-13
Th0
IL-4
IL-13
YM-1 Arginase-1
IL-5 Eosinophil differentiation
IL-4 Promotes Th2 differentiation, IgE production
IL-10 Role in immunomodulation
IL-9 Role in mastocytosis
Treg
Th2
IL-10
Bystander suppression (allergy, autoimmunity, etc.)
Tfh
B cell
IgG1
IgE
Basophil
IL-4

FIG 31.3 Regulation of the T-cell response in helminth infection. IL, interleukin; Th0, precursor T-helper cell; Th2, T-helper 2 cell; Treg, Regulatory T cell; TfH, T-follicular helper cell, RELM-α, Resistin-like molecule-α; Chi3l, Chitinase 3 like protein; IDO, indoleamine 2,3 dioxygenase; TSLP, thymic stromal lymphopoietin.

Helminths and Epithelial Cells

Epithelial cells are the first barrier layer exposed to or breached by most helminths, and the capacity of these cells to respond by initiating an "alarm" response has recently been recognized.[4] These epithelial cells mount a prototypical response comprising chemokines and cytokines, such as IL-1, IL-25, IL-33, and thymic stromal lymphopoietin (TSLP), as well as alarmins, such as uric acid, ATB, HMGB1, and S100 proteins. These signals program DCs to mount Th2 cell–mediated immunity and in doing so boost type 2 innate lymphoid cell (ILC2), basophil, and mast cell function. Epithelial cells produce chemokines, including CCL17 and CCL22 (acting on ILC2, basophils, Th2 cells, and Tregs), and eotaxins, such as CCL11, CCL24, and CCL26 (acting on eosinophils and Th2 cells). They also produce prostaglandin D2 (PGD2), which acts on the CRTH2 receptor to recruit ILC2, basophils, mast cells, and Th2 cells. More recently, tuft cells, a specialized secretory cell type of the intestinal epithelium, has been identified as a major player in anthelmintic immunity. In addition, epithelial cells in the intestine, for instance, are in constant contact with both beneficial and pathogenic bacteria and hence ideally located for immunological surveillance of the intestinal lumen. This recognition of signals by intestinal epithelial cells is essential to mucosal homeostasis, implicating these cells as central modulators of inflammatory responses. Finally, the production of mucus and mucus-associated bioactive molecules (Mucin5AC, trefoil factor-2, and resistin-like molecule-β [RELM-β]) are important in promoting protection against intestinal helminth infection.

Helminths and Innate Lymphoid Cells

The ILC family includes ILC1, which predominantly express IFN-γ; ILC2, which predominantly express IL-5 and IL-13; and ILC3; which predominantly express IL-22 and/or IL-17.[5] ILC2 are defined by their expression of the IL-33 receptor (IL-33R) and the transcriptional regulators, Id2, RORα, GATA-3, and Bcl11b. Unlike T cells, ILC2 rely on cytokines to drive activation rather than on cognate interactions mediated by antigen-specific receptors. ILC2 are a critical innate source of type 2 cytokines, including moderately large quantities of IL-5 and IL-13, but also of IL-4, IL-9, granulocyte macrophage–colony-stimulating factor (GM-CSF), and amphiregulin. These cytokines potently induce eosinophilia, mucus production from goblet cells, activation of AAM, muscle contractility, mastocytosis, and tissue repair.[5] They are dependent on IL-2 and IL-7 for their development and activation. In addition, the transcription factors GATA-3 and RORα have been found to be essential for the development of ILC2. Although the function of ILC2 and Th2 cells appear to be largely overlapping, the kinetic differences in the ability to secrete cytokines rapidly and in profuse amounts allows for a coordinated interaction between the two cell types. Moreover, ILC2 can directly regulate the activation of T cells through their expression of major histocompatibility complex (MHC) Class II molecules and the accessory molecules, CD80 and CD86, albeit less efficiently compared with DCs. Finally, recent reports have linked ILC2 with metabolic homeostasis, obesity, and dietary stress, providing an indirect link by which helminths might modulate host metabolic function.

Helminths and Dendritic Cells

DCs are professional antigen-presenting cells (APCs) that play an essential role in presenting antigen to T cells to initiate immune responses.[2] Although the role of DCs in inducing Th1, Th17, and Treg responses is well established, their role in inducing Th2 responses has remained relatively unclear. Nevertheless, a series of studies have shown that DCs are required for optimal Th2 responses *in vivo*. Thus *in vivo* depletion of DCs has been shown to inhibit the induction of Th2 responses to *S. mansoni* or *Heligmosomoides polygyrus.* Helminth products can prime DCs for the induction of Th2 responses by interaction with pattern recognition receptors (PRRs), such as Toll-like receptors (TLRs) and C-type lectin receptors (CLRs). This interaction, which depends on TLR and CLR signaling, can promote Th2 responses by suppressing antigen presentation, costimulation, and/or expression of Th1-promoting cytokines by directly interfering with these pathways. DCs that drive Th2 responses typically exhibit specialized markers, such as CD301b, PDL2, and CD11b, and several receptors for the Th2-related cytokines IL-4R, IL-13R, IL-25R, TSLP-R, and IL-33R. Additionally, the extracellular signal-regulated kinase (ERK) and signal transducer and activator of transcription 4 (STAT4) pathway upregulates the costimulatory molecules, CD40, OX40L, and Jagged. Activation of the major transcription factors interferon regulatory factor 4 (IRF4) and KLF4 inhibits IL-12 production and increased IL-10 secretion. In addition, DCs expressing FcεRIII can induce murine IgG1-related Th2 responses. These factors typically act individually or in concert to orchestrate Th2 responses in helminth infections. Although Th2 cell–mediated immunity requires IRF4-dependent CD301b$^+$CD11b$^+$ DCs in the mouse, Langerhans cells are the predominant inducers of Th2 cells *ex vivo* in humans. The modulation of DC function by helminth antigens appears to be generalizable and has been shown to impair their ability to respond to other infectious stimuli (*e.g., Mycobacterium tuberculosis*).

Helminths and Macrophages

Macrophages are the other important class of APCs that can serve as protective effector cells in bacterial and protozoan infections by their production of nitric oxide and other mediators. Helminth interaction with macrophages induces a population of cells preferentially expressing arginase instead of nitric oxide as a result of increased activation of *arginase-1* by IL-4 and IL-13.[6] These AAMs are characterized by their ability to upregulate arginase-1, chitinase 3-like proteins 3 and 4 (also known as Ym1 and Ym2, respectively), and RELM-α. These AAMs are known to be important in wound healing and have been postulated to play a potential role in repairing wound damage that occurs during tissue migration of helminth parasites. In fact, there appears to be two distinct populations of AAMs, one derived from blood and functioning in an immune regulatory role and the other derived from tissue-resident macrophages apparently responsible for much of the fibrosis seen in chronic helminth infections. By virtue of the expression of regulatory molecules, such as IL-10, TGF-β, and programmed cell death 1 ligand 2 (PDL2), these AAMs may have a predominantly regulatory role in helminth infections. These antiinflammatory macrophages function through arginase-1, PDL2, triggering receptor expressed on myeloid cells 2 (TREM2) and RELM-α to inhibit classic macrophage inflammation and recruitment and T-cell responses. Similarly, macrophage-derived human resistin is induced by helminth infection and promotes inflammatory responses and increased susceptibility.

Helminths and T Cells

Typically, infections with helminths induce a robust Th2 response manifested by enhanced expression of IL-4, IL-5, IL-9, IL-10, and IL-13 in response to live parasites, parasite antigens, or mitogens.[7] The central player in Th2 immunity is certainly the CD4$^+$ Th2 cell. It is clear that IL-4Rα, a component of both the IL-4 and IL-13 receptors, is at the epicenter of Th2 immunity, since IL-4 and IL-13, together or individually, are absolutely critical for resistance to most helminth parasites. Recent work has reported that the Th2 cell population is heterogeneous, containing both IL-5$^+$ and IL-5$^-$ Th2 cells that express IL-4 and IL-13. In addition, IL-4 and IL-13 production is spatially separated, with IL-13 expression being marked in tissues and IL-4 expression being pronounced in the lymph nodes within the Th2 cell compartment. Finally, induction of GATA-3 and downregulation of T-box expressed in T cells (T-bet) has been shown to be an important step in T resistin-like molecule cell differentiation to the Th2 phenotype in helminth infections. Interestingly, chronic helminth infections are associated with downmodulation of parasite antigen-specific proliferative responses as well as IFN-γ and IL-2 production but with intact IL-4 responses to parasite antigens and global downregulation of both Th1 and Th2 responses to live parasites. Finally, the receptor NLRP3 has been shown to be a key transcription factor in Th2 differentiation.

Although the role of tissue resident memory T (T_{RM}) cells is well established in viral and bacterial infections, very little is known about the role of these cells in helminth infections. However, it has been shown that tissue-resident Th2 cells can exert innate (TCR-independent and IL-33–dependent) functions upon appropriate stimuli and confer protection against helminth infection. In addition, although multifunctionality (ability to produce two or more cytokines) has not been well described in the Th2 cell compartment, helminth infections are known to be associated with an antigen-dependent enhancement of mono- and dual-functional Th2 cells and its reversal after treatment. Of interest, a stable subset of parasite induced T-bet$^+$, GATA-3$^+$, Th1/Th2 hybrid T cells has been described to develop directly from naïve precursors and to play a role in limiting pathological inflammation in animal models of helminth infection.

Recently, a new subset of T cells expressing IL-9 and IL-10, but not IL-4 (and therefore different from Th2 cells), has been described in allergic inflammation and in response to intestinal parasites.[8] These cells appear to be under the control of TGF-β and IL-4 and are dependent on STAT6, GATA-3, IRF4, and PU.1. Th9 cells have been recently shown to be associated with host protection in *Nippostrongylus brasiliensis* and *Trichuris muris* infection. Finally, Th9 cells have also been shown to be predominantly associated with lymphatic pathology in filariasis. T-follicular helper (Tfh) cells are a subset of CD4 T cells that migrate to B-cell follicles after activation and promote germinal center formation and B-cell isotype switching.[9] These cells, which form an independent lineage of CD4 T cells, have been recently identified to be the predominant IL-4 producing T cells early in helminth infection. In addition, Tfh are major producers of IL-21, a cytokine that plays a crucial role in supporting polarized Th2 responses *in vivo.*

Th17 cells, another subset of CD4 T cells, express the prototypical cytokine—IL-17. In terms of helminth infections, the role of Th17 cells has been primarily studied in animal models of *S.*

mansoni, where it has been strongly associated with infection-induced, immune-mediated pathology.[10] More recently, it has also been demonstrated in human infections, in which children with *S. hematobium*–associated pathology have higher Th17 responses compared with those who are pathology-free. Similarly, a strong association of Th17 responses with pathological responses has also been demonstrated in lymphatic filariasis. Finally, Th22 cells are yet another subset of CD4 T cells that typically secrete IL-22. To date, only a few studies have examined the role of Th22 cells in helminth infections.[11] IL-22 was shown to be induced in the intestinal mucosa after infection with *T. trichiura* or *Necator americanus* in humans, whereas the frequency of Th22 cells was shown to be higher in individuals with filarial infection compared with endemic healthy controls.

Helminths and B Cells

Helminth interactions with B cells occur both at the B-cell cytokine level and at the level of antibody production. Interactions at the cellular level primarily result in B-cell activation and cytokine production, most notably by the induction of IL-10. B cells have been shown to be important for the Th2 responses to certain helminths, with IL-2 producing B cells supporting optimal development of effector and memory Th2 cells and LTα1β2-expressing B cells supporting the recruitment of a Th2 promoting DCs.[12] Immune regulation by B cells has also been recognized in schistosome infection, where B-cell deficiency leads to enhanced Th2 cell–dependent immunopathology. However, it is at the level of antibody production that B cells play a profound role in helminth infections. Susceptibility to secondary infection is increased in the absence of B cells in infection with *Litomosoides sigmodontis, S. mansoni, T. muris,* and *Heligmosomoides polygyrus bakeri.*[13] IgG is reported as an antibody isotype that is important for protection against intestinal helminths, and IgM (typically produced in a T cell–independent manner) has been linked to timely elimination of filarial parasites. One of the most consistent findings in helminth infections, both in mice and humans, is the elevated level of IgE that is observed after exposure to helminths. Most of the IgE produced is not antigen specific, perhaps representing nonspecific potentiation of IgE-producing B cells or deregulation of a normally well-controlled immune response. Interestingly, these IgE antibodies persist many years after the infection has been treated, indicating the presence of long-lived memory B cells or plasma cells in helminth infections. IgE production both in mice and humans is absolutely dependent on IL-4 or IL-13. Other isotypes that are commonly elevated in humans with chronic helminth infection are IgG4 and IgG1, the former being most dependent on both IL-4 and IL-10. Recent studies have highlighted the role of regulatory B cells in suppression of immune responses to helminth parasites. This B-cell function involves the secretion of IL-10 and IL-35 and is similar to the regulatory activity of B cells in autoimmune diseases.

Helminths and Eosinophils

Blood and tissue eosinophilia is characteristic of helminth infection and is mediated by IL-5 (probably in concert with IL-3 and GM-CSF). Recruitment of eosinophils to the site of infection occurs very early in experimental helminth infection—as early as 24 hours after exposure. Kinetics of blood eosinophilia in humans is harder to determine but is postulated to occur as early as 2–3 weeks after infection, as demonstrated in experimental infections of volunteers.[14] Both basal eosinophil levels and tissue accumulation during helminth infection appear to be under the influence of ILC2. Apart from the rapid kinetics of recruitment, eosinophils in blood and tissue also exhibit morphological and functional changes attributable to eosinophil activation. Eosinophils possess a range of immunomodulatory factors that are released upon cell activation, including cytokines, growth factors, and chemokines. Unlike T and B cells, eosinophils can rapidly release cytokines within minutes in response to stimulation, since most of the cytokines are stored in a preformed fashion in secretory vesicles. Moreover, eosinophils can participate in the regulation of IgE and goblet cell mucus production; they also serve as effector cells in protective immune responses and as regulatory cells influencing both innate and adaptive immunity in helminth infections.

Helminths and Basophils/Mast Cells

Basophils are an important component of the immune response to helminth infections.[15] Basophils are capable of secreting a variety of mediators, including histamines, cytokines, chemokines, and lipid mediators that promote Th2 responses. Basophils in humans and mice also readily generate large quantities of IL-4 in IgE-dependent and IgE-independent manners. Basophils appear to play an important role in protective immunity to secondary infection (similar to eosinophils) with *N. brasiliensis, H. polygyrus bakeri,* and *L. sigmodontis;* they also play an active role in resistance to primary infection (through secretion of IL-4 and IL-13) with *T. muris* and *T. spiralis.* In addition, basophils have been shown to be critical APCs for driving Th2 cell differentiation in different models of helminth infection.

Mast cells may contribute to inflammatory reactions directed against invasive helminth parasites. These cells express high affinity Fcε receptors that are sensitized with parasite antigen–specific IgE and can be triggered by parasite antigens. It has been postulated that cytokines and other mediators released by sensitized mast cells contribute to *(i)* the recruitment and activation of effector eosinophils; *(ii)* increased local concentrations of antibody and complement; and *(iii)* enhanced mucus hypersecretion and increased peristalsis of the gastrointestinal (GI) tract that plays an important role in resistance to certain GI nematode infections.[16] More recently, a role for mast cells (in an IgE-independent manner) in mediating the secretion of epithelial-derived cytokines (IL-25, IL-33, and TSLP) and optimal migration of DCs was shown in *H. polygyrus bakeri* infection.

Helminths and Neutrophils

Although neutrophils are typically considered more important in bacterial and fungal infections, a number of studies have revealed that neutrophils can act in conjunction with macrophages to contain or kill helminth parasites.[17] Thus neutrophils are major components of the granulomas forming around filarial parasites and the cysts containing larvae of intestinal helminths. Neutrophils have been demonstrated to collaborate with macrophages in the immobilization and killing of *S. stercoralis* larvae in a process that is complement dependent and involving neutrophil extracellular traps (NETs). Similarly, neutrophils contribute in the early antifilarial response through oxidative burst, degranulation, and NETosis and protect against infective larvae in skin. A seminal study reported that neutrophils adopt an "N2" phenotype during experimental infection with *N. brasiliensis* in the lung and express the genes for IL-13, IL-33, RELM-α, and Ym1. These "N2" neutrophils can train macrophages to acquire a memory phenotype that protects against secondary infection. Finally, it was also shown that even during primary infection, the absence of neutrophils resulted in greater worm burdens because of lack of immunity in the lungs. Thus neutrophils appear to play an unexpected role in immunity to helminths that certainly merits further investigation.

PROTECTIVE IMMUNITY AGAINST HELMINTHS

The mechanism of protective immunity to helminths is dependent on the location of the helminth infection.[18] Clearly, T cells are central to resistance against helminths. For example, T cells are essential in mediating the expulsion of GI nematodes. Mice lacking T cells are defective in their ability to expel *T. muris,* but resistance can be reconstituted by transfer of T cells from normal mice. In addition, CD4 T cells from infected mice can transfer protective immunity to severe combined immunodeficiency (SCID) mice (lacking both B and T cells), indicating that CD4 T cells, not CD8 T cells, were important for protective immunity. Similarly, T cells were shown to be required for expulsion in *N. brasiliensis* infection. Both nude mice (lacking T cells alone) and SCID mice are susceptible to infection with Brugian parasites, whereas mice that lack either CD4 T cells or CD8 T cells are not. In schistosome-infected mice, T cells are essential in forming host-protective granulomas around the eggs deposited in the liver.

The role of cytokines in protective immunity has been extensively studied in murine models of both GI helminths and tissue-invasive helminths.[19] In general, type 2 (Th2) cytokines target epithelial cells, goblet cells, smooth muscle cells, and macrophages, which together coordinate parasite expulsion by increasing fluid and mucus production, encapsulation and barrier formation, epithelial cell turnover, smooth muscle cell contraction, and production of anthelmintic effector molecules, such as RELM-β. The cytokines involved in both responses are IL-4, IL-5, IL-9, and IL-13. Most of the studies examining resistance to intestinal helminths involve four parasitical GI nematode infections of rodent models—*T. spiralis, H. polygyrus bakeri, N. brasiliensis,* and *T. muris.* These studies show that *(i)* CD4 T cells are crucial for host protection; *(ii)* IL-4 is required for host protection and limiting host pathology; *(iii)* IL-13 can substitute for IL-4 in some but not all infections; *(iv)* IL-2 and IFN-γ inhibit protective immunity; and *(v)* IL-4 and IL-13 have multiple effects on the immune system and gut physiology leading generally to protection. Type 2 cytokines mobilize a broad range of downstream effector mechanisms. Epithelial cells in the gut, specifically tuft cells, promote goblet cell differentiation, enhancement of mucus secretion, and the production of RELM-β, an innate effector molecule with direct anthelmintic activity. Goblet cells can also secrete gel-forming mucins, which are major macromolecular components of the mucus barrier. Two of these mucins have been shown to be critical in resistance to intestinal nematode infection—Muc2 and Muc5AC. IL-4Rα activation also leads increased intestinal smooth muscle hypercontractility and accelerated epithelial turnover to promote an effector response akin to an "epithelial escalator," which, together with epithelial secretions, helps expel intestinal helminths. Mucosal mast cells release proteases that can degrade epithelial tight junctions, thereby increasing fluid flow as part of the "weep and sweep" response. AAMs in the gut can also entrap intestinal worms and cause death by compromising worm vitality.

Although the role of Th2 cytokines in immunity to GI helminth infection is well defined, their role in protective immunity to tissue-invasive helminths is not as clear.[18] In murine models of schistosomiasis, protective immune responses can be generated by vaccination with irradiated cercariae. This resistance is dependent on a Th1-mediated immune response consisting of macrophages and endothelial cells activated by IFN-γ and TNF-α, producing nitric oxide and Th1-associated antibodies—IgG2a and IgG2b. In contrast, studies in rats and epidemiological studies in humans suggest Th2-mediated effector mechanisms involving IgA and IgE antibodies as well as eosinophils are thought to be central to protective immunity. Protective immunity to filarial infections in mice is dependent primarily on Th2 responses in mice. Thus mice lacking IL-4, IL-4R, or Stat6 are all susceptible to infection with *Brugian* parasites.

In tissue-invasive helminth infections, effector mechanisms involve multiple innate immune cells, with antibodies acting as initiators of immunity by activating Fc receptor expressing cells. Basophils, by their ability to produce high levels of IL-4, act as effectors to promote helminth killing in secondary or challenge infections. For example, basophils are important in immunity to the skin invasive stages of intestinal helminths and, through IL-4 release, promote the activation of macrophages that trap larvae in an arginase-dependent manner. Although eosinophils are crucial players in producing IL-4 early in infection, they are also amplifiers of immune responses, rather than being critical mediators of primary immunity, since depletion of eosinophils does not alter the course of many helminth infections in murine models. The mechanism of protection mediated by eosinophils is thought to be by antibody-dependent, cell-mediated cytotoxicity, as observed in *S. mansoni* studies *in vitro* or through release of eosinophil granule contents. In addition, eosinophils play an important role in the protective immunity against primary infection with *B. malayi* and/or secondary infection with either *T. spiralis* or *N. brasiliensis.* The two most abundant granular proteins, major basic protein (MBP) and eosinophil peroxidase (EPO), are required for protective immunity against *S. stercoralis* and *L. sigmodontis.* Similarly, neutrophils can attack helminth larvae in response to IL-4 and IL-5, but their importance in resistance to primary helminth infections is not known.

Antibodies play a major role in mediating protection to some but not all helminth infections.[18] Antibody-mediated passive immunity has been demonstrated in animal models for *A. caninum,* Schistosoma species, Taenia species, *Ascaris suum, S. ratti, T. muris, N. brasiliensis,* and *H. polygyrus bakeri.* Passive immunity has also been shown by using IgG monoclonal antibodies (mAbs) specific for *Fasciola hepatica* and *S. mansoni;* IgM (mAbs) specific for *B. malayi;* and IgG or IgA mAbs specific for *T. spiralis.* Again using genetically manipulated mouse models, IgM has been shown to be crucial for host protection against *B. malayi* and to *S. stercoralis.* B1B cells, a subset of B cells that secrete IgM, appear to be an important component of this protective axis. Finally, antibodies have the capacity to trap tissue migrating helminth larvae and prevent tissue damage by driving an IL-4Rα–independent alternative differentiation of macrophages, in a process dependent on CD11b and FcγR1.

In terms of protection, a major mechanism appears to be the formation of multicellular, immune cell aggregates, called *granulomas,* around incoming infectious larvae or eggs.[18] In murine models of schistosomiasis and filariasis, granulomas are primarily composed of T cells (which help in the recruitment of other cell types and mediate alternative activation of macrophages), B cells (particularly the B1 subset), and macrophages and eosinophils. Although the exact mechanism by which granulomas mediate killing of the parasite remains unknown, it is clear that formation of these structures is an important host defense mechanism. One cell type that can mediate effector functions within granulomas

KEY CONCEPTS

Helminth-Induced Immune Responses

Characterized by immunoglobulin E (IgE) antibody production, tissue and peripheral blood eosinophilia, mast cell involvement, innate lymphoid cell type 2 and Th2 cell expansion, and production of type 2 cytokines.

Implicated both in pathogenesis of helminth infections and in mediating immunological protection.

In mucosal immunity to helminths, T-helper 2 (Th2) cell responses are initiated and sustained by innate populations (including tuft cells and innate lymphoid cells) through interleukin (IL)-25, IL-33, and thymic stromal lymphopoietin (TSLP).

In tissues, helminths are acted upon by the host innate effectors, including macrophages, neutrophils, eosinophils, and basophils.

Regulated by T cells and other cells producing IL-4, IL-5, IL-9, IL-10, and/or IL-13.

Characterized by the induction of regulatory T cells (Tregs) that mediate downmodulation of immune responses to helminth infections and impact bystander phenomena, such as allergy and autoimmunity.

is the AAM, which is exemplified by the targeting of the glycan chitin that is frequently expressed by helminths but not by the host. The chitinase and fizz family proteins (ChaFFs), which include chitinase and chitinase-like secreted proteins, are prime candidates for mediating host resistance. These proteins include acidic mammalian chitinase (AMCase) and the RELM family proteins and are capable of enzymatic activities that potentially damage certain helminths.

PATHOLOGY ASSOCIATED WITH IMMUNE RESPONSES IN PARASITIC HELMINTH INFECTION

Typically, pathological findings associated with each parasitic infection are different and relate to the presence of the parasites in host tissues, but there are pathological reactions that stem directly from the host response.

Immune Complexes

Immune complexes are potent mediators of localized inflammatory processes that form in many parasitic infections presumably as a result of the chronic low-dose antigen release seen in these infections. Circulating immune complexes have been identified in both experimental and human filarial and schistosomal infections. These have been shown to induce lymphatic inflammation and vasculitis in filarial infections as a result of their deposition. In addition, a common manifestation of immune complex–mediated pathology, immune complex glomerulonephritis (ICGN), has been documented by renal biopsy in patients with schistosomiasis and filarial infections. Other manifestations of immune complex–mediated damage, such as reactive arthritis and dermatitis, have also been described in patients with helminth infections.

Autoantibodies and Molecular Mimicry

Autoantibodies have been implicated as causing disease in a variety of helminth infections, including filarial infections, schistosomiasis, and hookworm infection, and are thought to reflect a polyclonal B cell expansion that often accompanies these infections.[20] Autoantibodies against nuclear material have been found in a vast majority of patients with chronic schistosomiasis, and antibodies against human calreticulin and defensin have been found in onchocerciasis.

Granulomatous Reactions

Granuloma formation is the mainstay of the protective immune response to certain helminths, but it can also lead to deleterious effects in the form of pathology. Although granulomatous reactions occur in many helminth infections (*e.g.,* toxocariasis, *Angiostrongylus* infections and lymphatic filariasis), parasitical granulomata have been best studied in *S. mansoni* infections, where granulomatous and fibrosing reactions against tissue-trapped eggs is orchestrated by CD4 T cells and the fibrosis that results from the cellular response is the principal cause of morbidity in infected individuals. The severity of the inflammatory process markedly varies both in humans and in experimental animal models, with severe pathology associated with Th1 and Th17 responses and milder pathology with Th2-dominant responses.[21] Studies in murine models of granuloma formation have demonstrated the important roles of IL-13 and TNF-α.

Fibrosis

Fibrosis is commonly associated with chronic helminth infections that result in chronic inflammation and dysregulated wound healing.[22] These infections activate macrophages and fibroblasts, resulting in the production of TGF-β, platelet-derived growth factor (PDGF), IL-1β, and other factors. Macrophages also promote inflammation by recruiting and activating monocytes and neutrophils, as well as activating CD4 T cells. In addition, fibroblasts are stimulated to synthesize matrix metalloproteinases (MMPs) and tissue inhibitors of metalloproteinases (TIMPs), leading to extracellular matrix remodeling and fibrosis. Another consequence of chronic schistosomiasis is pulmonary arterial hypertension, which has been shown to be associated with IL-4– and IL-13–mediated type 2 inflammation resulting in TGF-β–induced pulmonary vascular disease. In the same manner, IL-10 and IL-12 are known to modulate IL-13–mediated fibrosis; in the combined absence of IL-10, IL-12, and IL-13Rα, IL-13–dependent fibrosis in chronic schistosomiasis proceeds rapidly to lethal cirrhosis. Infection with *Wuchereria bancrofti* (one of the causative agents of lymphatic filariasis) is associated with similar fibrotic reactions.

Toll-Like Receptors

Immunopathology in lymphatic filariasis is associated with the presence of an endosymbiotic, *Rickettsia*-like bacteria called *Wolbachia. Wolbachia* are known to stimulate immune cells through TLR2 and TLR4 and release proinflammatory cytokines, as well as vascular endothelial growth factors (VEGFs), which might contribute to lymphatic pathology.[23] *Wolbachia*-TLR4 interaction has also been shown to be the major mechanism of corneal inflammation in onchocerciasis, and a TLR signaling molecule, IL-1 receptor associated kinase-2 (IRAK-2), regulates pathogenic Th17-cell development in *S. mansoni* infection.

Immediate Hypersensitivity Responses

Immediate hypersensitivity responses are associated with the early and/or acute phase of infections with invasive helminth parasites, such as *Ascaris,* hookworm, schistosomes, or filariae. Patients do manifest symptoms suggestive of allergic reactivity, such as wheezing or urticaria. Furthermore, in clinical syndromes associated with *Loa loa* infection (with its angioedematous Calabar swellings), with tropical pulmonary eosinophilia, and with larva currens in strongyloidiasis, IgE-mediated reactions are thought to underlie these signs and symptoms.[24] Anaphylaxis is a severe, life-threatening, generalized, or systemic hypersensitivity reaction, and is associated with IgE interaction with high-affinity IgE receptors on basophils and mast cells (Chapter 23). The risk of anaphylaxis in individuals with helminth infections can vary, depending on the parasite, tending to occur more frequently with echinococcosis or after Anisakis infection, while being extremely rare in most other helminth infections.

Wound Healing

Recent studies have shown a close association of type 2 cytokine responses with many aspects of wound healing and repair.[25,26] It has been proposed that the type 2 cytokine response has evolved to not only mediate resistance to helminth infection but also activate the wound healing apparatus to repair and reconstruct tissue, since tissue damage is intricately associated with helminth infections. Thus AAMs are intimately involved in this process as they produce MMPs, Arginase-1, insulin-like growth factor 1 (IGF1), VEGF, and TGF-β, which together promote myofibroblast activation, angiogenesis, epithelial cell turnover, and extracellular matrix deposition.

Lymphangiogenesis

The anatomical changes in the architecture of lymphatics, which range from lymphangiectasia and granulomatous responses to the development of collaterals, suggest that active lymphatic remodeling involving endothelial cell growth, migration, and proliferation is an important feature of early lymphatic filarial disease. Live filarial parasites (and their excretory/secretory products) have been shown to induce activation, proliferation, and tube formation in lymphatic endothelial cells (LECs) and their differentiation into tubelike networks.[27] This was found to be associated with significantly increased levels of MMPs and TIMPs. Recent studies have also implicated the VEGF family in lymphangiogenesis, with VEGF-C being associated with lymphedema and VEGF-A with hydrocele.[23] Finally, TLR-mediated events are considered to be the main drivers of this angiogenic/lymphangiogenic process in filarial disease.[27]

Carcinogenesis

Infection with *Opisthorchis viverrini, Clonorchis sinensis,* and *Schistosoma hematobium* are classified as group 1 biological carcinogens (*i.e.,* definitive causes of cancer). The former (liver fluke) is associated with cancer of the bile duct (cholangiocarcinoma) and cancer of the liver (hepatocarcinoma), and the latter is associated with carcinoma of the urinary bladder. The mechanisms of helminth-induced cancer include chronic inflammation, sustained cellular proliferation, modulation of the host immune system, reprogramming of glucose metabolism and redox signaling, induction of genomic instability and destabilization of tumor proteins, stimulation of angiogenesis, resistance to apoptosis, and activation of invasion and metastasis.[28]

Epileptogenesis

Neurocysticercosis, caused by the larval form of *Taenia solium,* is the most common preventable risk factor for epilepsy worldwide and accounts for nearly 30% of all epilepsies in some endemic areas. The manifestations are variable, depending on the location, number, and size of the cysts in the central nervous system as well as the degree of accompanying inflammation, provoked by cyst degeneration, calcification, and/or perilesional edema.

MECHANISMS OF EVASION AND IMMUNE REGULATION BY HELMINTH PARASITES

Helminths exert profound immunoregulatory effects on the host immune system with parasite antigen–specific immune suppression as well as more generalized levels of immune suppression. It has been shown that patients with schistosomiasis or filariasis have markedly diminished responses to parasite antigens and to some measurable attenuation in responses to bystander antigens and routine vaccinations. Thus host immunosuppression is usually antigen specific, whereas chronic infection can be associated with some spillover effects. Among the mechanisms utilized by parasites to avoid immune-mediated elimination are those of evasion—the use of sequestration, camouflage, and antigenic variation—and suppression, regulation, or blockade of immune effector pathways.

Parasite-Derived Factors

Parasite-derived products play a very important role in host immune evasion.[29,30] Parasite products, such as the schistosome-secreted proteins, alpha-1 and omega-1, promote Th2 differentiation. Alpha-1 (also known as IL-4–inducing principle of schistosome eggs [IPSE]), released by schistosome eggs, induces IL-4 release and degranulation by human and mouse basophils by cross-linking surface IgE. Omega-1 is a ribonuclease abundantly secreted by eggs, shown to condition DCs to drive Th2 polarization. Omega-1 binds to and is internalized by DCs in a mannose receptor–dependent process and then suppresses protein synthesis through degradation of messenger RNA (mRNA).

Phosphorylcholine (PC) is a small hapten-like moiety present in the excretory/secretory products of many helminths, and one particular PC-containing molecule, called ES-62, from filarial worms has been shown to have a wide variety of immunomodulatory properties. Thus ES-62 can inhibit the proliferation of CD4 T cells and conventional B cells, decrease IL-4 and IFN-γ production, promote proliferation and IL-10 production by B1B cells, modulate complement activation, and condition APCs to drive Th2 differentiation with concomitant inhibition of Th1 responses. ES-62 has also been shown to exhibit bystander antiinflammatory activity in collagen-induced arthritis, rheumatoid arthritis, chemical contact sensitivity, lupus-associated atherosclerosis, ear inflammation, chronic asthma, and airway hyperreactivity. Helminths utilize glycans within glycoproteins and glycolipids, which mimic host glycans, to regulate host immune responses. In addition, these host-like helminth glycans can directly interact with host glycan–binding proteins, such as C-type lectin receptors and galectins, to shape innate and adaptive immune responses. Similarly, helminth lipids have also been implicated in immune modulation; schistosome phosphatidylserine induces DCs to polarize IL-4–producing T cells, whereas schistosome lysophosphatidyl serine induces DCs to induce IL-10–secreting Tregs.

Helminth parasites utilize mechanisms involving cytokine mimicry and interference to establish chronic infection. Thus parasites produce cytokine- and chemokine-like molecules to interfere with the function of host innate immune products. The first helminth cytokines were found to be homologues of TGF-β expressed by *B. malayi*, and both schistosomes and filarial parasites express members of the TGF-β receptor family. Similarly, *E. granulosus* expresses a TGF-β ligand, and thus all helminth groups might have the potential to exploit TGF-β–mediated immune suppression. Various helminths, including *B. malayi*, produce homologues of macrophage migration inhibitory factors

(MIFs), which are known to activate an antiinflammatory pathway through SOCS-1, a molecule involved in cytokine signaling. *T. muris* is known to express a homologue of IFN-γ, which binds to the IFN-γ receptor *in vitro* and induces signaling. As *T. muris* is expelled by IL-4, secretion of an IFN-γ–like protein can prolong its survival.

Similarly, helminth parasites utilize chemokine- or chemokine-receptor like proteins to evade protective immunity. *Ascaris suum* is known to express a neutrophil chemoattractant with chemokine binding properties. *S. mansoni* eggs secrete a protein (*S. mansoni* chemokine-binding protein [smCKBP]) that binds the chemokines CXCL8 and CCL3 and inhibits their interaction with host chemokine receptors and their biological activity, resulting in suppression of inflammation.[31] Similarly, *B. malayi* (and all of the other filariae sequenced to date) has been shown to express galectins that can bind host immune cells in a carbohydrate-dependent manner.

Helminths secrete two major classes of protease inhibitors called *cystatins* and *serpins,* each with proposed immunomodulatory roles. Cystatins inhibit cysteine proteases (cathepsins and aspartyl endopeptidases) required for antigen processing and presentation and therefore inhibit T-cell activation. They also elicit the regulatory cytokine IL-10, leading to direct impairment of T-cell proliferation. The serpins are serine protease inhibitors, which can cause specific inhibition of the neutrophil proteinases cathepsin G and neutrophil elastase. Aspartic proteases from *Ascaris lumbricoides* have been shown to block efficient antigen processing that is dependent on proteolytic lysosomal enzymes.

Other parasite products mediate their effect by blocking effector functions, including recruitment and activation of inflammatory cells and limiting the destructive potential of activated granulocytes or macrophages in the local extracellular milieu. For example, the host chemoattractant platelet-activating factor (PAF) is inactivated by a complementary enzyme PAF hydrolase secreted by *N. brasiliensis.* Eotaxin-1, a potent eosinophil chemoattractant, is degraded by metalloproteases from hookworms. *A. caninum* secretes a protein called *neutrophil inhibitory factor,* which binds the integrins CD11b/CD18 and blocks adhesion of activated neutrophils to vascular endothelial cells and also the release of hydrogen peroxide (H_2O_2) from activated neutrophils. *N. americanus* ES products also bind to host NK cells and augment the secretion of IFN-γ, which might cross-regulate deleterious Th2 responses. Other modulators, such as prostaglandins, and other arachidonic acid family members, such as PGE2 and PGD2, are known to inhibit IL-12 production by DCs. Finally, helminths susceptible to oxidant-mediated killing express both secreted and membrane-associated enzymes, such as superoxide dismutase, glutathione-S-transferase, and glutathione peroxidase, molecules that are thought to play a significant role in assisting parasite survival in inflamed tissues. Recently, a family of helminth defense molecules secreted by parasitic helminths has been shown to exhibit biochemical and functional characteristics similar to human antimicrobial peptides. These molecules can modulate innate cell activation by classic TLR ligands, such as lipopolysaccharide.

It has been also been reported that parasitical helminths can produce exosomes and other secretory vesicles that facilitate the transfer of intracellular cargo. Exosomes derived from helminths have been shown to possess immunomodulatory capacity with the ability to modulate immune responses by altering the function of ILC2. In addition, secreted vesicles from *B. malayi, O. volvulus,* and *S. mansoni* have been shown to contain microRNAs that can then act as messengers of communication between the parasite and the host cell. These exosomes (containing microRNA) can then enter the host cell and modulate host gene expression.[32]

Host-Related Factors

Regulatory T and B Cells

Evidence for the involvement of Tregs in helminth-mediated downmodulation of the immune response has been accumulating in recent years[33] (Chapter 18). IL-10 and TGF-β, both factors associated with Tregs, are elicited in response to helminth infections, and *in vitro* neutralization of IL-10 and TGF-β at least partially restores T-cell proliferation and cytokine production in lymphatic filariasis. Similar reversals of immunosuppression are observed in onchocerciasis and schistosomiasis, with IL-10 producing natural Tregs (nTregs) from egg-induced granulomas (in *S. mansoni* infection) being important for host survival. More recently, Tregs from patients with filarial infections have also been shown to express high levels of other suppressive molecules, such as CCL4, IL-29, LAG-3, and Foxo3. Moreover, filarial infection is associated with an expansion of T cells expressing the IL-10 superfamily cytokine members (IL-19 and IL-24), and inhibition of these cytokines results in increased Th1 and Th2 responses. Tregs play a vital role in limiting host pathology by downregulating harmful Th1/Th17 responses in filarial infection and schistosomiasis. In addition, low level Treg activity is essential for type 2 effector immunity to expel certain intestinal helminths, and transdifferentiation of Th17 cells into Tregs during helminth infection is important for resolution of inflammation.

A number of studies have recently reported that B cells might have an active regulatory role in helminth infections. For example, purified B cells from mice infected with *B. malayi* produce IL-10 in response to filarial antigens. Similarly, IL-10 producing $CD1d^{hi}$ B cells are induced in both mice and humans and are suppressed after anthelmintic treatment. This induction occurs, in part, through a mechanism involving the ICOS-B7RP-1 pathway.

Hyporesponsive T Cells

Effector T cell responses can be turned off or modulated through a variety of mechanisms, including through cytotoxic T lymphocyte antigen-4 (CTLA-4) and programmed death-1 (PD-1). Interestingly, increased expression of CTLA-4 and PD-1 has been demonstrated in filarial infections, and blocking of CTLA-4 can partially restore a degree of immunological responsiveness in cells from infected individuals. Moreover, T cells have decreased induction of *T-bet,* the Th1 master regulatory gene indicating a failure at the transcriptional level to differentiate into Th1 cells. Finally, T cells from individuals with filarial infection exhibit classic signs of anergy, including diminished T-cell proliferation to parasite antigens, lack of IL-2 production, and increased expression of E3 ubiquitin ligases. Similarly, anergic T cells are found in both humans and mice with *F. hepatica* infection and schistosomiasis; in the latter case, these T cells express high levels of the anergy molecule GRAIL (gene related to anergy in lymphocytes).

Modulation of Apc Function

DCs are the first APCs usually to encounter parasites, and helminth modulation of DC function has been well characterized.[34] Filarial parasites induce downregulation of MHC class I and class II molecules, as well as cytokines and other genes involved in antigen presentation, thereby rendering DCs suboptimal in their ability to activate CD4 T cells. Schistosomes have similar effects on DCs, with subsequent Th2 polarization and inhibited responses to Th1-inducing TLR ligands. In addition, schistosomes modulate the activation of Nlrp3 (NLR family, pyrin domain containing 3) inflammasome and thus IL-β production. Excretory/secretory antigens produced by helminths can inhibit DC synthesis of proinflammatory cytokines, chemokines, and costimulatory molecules and promote DC production of the regulatory cytokines IL-10 and TGF-β. Helminth infection has also been shown to induce *in vivo* differentiation of a $CD103^-$ $CD11c^{lo}$ population of regulatory DCs, which are inefficient in priming effector T cells and instead favor the generation of Tregs. AAMs are able to markedly suppress target cell proliferation, as well as mediate repair of tissue that has been damaged by parasites. In addition, human filarial infection is associated with the expansion of the nonclassic monocyte subset, as well as an immunoregulatory monocyte subset. Helminth antigens can modulate MHC class II and CD80/86 expression on "antigen-presenting" basophils to induce the development of Th2 cells. Finally, a heterogeneous population of immature myeloid cells that share the common property of suppressing immune responses are termed *myeloid-derived suppressor cells* (MDSCs). Although these MDSCs have been well characterized in cancer immunology, their role in helminth infections is still being explored.

Apoptosis

Another mechanism of immune evasion is the ability of some helminths to induce host cell apoptosis (Chapter 13). Apoptosis has been described as a host regulatory mechanism in various helminth infections, including schistosomiasis, lymphatic filariasis, and onchocerciasis.

HELMINTHS AND THE MICROBIOTA

Recent work has highlighted the importance of the microbiota in influencing host immunological and metabolic functions (Chapter 14). Helminths secrete a variety of products that can directly influence the composition and function of the microbiota, whereas changes in microbiota can have an impact on susceptibility to helminth infection, indicating that helminth–microbiota cross-talk can regulate a variety of host processes.[35] Recently, it was reported that infection with helminths (*H. polygyrus* or *T. muris*) results in alteration of the composition of the microbiota, with expansion of Lactobacillacea and Enterobacteriacea in the gut, and that the removal of helminths resulted in restoration of the original composition of the microbiota. The consensus in humans is yet to emerge, with some studies showing an effect of helminth infection on microbial diversity and others failing to reveal any differences. Conversely, it has also been demonstrated that introducing higher level of certain commensal microbes enhances susceptibility, whereas removal of microbiota by antibiotics decreases susceptibility to helminth infections.

REGULATION OF ALLERGY, AUTOIMMUNITY, AND METABOLIC DISEASES IN HELMINTH INFECTION

The hygiene hypothesis postulates that the stimulation of the immune system by microbes or microbial products protects from the development of inflammatory and atopic disorders[36] (Chapter 1). Human studies have demonstrated that people living in areas endemic for helminth infections have a decreased reactivity to skin tests for allergens and milder forms of asthma.[37] Experimental animal models have revealed the protective effect of helminth infections against atopy and asthma. Several mechanisms have been proposed for the helminth-induced protection, the chief of which are the induction of Treg activity, regulatory B-cell activity, and immunosuppressive cytokines including IL-10 and TGF-β. Similarly, exposure to helminth parasites has been shown to prevent the onset of Th1-mediated diseases, such as multiple sclerosis (MS), diabetes mellitus, and Crohn disease in experimental animal models.[37] Finally, recent studies in mice have shown that type 2 immunity, induced by helminth infection, can maintain adipose tissue homeostasis and promote adipose tissue beiging, protecting against obesity and metabolic dysfunction; that the immunomodulatory glycan LNFPIII, which is secreted by helminths, can alleviate hepatosteatosis and insulin resistance; and that there exists an inverse association between the presence of helminth infections and the prevalence of type 2 diabetes.

HELMINTH THERAPY FOR INFLAMMATORY DISEASES

To date, two species of helminths have been tested as clinical treatment for therapy of inflammatory diseases: *T. suis* ova and infection with *N. americanus*. Currently, 28 clinical trials of helminth therapy in 10 autoimmune diseases and allergic or related conditions have been planned, started, or completed.[38] These include Crohn disease, ulcerative colitis, MS, celiac disease, autism, plaque psoriasis, peanut and tree nut allergy, asthma, rhinoconjunctivitis, and rheumatoid arthritis. Although well tolerated, recent results suggested that this therapy did not achieve improvement in disease activity or remission rates in Crohn disease. Studies using *N. americanus* include fairly small trials with patients with Crohn disease or individuals with celiac disease (gluten allergy). Additionally, a few studies have been performed examining the effect of helminth therapy on asthma/allergy. The one minor success of helminth therapy in humans is in the treatment of MS, wherein small trials have successfully demonstrated lower relapses and lower magnetic resonance imaging (MRI) activity, prompting larger phase I or II trial involving either *T. suis* or *N. americanus*.

ON THE HORIZON

- Identification and synthesis of helminth products that can be useful as immune therapy in a variety of inflammatory disorders
- Deciphering the three way cross-talk among host immunity, helminths, and the microbiota
- Elucidation of the detailed mechanisms by which helminths manipulate immune responses to bystander antigens
- Development of clues to production of novel vaccine candidates to protect against not only helminth infection but also helminth-induced morbidity
- Combined approaches involving genomics, transcriptomics, proteomics, and metabolomics for assessment of host–helminth interactions

VACCINES AGAINST HELMINTH PARASITES

Vaccines against helminth infections are a necessary tool for their elimination and eradication for several different reasons.[39] Currently, a total of five human anthelmintic vaccines have advanced from discovery through manufacture and are now in phase I or II clinical testing.[40] These include three Schistosome vaccine candidates: *(i) S. hematobium* Sh28GST (phase II); *(ii) S. mansoni* Sm-TSP-2 (phase I); *(iii) S. mansoni* Sm-14 (phase I); and two hookworm vaccine candidates: *(i)* Na-GST-1 (phase I); *(ii)* Na-APR-1 (phase I). In addition, the two *O. volvulus* vaccine candidates—Ov-103 and Ov-RAL2—and the *S. mansoni* candidate—Sm-p80—are in preclinical testing. With rapid advances in the parasite genomics and proteomics, as well the newer, better vaccine delivery systems offering more efficient and quicker assessment, the prospects for newer anthelmintic vaccines are excellent, although the potential lack of commercial markets imposes a significant impediment to their development.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Keusch GT. Immune responses in parasitic diseases. Part A: general concepts. Rev Infect Dis 1982;4:751–5.
2. Allen JE, Maizels RM. Diversity and dialogue in immunity to helminths. Nat Rev Immunol 2011;11:375–88.
3. Mishra PK, Palma M, Bleich D, et al. Systemic impact of intestinal helminth infections. Mucosal Immunol 2014;7:753–62.
4. Hammad H, Lambrecht BN. Barrier Epithelial cells and the control of type 2 immunity. Immunity 2015;43:29–40.
5. McKenzie AN, Spits H, Eberl G. Innate lymphoid cells in inflammation and immunity. Immunity 2014;41:366–74.
6. Kreider T, Anthony RM, Urban JF Jr, et al. Alternatively activated macrophages in helminth infections. Curr Opin Immunol 2007;19:448–53.
7. Grencis RK. Immunity to helminths: resistance, regulation, and susceptibility to gastrointestinal nematodes. Annu Rev Immunol 2015;33:201–25.
8. Noelle RJ, Nowak EC. Cellular sources and immune functions of interleukin-9. Nat Rev Immunol 2010;10:683–7.
9. Crotty S. Follicular helper CD4 T cells (TFH). Annu Rev Immunol 2011;29:621–63.
10. Larkin BM, Smith PM, Ponichtera HE, et al. Induction and regulation of pathogenic Th17 cell responses in schistosomiasis. Semin Immunopathol 2012;34:873–88.
11. Bouchery T, Kyle R, Ronchese F, et al. The differentiation of CD4(+) T-helper Cell subsets in the context of helminth parasite infection. Front Immunol 2014;5:487.
12. Shen P, Fillatreau S. Antibody-independent functions of B cells: a focus on cytokines. Nat Rev Immunol 2015;15:441–51.
13. Harris N, Gause WC. To B or not to B: B cells and the Th2-type immune response to helminths. Trends Immunol 2011;32:80–8.
14. Klion AD, Nutman TB. The role of eosinophils in host defense against helminth parasites. J Allergy Clin Immunol 2004;113:30–7.
15. Karasuyama H, Mukai K, Obata K, et al. Nonredundant roles of basophils in immunity. Annu Rev Immunol 2011;29:45–69.
16. Galli SJ, Tsai M. Mast cells in allergy and infection: versatile effector and regulatory cells in innate and adaptive immunity. Eur J Immunol 2010;40:1843–51.
17. Allen JE, Sutherland TE, Ruckerl D. IL-17 and neutrophils: unexpected players in the type 2 immune response. Curr Opin Immunol 2015;34:99–106.
18. Anthony RM, Rutitzky LI, Urban JF Jr, et al. Protective immune mechanisms in helminth infection. Nat Rev Immunol 2007c;7:975–87.
19. Finkelman FD, Shea-Donohue T, Goldhill J, et al. Cytokine regulation of host defense against parasitic gastrointestinal nematodes: lessons from studies with rodent models. Annu Rev Immunol 1997;15:505–33.
20. Zandman-Goddard G, Shoenfeld Y. Parasitic infection and autoimmunity. Lupus 2009;18:1144–8.
21. Wynn TA, Thompson RW, Cheever AW, et al. Immunopathogenesis of schistosomiasis. Immunol Rev 2004;201:156–67.
22. Allen JE, Wynn TA. Evolution of Th2 immunity: a rapid repair response to tissue destructive pathogens. PLoS Pathog 2011;7:e1002003.
23. Pfarr KM, Debrah AY, Specht S, et al. Filariasis and lymphoedema. Parasite Immunol 2009;31:664–72.
24. Pritchard DI. The pro-allergic influences of helminth parasites. Mem Inst Oswaldo Cruz 1997;92:15–18.
25. Gause WC, Wynn TA, Allen JE. Type 2 immunity and wound healing: evolutionary refinement of adaptive immunity by helminths. Nat Rev Immunol 2013;13:607–14.
26. Wynn TA, Vannella KM. Macrophages in tissue repair, regeneration, and fibrosis. Immunity 2016;44:450–62.
27. Babu S, Nutman TB. Immunopathogenesis of lymphatic filarial disease. Semin Immunopathol 2012;34:847–61.
28. Fried B, Reddy A, Mayer D. Helminths in human carcinogenesis. Cancer Lett 2011;305:239–49.
29. Harnett W, Harnett MM. Helminth-derived immunomodulators: can understanding the worm produce the pill? Nat Rev Immunol 2010;10:278–84.
30. Hewitson JP, Grainger JR, Maizels RM. Helminth immunoregulation: the role of parasite secreted proteins in modulating host immunity. Mol Biochem Parasitol 2009;167:1–11.
31. Smith P, Fallon RE, Mangan NE, et al. Schistosoma mansoni secretes a chemokine binding protein with antiinflammatory activity. J Exp Med 2005;202:1319–25.
32. Coakley G, Buck AH, Maizels RM. Host parasite communications-Messages from helminths for the immune system. Mol Biochem Parasitol 2016;208:33–40.
33. Belkaid Y, Tarbell K. Regulatory T cells in the control of host-microorganism interactions. Annu Rev Immunol 2009;27:551–89.
34. Everts B, Smits HH, Hokke CH, et al. Helminths and dendritic cells: sensing and regulating via pattern recognition receptors, Th2 and Treg responses. Eur J Immunol 2010;40:1525–37.
35. Gause WC, Maizels RM. Macrobiota—helminths as active participants and partners of the microbiota in host intestinal homeostasis. Curr Opin Microbiol 2016;32:14–18.
36. Yazdanbakhsh M, Kremsner PG, van Ree R. Allergy, parasites, and the hygiene hypothesis. Science 2002;296:490–4.
37. Wilson MS, Maizels RM. Regulation of allergy and autoimmunity in helminth infection. Clin Rev Allergy Immunol 2004;26:35–50.
38. Fleming JO, Weinstock JV. Clinical trials of helminth therapy in autoimmune diseases: rationale and findings. Parasite Immunol 2015;37:277–92.
39. Bethony JM, Cole RN, Guo X, et al. Vaccines to combat the neglected tropical diseases. Immunol Rev 2011;239:237–70.
40. Hotez PJ, Strych U, Lustigman S, et al. Human anthelminthic vaccines: Rationale and challenges. Vaccine 2016;34:3549–55.

MULTIPLE-CHOICE QUESTIONS

1. The immunity to helminths typically involves a T-helper 2 (Th2) cell response with what other component of the T-cell response?
   A. Th1 response
   B. Th17 response
   C. Regulatory T-cell response
   D. Th22 response

2. The classic antibody isotypes induced by helminths typically include:
   A. Immunoglobulin E (IgE) and IgG4
   B. IgE and IgG3
   C. IgE and IgD
   D. IgG1 and IgG2

3. These cells are a novel cell type recently shown to play a pivotal role in the initiation of immunity to helminth parasites:
   A. Natural killer (NK) cells
   B. Type 1 innate lymphoid cells
   C. Type 2 innate lymphoid cells
   D. Type 3 innate lymphoid cells

32

Approach to the Evaluation of the Patient With Suspected Immunodeficiency

Javier Chinen, Mary E. Paul, William T. Shearer

KEY CONCEPTS

Secondary Immunodeficiencies

Immunodeficiency is often secondary or transient, caused by nonimmune factors, including:

- Previous use of high-dose steroids, or other immunosuppressive medications
- Previous use of monoclonal antibodies (mAbs), such as rituximab (anti-CD20)
- Immunoglobulin losses via the gastrointestinal or urinary tract
- Severe illness requiring critical care
- Malnutrition
- Human immunodeficiency virus (HIV) infection

Clinical immunologists are often consulted to evaluate patients for suspected immune defects, usually because such patients have an unusual frequency or severity of infectious illnesses. Indeed, immunodeficiency presents with increased susceptibility to infection but may also manifest with conditions that reflect dysregulation of the immune response, such as allergies, autoimmunity, or lymphoproliferation. Prompt diagnosis is essential to reduce the risk of organ damage caused by preventable severe infections. Primary immunodeficiencies (PIDs) are congenital diseases that might affect any aspect of the immune response and are often diagnosed in childhood. Examples of PIDs include severe combined immune deficiencies (SCIDs), complete DiGeorge syndrome, and chronic granulomatous disease (CGD). In contrast to PIDs, secondary immunodeficiencies (Chapters 38, 39) present at any age, as a result of a wide variety of factors that affect the immune function, such as environmental factors, metabolic disease, anatomical abnormalities, or infectious agents. The most known and significant secondary immunodeficiency is caused by human immunodeficiency virus (HIV). The assessment of a patient for PID should include history and physical examination to direct immunology laboratory testing to confirm a diagnosis.

EPIDEMIOLOGY—PRIMARY IMMUNODEFICIENCIES ARE NOT UNCOMMON

Estimates of the incidence of PIDs or congenital immunodeficiencies varies from selective immunoglobulin A (IgA) deficiency, a relatively common condition, (1/223–1/1000 people)[1] to the less common SCID. Recent analysis from 11 state programs established for universal newborn screening for T-cell deficiencies in the United States reported an incidence of SCID of 1/58 000 live births, comparable with childhood leukemia.[2] A household-based telephone survey suggested that 1 in 1200 persons in the United States has been diagnosed with a PID.[3] Although significant progress has been made to stop the acquired immunodeficiency syndrome (AIDS) epidemic, HIV infection continues to be the most prevalent cause of immunodeficiency worldwide, with an estimated 36.9 million people living with HIV.[4]

PRIMARY VERSUS SECONDARY IMMUNODEFICIENCY

Because of their common occurrence, acquired and nonimmunological causes for recurrent infections should be first considered in the differential diagnosis of the patient with a suspected immune disorder. Acquired conditions that might increase the frequency of infections include allergic inflammation, HIV infection, and the use of immunosuppressive drugs. Examples of nonimmunological conditions include those that disrupt the usual mucosal clearance mechanisms, such as posterior urethral valves or urethral stenosis in a patient with recurrent urinary tract infection, or cystic fibrosis in individuals who have recurrent sinusitis or pneumonia and/or diarrhea. Disruption of natural barriers can lead to increased risk of infection, for example, in patients with skin lesions caused by eczema or burns or in individuals with cerebrospinal fluid (CSF) leaks following bony injury to the cranium. Patients presenting with low Ig levels might have loss of antibodies as a result of a protein-losing enteropathy, nephropathy, or massive protein loss through skin, such as in severe eczema or burns.[5] Secondary immunodeficiency can also result from other conditions affecting cell metabolism (*e.g.*, malnutrition, diabetes mellitus, and sickle cell anemia) or could be secondary to predictable or idiopathic adverse effects of drugs. Optimal management of these conditions often results in improved immunity.

EVALUATING PATIENTS FOR IMMUNODEFICIENCY

The evaluation of patients for immunodeficiency is based on a careful assessment of patient history and physical examination, with very limited initial laboratory testing. With this information, the clinician can often tell patients (or parents) whether their (or their child's) immune system is significantly compromised. The immunologist must take on a role of counselor and advisor of patients and explain the many factors that may result in increased frequency of infections (Fig. 32.1). The limitations of available clinically validated testing need to be considered, as such tests may not be sensitive or specific to identify uncommon

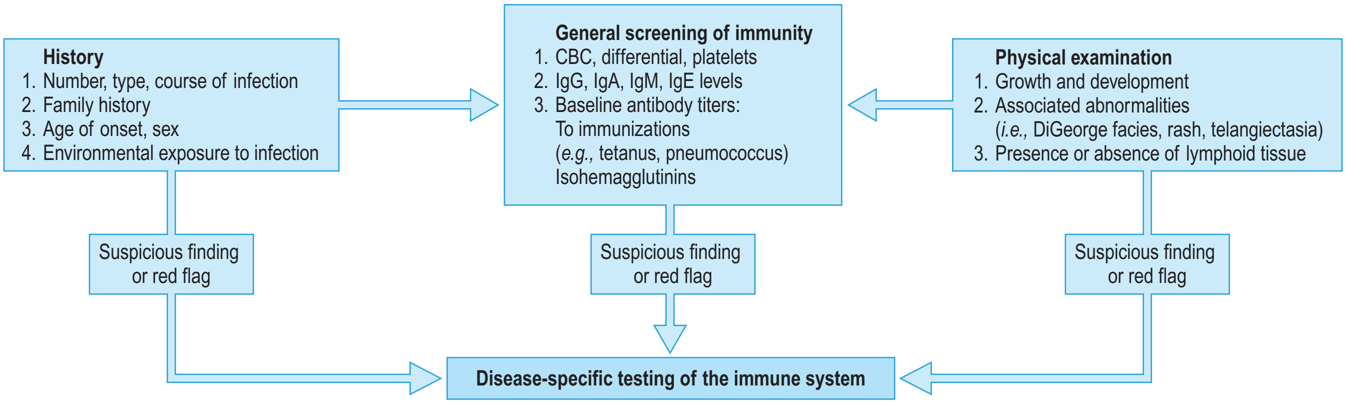

History
1. Number, type, course of infection
2. Family history
3. Age of onset, sex
4. Environmental exposure to infection
General screening of immunity
1. CBC, differential, platelets
2. IgG, IgA, IgM, IgE levels
3. Baseline antibody titers:
To immunizations
(e.g., tetanus, pneumococcus)
Isohemagglutinins
Physical examination
1. Growth and development
2. Associated abnormalities
(i.e., DiGeorge facies, rash, telangiectasia)
3. Presence or absence of lymphoid tissue
Suspicious finding or red flag
Suspicious finding or red flag
Suspicious finding or red flag
Disease-specific testing of the immune system

FIG 32.1 **Evaluation of Immunity in a Patient for Immunodeficiency Starts With a Careful History and Physical Examination.** Clues in the history for further evaluation include an excess of respiratory tract infections, of unusual severity; life-threatening infections; infections with unusual organisms; a family history of immunodeficiency. The outlined laboratory tests provide an adequate screen of immunity in a patient with no specific findings.

immune defects, such as impairments in phagocyte function other than oxidative burst deficiency. The medical history and initial laboratory testing often provide clues that suggest a specific immune disorder, and the examination of specific component of the immune response or a diagnostic test for specific immunodeficiencies may be indicated. For example, an increased frequency of infections affecting only the respiratory tract and caused by encapsulated bacteria direct the exploration to defects in humoral immunity and complement; in contrast, a history of *Aspergillus* pneumonia would suggest neutropenia and CGD. According to the severity of the illness, clinical immunologists may recommend an initial exploration of the major components of the immune system: lymphocyte subset distribution, antibody responses, T-cell function, phagocyte oxidative burst, and the complement system.

EXPLORING THE MEDICAL HISTORY

Age and Environment

The differential diagnosis of immunodeficiency varies with the age of onset of symptoms. Pediatric patients are more likely to present a PID than a secondary immunodeficiency. Infants from birth to 3 months of age have maternal Igs acquired through the placenta, unless they were born prematurely. Therefore deficiencies in the immune system at this age presenting with frequent infections most probably result from severe deficiencies in other immune components, such as neutrophils, complement components, or T cells. Older patients might present with increased risk of infections secondary to comorbid conditions, such as allergic inflammation or diabetes mellitus, or to a normal decline of immune responses, a process known as *immunosenescence* (Chapter 38).[6] Environmental conditions can influence the risk of infection. Infants frequently exposed to other infants with infections, such as in the setting of a daycare facility, have more infectious illnesses compared with those who are not exposed. By inducing an inflammatory response of the respiratory mucosa, passive cigarette smoke inhalation also predisposes to infections, including otitis media, pneumonia, and bronchitis.[7] The hygiene practices of the patient, caregivers, and family members have an impact on the frequency of infections, such as impetigo and furunculosis. For patients with allergies, exposure to indoor allergens, such as dust mites and molds, often worsens mucosal congestion and increases the risk of sinusitis and otitis media.

Immunization and Previous Infections

The immunization history provides significant clues, since the efficacy of vaccines depends on intact immunity. An incomplete schedule of immunizations easily explains the incidence of preventable diseases. A history of an adverse reaction to a live viral vaccine is suspicious for immunodeficiency, as infants with T-cell defects, B-cell defects, and combined T- and B-cell defects are susceptible to potentially fatal or severe infections from live attenuated vaccines. These infections include measles and chicken pox pneumonitis, rotavirus vaccine-induced diarrhea,[8] and lymphadenitis caused by the Bacille Calmette-Guérin (BCG) vaccine.[9] Historical information regarding the frequency, type, and severity of illnesses and infections should be sought. Individuals with immunodeficiency may have infections with unusually prolonged courses or unusual severity or that may present as unexpected complications (Fig. 32.2). Recurrent infections that involve multiple sites are more suspicious for immune deficiency than those involving a single site. Also suggestive of immune compromise are severe and invasive infections, such as recurrent pneumonia, meningitis, sepsis, septic arthritis, osteomyelitis, or abscess and infections with organisms of low pathogenicity in normal individuals, such as *Candida albicans* or *Pneumocystis jiroveci* (Table 32.1) Patients with antibody deficiency disorders tend to present with infections caused by extracellular pyogenic organisms, such as *Haemophilus* spp., *Pneumococcus* spp., and *Streptococcus* spp. In contrast, patients with defects in cell-mediated immunity are more likely to also present with recurrent infections with viruses, fungi, protozoa, and mycobacteria. Furthermore, infections by catalase-positive bacteria, such as *Serratia marcescens,* may indicate a possible neutrophil oxidative burst defect. Recurrent neisserial infections may be found in individuals deficient in the terminal complement components. A relatively normal incidence of infections followed by a sudden occurrence of repeated infections in an adult or adolescent suggests a secondary immunodeficiency, including HIV infection.

TABLE 32.1 Clinical Clues of Significance for the Diagnosis of Immunodeficiencies

	T-Cell Function Defect	Antibody Defect	Granulocyte Defects	Complement Defect	IFN-γ/IL-12 Defect
Recurrent or severe bacterial infections	X	X	X (catalase-positive)	X (encapsulated bacteria)	
Systemic mycobacterial infections	X				X
Recurrent or severe viral infections	X	X	X		X
Invasive fungal infections	X		X		
Opportunistic infections	X		X		X
Failure to thrive	X	X	X		
Autoimmunity	X	X	X	X	
Lymphoma	X	X (CVID)			

CVID, common variable immunodeficiency; IFN, interferon; IL, interleukin.

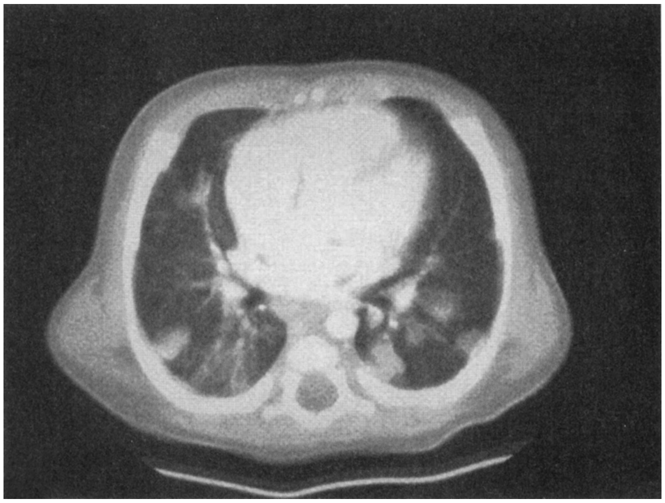


FIG 32.2 **Computed Tomography (CT) Scan in an Infant With Chronic Granulomatous Disease (CGD).** Multiple nodular opacities are seen throughout both lung fields as a result of fungal pneumonia in an infant with CGD. (From Seeborg FO, et al. A 5-week-old HIV-1-exposed girl with failure to thrive and diffuse nodular pulmonary infiltrates. J Allergy Clin Immunol 2004; 113: 629, with permission from Elsevier.)

TABLE 32.2 Nonimmunological Clinical Findings Present in Immunodeficiency Syndromes

Nonimmunological Clinical Finding	Immunodeficiency
Small platelets, thrombocytopenia, eczema	Wiskott-Aldrich syndrome (WAS)
Conical teeth, ectodermal dysplasia	Nuclear factor κB (NF-κB) essential modulator (NEMO) defect
Delayed shedding of primary teeth, frequent fractures, hyperextensibility	Autosomal dominant hyper-IgE syndrome (AD-HIES)
Cerebellar ataxia, telangiectasia	Ataxia–telangiectasia (AT)
Hypoparathyroidism, conotruncal heart defect, velopalatal insufficiency	DiGeorge syndrome
Short limbs	Cartilage-hair hypoplasia
Microcephaly	DNAse IV–deficient severe combined immunodeficiency (SCID), Nijmegen syndrome
Silvery hair, albinism	Pigment dilution disorders
Pectum carinatum, skeletal dysostosis, pancreatic insufficiency	Shwachman-Bodian-Diamond syndrome (SBDS)

Comorbid Conditions

Apart from frequent infections, other important aspects of the history and physical examination may suggest congenital syndromes associated with immune defects (Table 32.2). Neutrophil adhesion defects lead to delayed (beyond 2 weeks of age) umbilical cord separation because of omphalitis and poor wound healing. Patients with DiGeorge syndrome are usually diagnosed in the neonatal period with hypocalcemic seizures associated with hypoparathyroidism, velopalatal insufficiency, or cardiovascular malformations, rather than with recurrent infections. Infants with Wiskott-Aldrich syndrome (WAS)—immunodeficiency, thrombocytopenia, and eczema—present with petechiae and bruises similarly in the neonatal period. Obtaining a history of atopic disease is helpful. Atopic dermatitis has been associated with high risk of infection, including nondermatological infections.[10] For some patients with allergic rhinitis, management of allergies reduces the frequency of upper respiratory tract infection. In addition, it is important to inquire for a history of recurrent wheezing. At times, an initial history of recurrent pneumonia is, in fact, secondary to reactive airways disease or asthma. Other significant conditions include renal disease causing proteinuria or enteropathies that result in protein losses and secondary hypogammaglobulinemia.

Use of Medications

Use of particular drugs might also cause immunodeficiency, which could be *predictable,* such as the use of rituximab (anti-CD20 antibody) resulting in B-cell depletion and potential antibody deficiency, or *idiopathic,* such as hypogammaglobulinemia that might develop with the use of anticonvulsants.[11]

Family and Social Histories

Family history is essential in the evaluation of suspected immunodeficiency. A history of early infant deaths and possible consanguinity should be sought. A clear pattern of inheritance may be found to define an X-linked, autosomal dominant, or autosomal recessive genetic syndrome. Many of the most common PIDs have X-linked inheritance patterns. Family members of patients with immunodeficiencies might also have a history of autoimmune disease or of connective tissue disease. Familial cases of selective IgA deficiency and common variable immunodeficiency (CVID) have been reported, and a susceptibility trait can sometimes be traced back to many generations. A social history should be obtained for risk factors associated with increased risk of acquiring HIV infection. Socioeconomic factors often determine malnutrition, known to be of a significant impact on immune function.[12]

KEY CONCEPTS

Genetic Immunodeficiencies

- Diagnosis of genetic immunodeficiencies require high clinical suspicion and confirmation with molecular studies.
- In a minority of cases, genetic immunodeficiencies might not present with all clinical elements of classic descriptions. For example, patients with Bruton agammaglobulinemia might have residual B-cell and immunoglobulin production.

PHYSICAL EXAMINATION FINDINGS

The physical examination might provide findings that indirectly address the immune system; for example, scarred bilateral tympani suggest recurrent ear infections. More commonly, patients with immunodeficiency might otherwise look like normal individuals, unless severe infections had produced an organ damage or had delayed growth and development. However, attention to details in the physical examination may supply important clues that suggest immune dysfunction. In a normal child, a paucity of lymphoid tissue, such as tonsils and lymph nodes, might reflect impaired development resulting from immune deficiency. This is especially seen in patients with X-linked agammaglobulinemia. Certain physical findings are suggestive of syndromic immunodeficiencies, such as with telangiectasia over the bulbar conjunctivae and face with or without ataxia in ataxia–telangiectasia (AT); chronic eczema and delayed shedding of primary teeth in hyper-IgE syndrome (HIES); severe eczema in immunodeficiency, polyendocrinopathy and enteropathy, X-linked (IPEX) syndrome, and WAS; chronic periodontitis in defects of the neutrophils; or silvery hair, pale skin, and photophobia in Chediak-Higashi syndrome. Investigation for Shwachman-Bodian-Diamond syndrome (SBDS) should be considered in patients with neutropenia, especially if they also present with skeletal dysplasia. Patients with DiGeorge syndrome and nuclear factor-κB (NF-κB) essential modulator (NEMO) deficiency present with characteristic facies.

Children with severe immune defects are small for their age, with growth delay secondary to recurrent infections. Hepatosplenomegaly and diffuse lymphadenopathy might suggest HIV infection or a disorder of immune dysregulation. Children with leukocyte adhesion defect (LAD) can present with severe gingivostomatitis and dental erosion as a consequence of abnormal leukocyte function (Fig. 32.3). Multiple scars from skin abscesses might suggest neutrophil defects, and scarred tympani with reduced hearing might indicate a history of recurrent otitis media, which can be associated with antibody deficiency.

LABORATORY TESTING FOR IMMUNE FUNCTION

Results from commonly ordered tests might provide with a great deal of information about the immune system. The complete blood count (CBC) with differential and platelet determination is ordered to quantitate the total white blood cell (WBC) count and total numbers of neutrophils, lymphocytes, eosinophils, and platelets. Abnormal counts should be determined by using age-specific ranges. Leukocytosis, neutropenia, lymphopenia, and abnormalities in WBC morphology can be detected from this test. Persistent neutrophilia might suggest LAD. Anemia may be present in children with chronic disease. Platelet counts may be abnormally low in children with poor bone marrow function or autoimmune disease, and platelets will be reduced in number and morphologically small in children with WAS. Chemistry panels, including serum liver enzymes levels, might suggest organ compromise as a result of infections or autoimmunity associated with immunodeficiency. Low protein levels suggest malnutrition and conditions associated with protein losses, which may cause hypogammaglobulinemia. Examination of the posteroanterior and lateral chest radiographs to look for a thymic shadow can be helpful because its absence suggests impaired T-cell development. This is especially useful in infants because the thymus mass normally involutes with age. In addition, the thymus may shrink in response to such stresses as surgery, infection, or high-dose steroid treatment.

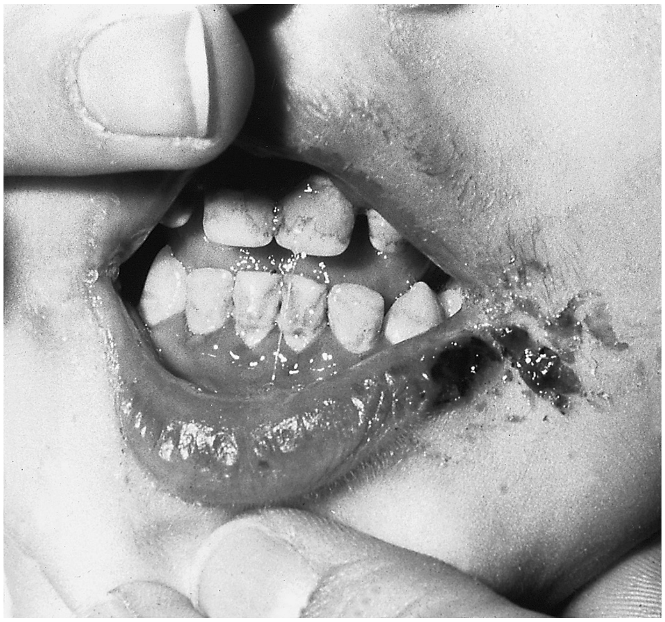


FIG 32.3 Severe gingivostomatitis and dental erosion in a 2-year-old child with leukocyte adhesion defect (LAD). (Courtesy of Dr. D.C. Anderson.)

HIV infection can be ruled out by screening with measurement of anti-HIV antibodies, by the enzyme-linked immunosorbent assay (ELISA) or the rapid HIV test. In those individuals with suspected humoral immunity defect and in children younger than 18 months of age, a polymerase chain reaction (PCR)–based test to detect HIV viremia should be performed to avoid false-negative results and confounding maternal anti-HIV antibodies, respectively.

Immunology Testing

Specific immunological testing is guided by clues obtained from the history and physical examination and common screening laboratory tests.

Serum Immunoglobulin Levels

The levels of IgG, IgA, IgE, and IgM can be measured in serum. The IgA level is especially helpful in that it is low in all permanent types of agammaglobulinemia and in selective IgA deficiency.

KEY CONCEPTS

Diagnosis of Immune Deficiencies

Features of Congenital Antibody Deficiencies

- Free of infections until 6–9 months of age, when maternal antibodies that passed through the placenta to infant are below protective levels
- Severe infections with bacterial organisms, especially sinusitis, otitis media, and pneumonias caused by encapsulated bacteria, such as *Streptococcus pneumoniae*

Features of Congenital T-Cell Immunodeficiency

- Onset of thrush, diarrhea, and failure to thrive in the first months of life
- Severe infections with opportunistic microorganisms, such as *Pneumocystis jiroveci, Candida albicans,* adenovirus, cytomegalovirus (CMV), Epstein-Barr virus (EBV). "BCG-itis" in areas where Bacille Calmette-Guérin (BCG) immunization is mandatory
- Absolute and relative lymphopenia

Screening Tests for Suspected Immunodeficiency

- Evaluation for neutropenia, lymphopenia, thrombocytopenia, and/or small platelets
- Immunoglobulin levels and specific antibodies to childhood immunizations
- Lateral chest radiography in infants for thymus shadow
- Consider flow cytometry to quantify T cells, T-cell subsets, B cells, and natural killer (NK) cells (especially in infants)
- Measurement of CH50 activity
- Test for oxidative burst in phagocytes

IgE level measurement is of significance for the diagnosis of HIES. Serum IgG subclasses levels can be determined. However, rather than using IgG subclass levels to screen for immunodeficiency, they are best utilized when patients have clinical conditions associated with specific antibody deficiencies but normal total IgG levels. In some of these patients, IgG subclass deficiency, particularly IgG2 and IgG3 deficiencies, might be present. IgG2 subclass deficiency has been linked with selective IgA deficiency and deficiency of antipolysaccharide antibodies. IgA subclass low levels, IgA1, IgA2, have not been associated with a specific immune defect, and there is no validity for measuring these. The variation of normal ranges of human serum Igs with age is an important consideration in children, since IgA and IgG subclass levels may not reach normal adult reference ranges until 6 years of age.[1]

B-Cell Function: Specific Antibody Production

To properly assess B-cell function, specific antibody production must be measured. Patients with normal Ig and Ig subclass levels might exhibit deficient antigen-dependent antibody responses. An initial screen of antibody production may involve the quantification of isohemagglutinins. Isohemagglutinins occur in all individuals except those with blood type AB; isohemagglutinins are natural IgM antibodies to polysaccharide blood group antigens A and/or B, which are not expressed in the red blood cells (RBCs) of the patient tested. Individuals form isohemagglutinins as a result of environmental exposure to ubiquitous antigens that share epitopes with blood antigens. Children less than 1 year of age do not reliably have measurable serum isohemagglutinins because of the limited exposure to the environment. A patient with blood type A should have anti-B IgM; patients with blood type B should have anti-A IgM; and patients with blood type O should have both anti-A and anti-B IgM. These antibodies are normally present in titers greater than 1:10; individuals with poor antibody production may have low or absent titers. Specific IgG antibody production can be measured following immunization with protein antigens, such as toxoids derived from *Tetanus* and *Diphtheria* organisms, and polysaccharide antigens, such as those produced by pneumococci and *Haemophilus influenzae.* For pneumococcal immunization, there are two vaccines that need to be differentiated. The conjugated vaccine containing 13 pneumococcal serotypes (PREVNAR13, Wyeth) is currently included in the universal schedule of immunizations for infants and toddlers and induces a robust, T cell–dependent immune response. The unconjugated 23-valent pneumococcal polysaccharide vaccine (Pneumovax, Merck) is available for immunization to adults and children aged 2 years and older. The immune response for this vaccine is considered less dependent on T cells and also less lasting than the conjugated vaccine. The pneumococcal antigen challenge using the unconjugated vaccine is not recommended for children under 2 years of age because healthy children are not thought to reliably respond to the unconjugated pneumococcal antigen at this age. However, this view has been challenged by data showing that 1-year-old children produce normal antibody responses to this unconjugated vaccine.[13,14] Normal antibody responses are usually demonstrated with an over twofold rise in specific antibody levels within 2–3 weeks for protein antigens and within 4–6 weeks for polysaccharide antigens.[15] Patients with agammaglobulinemia are expected not to produce antibody responses, whereas others, such as those with IgG2 subclass deficiency and normal levels of total IgG, may only have difficulty with antibody production following immunization with polysaccharide antigens. Patients with selective IgA deficiency, alone or with transient hypogammaglobulinemia of infancy, have normal specific IgG antibody production, by definition. The pneumococcal serotypes included in the current conjugated antipneumococcal vaccine, serotypes 1, 3, 4, 5, 6A, 6B, 7F, 9V, 14, 18C, 19A, 19F, and 23F, were estimated to be responsible for approximately 90% of invasive pneumococcal disease in children less than 5 years of age worldwide.[16] Previous immunization with the conjugate vaccine does not preclude use of the unconjugated pneumococcal polysaccharide vaccine. The 23-valent polysaccharide vaccine provides the potential for stimulation and measurement of a protective immune response to additional 11 serotypes (2, 8, 9N, 10A, 11A, 12F, 15B, 17F, 20, 22F, 33F) not included in the conjugated vaccine. Testing for antibodies against serotypes not included in the two vaccines and comparing the antibody titers in the pre- and postimmunization blood samples helps in the assessment of specific increase of particular antiserotype antibody titers as a response to the vaccine administration.

Evaluation of Cellular Immunity

Cellular immune function had been historically screened with the use of the delayed-type hypersensitivity (DTH) skin tests. DTH reactions occur 48–72 hours after antigen exposure, although antigens used to test DTH can occasionally produce immediate hypersensitivity reactions. The clinical response and time of observation were to discriminate between immediate and delayed reactions. DTH reactions involve the production of local edema and vasodilation as a result of inflammatory cytokines secreted by antigen-specific T cells, followed by lymphocyte infiltration and maximal induration 48 hours after antigen intradermal injection of an antigen. Generally, DTH skin tests are performed using vaccine antigens or microbial agents to which the patient

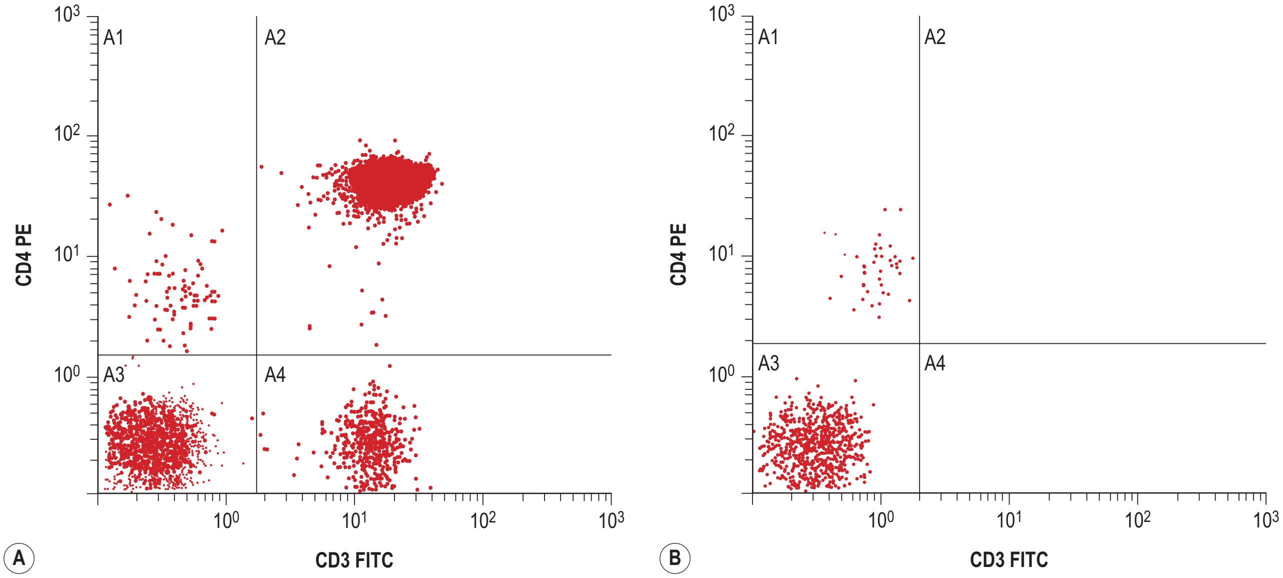


FIG 32.4 Histograms of Fluorescently Stained Lymphocytes. The quadrant of interest, A2, shows lymphocytes that are positive for labeling with both fluorescein isothiocyanate (FITC)–tagged monoclonal antibodies specific for CD3 and phycoerythrin (PE)–tagged monoclonal antibodies (mAbs) specific for CD4. The histogram on the left shows normal fluorescence as a result of CD4 T lymphocytes present in quadrant A2. The histogram on the right shows absence of CD4 T lymphocytes in an infant with SCID.

has had previous exposure. Commonly used antigens include tetanus toxoid, mumps, and extracts from *Candida albicans* and *Trichophyton* spp. The purified protein derivative (PPD), or tuberculin, can serve as a negative control in most patients in developed countries with a low incidence of tuberculosis. Likewise, a positive PPD result indicates sensitivity to mycobacteria as well as robust cellular immunity. Virtually all children and adults with previous exposure should respond to at least one antigen in a panel of tetanus toxoid, mumps, and *C. albicans*. *Anergy*, or nonresponse to the antigen following previous exposure, may indicate a cellular defect. A nonresponder should be retested with an *in vitro* lymphocyte evaluation.

Lymphocyte subset enumeration. Quantitation of B- and T-cell subsets narrows the differential diagnosis and provides evidence for the diagnosis of combined, cellular, or antibody immunodeficiency (Chapters 34, 35, 93, 94, 95,-96). Both T and B cells can be identified and labeled by using flow cytometry and fluorescent monoclonal antibodies (mAbs) (Chapter 92). T-cell enumeration involves the use of a pan–T cell mAb specific for CD3. The CD4 marker serves as identification for T-helper (Th) cells. CD8 marker characterizes cytotoxic T cells. B cells can be identified by using mAbs against the cell surface markers CD19 or CD20. Natural killer (NK) cells can be identified by using mAbs against CD16 and CD56. Specialized clinical laboratories are available to measure lymphocyte markers of importance to particular diseases; for instance, the proportion of αβ T-cell receptor (αβTCR) and γδTCR double-negative $CD3^+$ T cells are of relevance in the diagnosis of autoimmune lymphoproliferative syndrome (ALPS). T cell subsets can be characterized as naïve or activated based on the expression of CD45RA and CD45RO antigens.

B-cell panels and NK-cell panels. These panels have been designed to characterize the maturation stage of these cells and provide support for the diagnosis of an specific immunodeficiency. For example, the proportion of class-switched B cells has a predictive value for autoimmune and granulomatous complications in CVID.

In flow cytometry, the fluorescence intensity corresponding to cells labeled with each specific antibody is obtained (Fig. 32.4) and the percentage of the specific lymphocyte subset can be estimated. A reference range is available for each subset defining normal values as those whose values fall between the fifth and 95th percentages for this population. Separate ranges should be used for children because infants and children generally have higher absolute numbers of T-cell subsets and higher percentages of CD4 T cells (Fig. 32.5 and Appendix 2). Nonimmune factors, such as age, gender, and adrenocorticoid levels, influence the expression of blood lymphocyte subset populations. Therefore interpretation of lymphocyte phenotyping should take into consideration the clinical status of the patient. For example, transient moderate lymphopenia with predominance of T cells and NK cells might be seen in patients admitted to intensive care units. HIV infection causes progressive depletion of CD4 T cells. These abnormalities resolve when the patient's condition improves.

Lymphocyte functional analysis. To test lymphocyte function in the laboratory, mitogen- and antigen- induced lymphocyte proliferation or transformation studies are performed (Chapter 93). For these studies, lymphocytes are stimulated to proliferate involving new DNA synthesis and cell division. Lymphocytes from immunized or previously exposed individuals will normally proliferate in response to antigens to which they are sensitized. This response *in vitro* correlates with the *in vivo* DTH response. Mitogens, such as concanavalin A (ConA), phytohemagglutinin (PHA), and pokeweed mitogen (PWM), stimulate proliferation of normal T cells, as can allogeneic histocompatibility antigens when leukocytes from two donors are mixed in culture. Proliferation

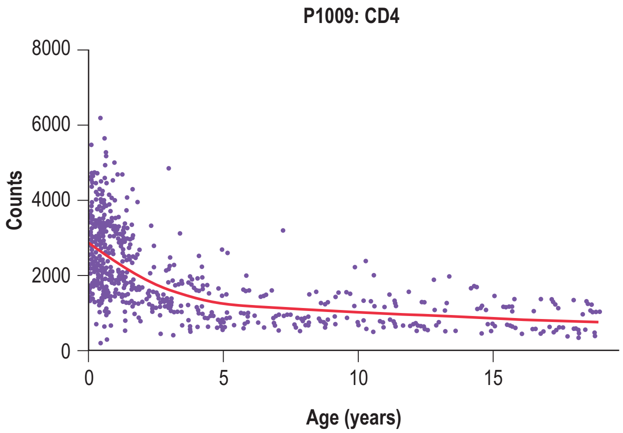


FIG 32.5 **Change in Distribution of Peripheral Blood CD4 T-Cell Subsets With Age in Healthy Children.** Scatter plot indicates peripheral blood CD4 T-cell counts (cells/μL) by age, with lowest curves in healthy children from birth to 18 years of age. (From Shearer WT, et al. Lymphocyte subsets in healthy children from birth through 18 years of age: The Pediatric AIDS Clinical Trials Group P1009 study. J Allergy Clin Immunol 2003; 112: 973, with permission from Elsevier.)

of lymphocytes can be evaluated by the demonstration of cell division or by increased DNA synthesis reflecting this cell process. Increased DNA synthesis is monitored by the incorporation of radiolabeled nucleotides, usually tritiated thymidine, in culture media. A measure of the amount of radioactivity in the cells correlates with DNA synthesis. Other assays to assess mitogen-induced cell proliferation measure deoxybromouridine incorporation, change in pH, or adenosine triphosphate (ATP) concentration of the culture media. These assays are being increasingly used as surrogate markers of cellular immunity; however, a comparison with the traditional assay based on radiolabeled nucleotide is not available. Of note, a flow cytometry assay that measures cell division with the use of carboxyl fluorescein succinyl ester (CFSE), a fluorescent compound that distributes evenly in cells and specific antibodies, is also increasingly used in clinical immunology.[17] CFSE is distributed equally in dividing cells, and each progeny cell has half the fluorescence intensity of CFSE compared with the parent cell, providing the basis to identify these dividing cells. After mitogen or antigen stimulation, mononuclear cells can be stained with specifically labeled antibodies, allowing the identification of cell subsets that proliferate.

Phagocytes

The laboratory evaluation of a patient with a suspected phagocyte deficiency (Chapters 22, 94) should always begin with a CBC. Neutropenia is the most frequently encountered disorder of the phagocyte system.[18] Neutrophilia, at values exceeding those associated with acute infection, is a common finding in LAD type 1 (LAD-1). Abnormalities of WBC function involve difficulty with adherence, locomotion, deformability, recognition, attachment, engulfment, phagosome formation, phagocytosis, degranulation, microbial killing, and elimination of engulfed material. Clinical assays to evaluate neutrophil function are limited in number. CGD is diagnosed by demonstrating absent or markedly reduced oxidase activity in neutrophils in response to stimulation. Oxidase activity can be detected by a flow cytometry assay measuring the oxidation of dihydrorhodamine (DHR) 123 in phagocytes, resulting in fluorescent rhodamine-123.[19] The nitroblue tetrazolium (NBT) test measures oxidative burst activity as well, but it is a more subjective test and can miss the diagnosis of CGD. For patients with suspected LAD-1 deficiency, neutrophils are labeled with mAb directed against the adhesion molecule CD11/CD18 heterodimer. Absence of fluorescence intensity indicates lack of expression of the adhesion molecule. In addition, an increase of fluorescence intensity after stimulation can be documented in normal individuals, indicating the normal upregulation of this molecule after cell activation.[20] Other laboratory techniques used to identify phagocytic defects include assays for chemotaxis and bactericidal activity. A major pitfall for neutrophil studies is the spontaneous cell activation that might occur *in vitro* when cells are not tested within a few hours of when the sample was drawn, resulting in artifactual values that might falsely suggest poor function.

Complement

Laboratory tests for complement components include tests for functional activity of the classical pathway with a CH50 assay and the alternative pathway with an AH50 assay, as well as immunochemical methods to measure complement component levels.[21] The CH50 evaluation tests the ability of fresh serum from the patient to lyse antibody-coated sheep erythrocytes. This reflects the activity of all numbered components of the classical complement pathway, C1–C9, and terminal components of the alternative complement pathway. A total deficiency of one of the classical complement pathway components will result in a CH50 assay approaching zero (Chapter 21). Patients with complement deficiency are rare, and complement test abnormalities are often transient because of increased consumption or activation. It is usually recommended that that in case of an abnormal result, the complement test be repeated if the sample was taken when the patient had an acute illness. Quantitative tests for components C3 and C4 are utilized in testing for complement deficiencies and for evaluation of complement activation (Chapter 21).

Innate Immunity: Interferon-γ Levels, Toll-Like Receptor Assay

The importance of the many components of innate immunity are increasingly recognized, as single gene defects in this immune compartment have been found to cause susceptibility to specific infections.[22] For example, patients with defects in the proteins that are part of the interferon-γ (IFN-γ) receptor may have elevated serum IFN-γ levels, even when there is no infection to explain these levels. The IFN-induced response associated kinase 4 (IRAK4) defect, observed with susceptibility to pneumococcal infection, might be accompanied with abnormal Toll-like receptor (TLR) assay responses. It should be noted that the clinical value of most of these innate immunity tests as screening or diagnostic tools for immune defects has not been clearly established.

The testing of lymphocyte apoptosis in a patient who may have ALPS and the evaluation of NK-cell function for suspected familial hemophagocytic lymphohistiocytosis are examples of specific functional assays that suggest immunodeficiency syndromes. Many patients with increased frequency of infections may not have abnormal results in clinically available immunological testing, which may not give clear evidence of a secondary etiology in the medical evaluation. In these difficult cases, referral to tertiary care and research centers for investigation of rare diseases is recommended.

MOLECULAR TESTING FOR PRIMARY IMMUNE DEFECTS

Molecular testing for specific PIDs is available through commercial and research laboratories.[23] Biochemical and genetic testing should be considered. If autosomal recessive SCID is suspected, the adenosine deaminase (ADA) and purine nucleoside phosphorylase (PNP) enzyme activities in the RBCs should be determined. White blood cells must be used to measure the activity of these enzymes in recently transfused individuals, since donor RBCs will elevate the enzyme activity in deficient patients. AT has the consistent laboratory finding of elevated alpha-fetoprotein (AFP) levels along with variable abnormalities in B- and T-cell function. Nearly 300 defective genes and gene products have been identified to result in congenital immunodeficiency syndromes.[24] In these cases, the diagnosis can be confirmed with molecular genetic analysis (Chapter 33). For example, gene mutations in *BTK* leading to the absence of Bruton tyrosine kinase (BTK) results in arrest of B-cell development at the pre–B cell stage in congenital X-linked agammaglobulinemia. Similarly, abnormal T-cell development leading to SCID results from mutations in at least 15 genes, including *IL2RG* and *JAK3*. Patients with gene mutations that may result in disease but have not been investigated need to be carefully evaluated to demonstrate the pathogenic nature of the genetic change. Some genetic changes do not have clinical significance and are known as single nucleotide polymorphisms (SNPs). Standard protocols using Southern, Northern, and Western blot analyses, PCR analysis, and DNA sequence analysis are helpful to identify affected patients, affected fetuses prenatally, and carriers of genetic mutations (Chapter 96). Most recently, the use of whole exome sequencing for immunodeficiency syndromes has facilitated the identification of new genes causing immunodeficiencies by examining all known gene exons without bias, and simultaneously. This methodology for diagnosis is particularly helpful when the clinical presentation does not match any of the already described immunodeficiency syndromes.[25]

CONCLUSIONS

CLINICAL PEARLS

Management of Immunodeficiency Patients

- Lymphopenia is a hallmark of T-cell immunodeficiency in infancy. Unexplained lymphopenia should be recognized and evaluated.
- Normal range for immunoglobulin levels and lymphocyte counts varies with age; age-matched controls should be used for interpretation.
- Survival and morbidity outcomes of hematopoietic stem cell transplantation (HSCT) for severe combined immunodeficiency (SCID) are best when performed in the first 4 months of age.
- Delays in treatment of immunodeficiency leads to infections causing permanent damage to organ systems, such as the lungs, where bronchiectasis or bronchiolitis obliterans may develop secondary to recurrent pneumonias.
- Implementation of universal screening of newborn infants by DNA analysis of dried blood spots on Guthrie cards (T-cell receptor excision circle [TREC]) detect all T-cell deficiencies.

The approach to the patient with suspected immune deficiency requires knowledge of developmental pathways and function of the different compartments of the immune system, as well as the clinical presentation of these disorders. The medical history, particularly the frequency, severity, and etiology of infections, is most helpful to orient the diagnostic workup. Commonly ordered tests in primary care, such as a CBC and serum Ig levels, are helpful to support possible diagnosis and referral to the clinical immunologist. Immunological testing according to clues obtained from the medical history helps narrow the differential diagnoses to specific immunodeficiencies, which are confirmed by molecular methods. Description of new T-cell subsets (*e.g.,* Th17 and regulatory T cells [Tregs]) has helped explain the immunopathogenesis of certain clinical manifestations, such as the occurrence of autoimmunity in patients with combined immunodeficiency, and "cold abscesses" in the autosomal dominant HIES. Testing for these lymphocyte phenotypes is being integrated in the clinical evaluation. Identification of genetic defects can now be accomplished by increased availability of whole exome sequencing, as an alternative to genetic analysis of candidate genes. Technological advances are making molecular diagnosis available for most patients with immunodeficiency conditions.

ON THE HORIZON

- Early diagnosis of severe combined immunodeficiency (SCID) and non-SCID is the key to successful hematopoietic stem cell transplantation (HSCT).
- Development of DNA sequence analysis for >300 types of primary immunodeficiency (PID).
- Wide availability of whole-genome sequencing methods to determine molecular defects for difficult to diagnose immunodeficiency syndromes

ILLUSTRATIVE CASES

Case 1

A 6-month-old Caucasian female infant was brought with a history of rash and otitis media that had been recurrent since 2 weeks of age. The infant had poor weight gain, frequent spitting up, coughing spells, and persistent diarrhea. Maternal HIV testing showed negative results. Physical examination revealed an emaciated infant without palpable lymphoid tissue and severe oral thrush. Cystic fibrosis was ruled out by using genetic studies. Stool viral cultures were persistently positive for rotavirus.

Immunodeficiency was suspected because of the infant's failure to thrive, persistent chronic diarrhea, and recurrent infection. An evaluation of the immune system was performed (Table 32.3). Humoral immunity testing showed Ig levels below or just above the lower limit of normal. Isohemagglutinins were not tested as the patient was less than 1 year old. As she had not been immunized, specific antibody titers to vaccines were not tested. Stool α_1-antitrypsin levels were normal, suggesting that protein was not being lost in stool, and urinalysis did not show protein loss. CBC revealed profound lymphopenia and absolute neutropenia. The neutropenia had resolved upon subsequent evaluations, but the lymphopenia persisted. A thymic shadow was not present on the chest radiograph. Evaluation of lymphocyte subsets by flow cytometry revealed that the CD4 and CD8 T-cell and B-cell numbers were all markedly low. The patient's lymphocytes had no proliferative response when stimulated with phytohemagglutinin. The history and the physical examination, along with laboratory values, including low Ig levels and markedly low lymphocyte numbers with poor proliferative response to mitogen stimulation, strongly suggested the diagnosis of SCID.

TABLE 32.3 Results of Screening Tests and Special Evaluations of Illustrative Cases*

	CASE NUMBER			
Test	**1**	**2**	**3**	**4**
Antibody Function				
Serum immunoglobulin (mg/dL)				
IgG	235 (208–686)	Normal	41 (256–1067)	748 (520–1340)
IgA	<8 (10–62)	Normal	7 (12–103)	40 (25–81)
IgM	<7 (43–183)	Normal	14 (47–173)	78 (64–275)
Isohemagglutinins	ND	ND	Negative	Normal
Functional antibody tests				
Antitetanus	ND	Normal	Undetectable	0.28
Antidiphtheria	ND	Normal	Undetectable	0.12
Anti-*Pneumococcus*	ND	2 of 14 protective	ND	Normal
Lymphocyte Numbers				
Lymphocyte count ($\times 10^3$ cell/μL)	0.217 (2.5–16.5)	2.8 (2.3–5.4)	2.56 (4.0–13.5)	10.60 (3.0–9.5)
CD4 number ($\times 10^3$ cell/μL)	0.009 (1.6–4.0)	1.27 (0.9–2.4)	1.72 (1.4–4.3)	ND
CD8 number ($\times 10^3$ cell/μL)	0.008 (0.5–1.7)	1.5 (0.5–1.5)	0.73 (0.5–1.7)	ND
CD19 number ($\times 10^3$ cell/μL)	0.010 (0.3–2.0)	0.023 (0.4–1.4)	0.107 (0.6–2.6)	ND
Lymphocyte Proliferation (counts/min)				
PHA, Phytohemagglutin	546 (55494)	Normal	210973 (177195)	ND
ConA, Concanavalin A	ND	Normal	124680 (143591)	ND
PWM, Pokeweed mitogen	ND	Normal	162646 (138684)	ND
Complement Function				
Total hemolytic complement (U/mL)	ND	ND	618 (300–500)	<100 (300–500)
Phagocyte Function				
Neutrophil count (x10e3 cell/mL)	0.109 (1.0–9.0)	0.570 (1.5–8.5)	3.3 (1.0–8.5)	4.2 (1.5–8.5)
Monocyte count (x10e3 cell/mL)	0.546	Normal	0.11	0.65
Oxidative activity (DHR)	ND	Normal	Normal	ND

*Normal values are reported in parentheses. Normal values may vary from laboratory to laboratory as a result of different techniques used and may vary according to patient age.
ConA, concanavalin A; DHR, dihydrorhodamine; ND, not done; PHA, phytohemagglutinin; PVVM, .

The patient had low levels of ADA and increased levels of the toxic nucleotide metabolites. SCID as a result of ADA deficiency was diagnosed.

Following ADA enzyme replacement therapy and Ig infusions, the patient thrived and had fewer infections. Her lymphocyte numbers and function improved.

Case 2

A 4-year-old female child was presented with a history of short stature and frequent upper and lower respiratory infections. Family history was not contributory. Physical examination was remarkable for height and weight less than the fifth percentile for age and a significant chest deformity (Fig. 32.6). Radiological examination demonstrated metaphyseal dysplasia and *pectum carinatum*, with significant bilateral lung atelectasis. Hematological evaluation showed severe neutropenia, a normal lymphocyte count, and mild thrombocytopenia (see Table 32.3). Examination of bone marrow revealed mild dysplasia with the presence of three lineages and no evidence of chronic viral infections. Evaluation for failure to thrive was significant for low levels of pancreatic enzymes. A sweat test was done to screen for cystic fibrosis, and the result was normal.

Immunological evaluation showed normal serum levels of IgG, IgA, and IgM, with mildly increased IgE levels. Specific antibody responses to childhood vaccines were protective against tetanus, diphtheria, and *H. influenzae* antigens, and to only two of 14 pneumococcal serotypes were measured despite the child being fully immunized. Lymphocyte phenotyping revealed low B cells and NK cells. A diagnosis of Schwachman-Bodian-Diamond syndrome (SBDS) was made on the basis of the patient's unique clinical manifestations. Genetic sequencing analysis of the *SBDS* gene identified one mutated allele that had not been previously reported, and molecular investigations for chromosomal breakage defects, other skeletal genetic syndromes, and mitochondrial defects showed negative results. The patient was treated with pancreatic enzyme replacement, antibiotic prophylaxis, and close hematological follow-up.

Case 3

A 12-year-old Caucasian male child had a history of diarrhea that lasted for 1 month, nuchal rigidity, left facial weakness, and a left sixth cranial nerve palsy when he was 8 months of age, although he had been well for the first 6 months of life. The family history was negative for immunodeficiency. The patient had received his childhood immunizations, including the live polio virus vaccine. Physical examination revealed absence of tonsillar tissue, and the patient had no palpable lymph nodes. The presenting neurological symptoms resolved, but the patient subsequently developed a lower-extremity flaccid paralysis that has persisted.

Immunodeficiency was suspected because of the chronicity of the illness, the lack of lymphoid tissue on physical examination, and the finding of an unusual infectious agent causing severe sequelae to infection. The age at presentation was typical for

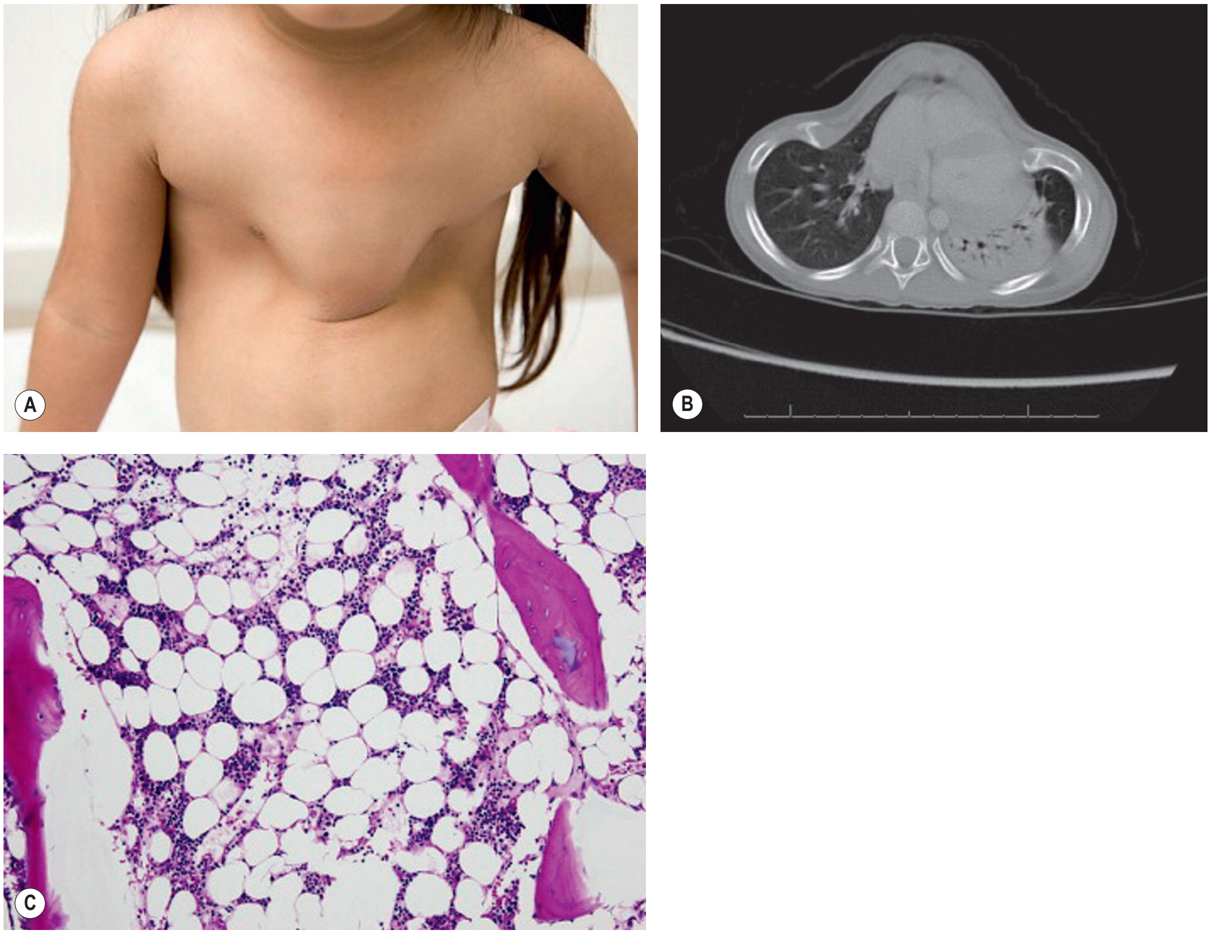


FIG 32.6 **A 4-Year-Old Girl Presenting With Congenital Neutropenia, Small Stature, and Pancreatic Insufficiency.** Genetic sequencing analysis confirmed diagnosis of Schwachman-Bodian-Diamond syndrome. (A) *Pectus carinatum.* (B) Chest computed tomography (CT) scan. (C) Bone marrow biopsy (hematoxylin and eosin stain; original magnification, ×100). (From Shah SS, et al. Diagnosis of primary immunodeficiency: let your eyes do the talking. J Allergy Clin Immunol 2009;124:1363–4, with the permission from Elsevier.)

patients with a defect in humoral immunity. Immune system evaluation (see Table 32.3) showed a normal absolute neutrophil count and a low absolute lymphocyte count. The CH50 was mildly elevated; this result ruled out a complement component deficiency from the differential diagnosis. Ig levels were low. Antibody levels to diphtheria and tetanus were very low. A thymus shadow was present on the chest radiograph. Lymphocyte evaluation showed normal numbers and percentages of T cells and absent numbers and percentages of B cells. Proliferative responses of lymphocytes to mitogens were normal. Molecular testing showed a mutation in the *BTK* gene. The CSF was remarkable for lymphocytic pleocytosis, but the protein and glucose levels were normal. Stool and throat cultures revealed vaccine-strain poliovirus. The patient was diagnosed with vaccine-acquired polio virus infection and X-linked agammaglobulinemia.

The patient receives monthly Ig replacement therapy and is otherwise well.

Case 4

The patient was a 3-year-old girl who was hospitalized for pneumonia. Her past medical history was significant for recurrent otitis media and bronchitis. The family history was negative for diagnosed immunodeficiency, although two siblings had a history of recurrent otitis media and bronchitis. The physical examination revealed normal tonsillar tissue and the presence of small lymph nodes in the cervical region.

Laboratory values (see Table 32.3) included normal numbers of granulocytes and lymphocytes. Serum Ig levels were normal. The patient had been immunized and had adequate antibody titers to tetanus, diphtheria, and *Pneumococcus* antigens. The CH50 level was below the lowest measurable level. C3 and C4 levels were normal, but the level of C2 was 0. The patient's father, mother, and brother had levels of C2 that were half the normal value, consistent with being heterozygous for an abnormal C2 allele. C2 deficiency was diagnosed. The patient has had chronic sinusitis and bronchitis despite antibiotic prophylaxis. C2 deficiency is most frequently associated with autoimmune disease; this patient has not had autoimmune disease and undergoes yearly surveillance examinations.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Yel L. Selective IgA deficiency. J Clin Immunol 2010;30:10–16.
2. Kwan A, Abraham RS, Currier R, et al. Newborn screening for severe combined immunodeficiency in 11 screening programs in the United States. JAMA 2014;312:729–38.
3. Boyle JM, Buckley RH. Population prevalence of diagnosed primary immunodeficiency diseases in the United States. J Clin Immunol 2007;27:497–502.
4. UNAIDS. AIDS by the numbers 2015. Joint United Nations Programme on HIV/AIDS (UNAIDS) and World Health Organization (WHO); 2015.
5. Agarwal S, Mayer L. Diagnosis and treatment of gastrointestinal disorders in patients with primary immunodeficiency. Clin Gastroenterol Hepatol 2013;11:1050–63.
6. Goronzy JJ, Weyand CM. Understanding immunosenescence to improve responses to vaccines. Nat Immunol 2013;14:428–36.
7. Jones LL, Hashim A, McKeever T, et al. Parental and household smoking and the increased risk of bronchitis, bronchiolitis and other lower respiratory infections in infancy: systematic review and meta-analysis. Respir Res 2011;12:5.
8. Shearer WT, Fleisher TA, Buckley RH, et al. Medical Advisory Committee of the Immune Deficiency Foundation, Recommendations for live viral and bacterial vaccines in immunodeficient patients and their close contacts. J Allergy Clin Immunol 2014;133:961–6.
9. Marciano BE, Huang CY, Joshi G, et al. BCG vaccination in patients with severe combined immunodeficiency: complications, risks, and vaccination policies. J Allergy Clin Immunol 2014;133:1134–41.
10. Silverberg JI, Silverberg NB. Childhood atopic dermatitis and warts are associated with increased risk of infection: a US population-based study. J Allergy Clin Immunol 2014;133:1041–7.
11. Guerra IC, Fawcett WA, Redmon AH, et al. Permanent intrinsic B cell immunodeficiency caused by phenytoin hypersensitivity. J Allergy Clin Immunol 1986;77:603–7.
12. Rytter MJ, Kolte L, Briend A, et al. The immune system in children with malnutrition—a systematic review. PLoS ONE 2014;9:e105017.
13. Ozen A, Baris S, Karakoc-Aydiner E, et al. Outcome of hypogammaglobulinemia in children: immunoglobulin levels as predictors. Clin Immunol 2010;137:374–83.
14. Balloch A, Licciardi PV, Russell FM, et al. Infants aged 12 months can mount adequate serotype-specific IgG responses to pneumococcal polysaccharide vaccine. J Allergy Clin Immunol 2010;126:395–7.
15. Orange JS, Ballow M, Stiehm ER, et al. Use and interpretation of diagnostic vaccination in primary immunodeficiency: a working group report of the Basic and Clinical Immunology Interest Section of the American Academy of Allergy, Asthma & Immunology. J Allergy Clin Immunol 2012;130:S1–24.
16. Johnson HL, Deloria-Knoll M, Levine OS, et al. Systematic evaluation of serotypes causing invasive pneumococcal disease among children under five: the pneumococcal global serotype project. PLoS Med 2010;7:pii: e1000348.
17. Rakha A, Todeschini M, Casiraghi F. Assessment of anti-donor T cell proliferation and cytotoxic T lymphocyte-mediated lympholysis in living donor kidney transplant patients. Methods Mol Biol 2014;1213:355–64.
18. Henry M, Sung L. Supportive care in pediatric oncology: oncologic emergencies and management of fever and neutropenia. Pediatr Clin North Am 2015;62:27–46.
19. Holland SM. Chronic granulomatous disease. Hematol Oncol Clin North Am 2013;27:89–99.
20. Etzioni A. Defects in the leukocyte adhesion cascade. Clin Rev Allergy Immunol 2010;38:54–60.
21. Grumach AS, Kirschfink M. Are complement deficiencies really rare? Overview on prevalence, clinical importance and modern diagnostic approach. Mol Immunol 2014;61:110–17.
22. Rosenzweig SD, Holland SM. Recent insights into the pathobiology of innate immune deficiencies. Curr Allergy Asthma Rep 2011;11:369–77.
23. https://www.genetests.org/tests/.
24. Picard C, Al-Herz W, Bousfiha A, et al. Primary Immunodeficiency Diseases: an Update on the Classification from the International Union of Immunological Societies Expert Committee for Primary Immunodeficiency 2015. J Clin Immunol 2015;35:696–726.
25. Platt C, Geha RS, Chou J. Gene hunting in the genomic era: approaches to diagnostic dilemmas in patients with primary immunodeficiencies. J Allergy Clin Immunol 2014;134:262–8.

MULTIPLE-CHOICE QUESTIONS

1. Which of the following diagnoses is most likely to present with *Pneumocystis jiroveci* pneumonia during the first year of life?
   A. Partial DiGeorge syndrome
   B. Human immunodeficiency virus (HIV) infection
   C. Common variable immunodeficiency (CVID)
   D. Chronic granulomatous disease (CGD)
   E. Complement deficiency

2. Deficiencies of terminal component of the complement system is associated with frequent or severe infections caused by:
   A. *Candida* spp.
   B. Epstein-Barr virus
   C. *Neisseria* spp.
   D. *Mycobacterium tuberculosis*
   E. *Cryptosporidium parvum*

3. Which of the following conditions is MOST likely to contribute to frequent respiratory infections?
   A. Adolescence
   B. Peanut allergy
   C. Dust mite allergy
   D. Homeschooling
   E. Protein-rich diet

Human Genomics in Immunology

Robert L. Nussbaum, Jennifer M. Puck

The completion of the Human Genome Project in 2003, 50 years after the landmark 1953 publication of the double-helical structure of DNA by James Watson and Francis Crick, is correctly deemed a major milestone in modern biology. The sequence provided a first comprehensive and accurate view of the genetic makeup of humans, with a low error rate and only a few hundred gaps of indeterminate sequence. The first detailed picture of the structure of a composite human genome is for human genetics what Vesalius's publication of a consensus structure of the human body, *De humani corporis fabrica,* was for anatomy. And, like Vesalius's work, it continues to serve as a foundation for further scientific discovery in such areas as genetic variation, gene function, human physiology, and the genetic basis for disease. The years following the completion of the first human genome sequencing have seen progress in all of these areas.

KEY CONCEPTS

The sequence of the human genome and catalogue of extensive genome variation among humans has revolutionized our approach to heritable immune disorders.

Genomic variation between humans and other species illuminates conserved areas that are critical for normal function of gene products.

A variety of types of DNA variation are recognized, including single nucleotide changes, insertions or deletions of a few to many nucleotides, copy number variants, and structural variants.

Interpretation of the significance of an observed variant in DNA sequence may require consideration of its location, frequency in the population, inheritance in a family, and specific effect on the resulting gene product.

GENOME ANNOTATION

A consensus sequence of the human genome is only the first step in furthering our understanding of normal biological functions and how mutations lead to abnormal functions that cause disease. The Human Genome Project has now matured into a number of important basic and applied research areas: *(i)* acquiring a comprehensive catalogue of human variation and the impact of such variation on phenotype, including disorders of human development; *(ii)* comparing the genomes of humans with those of other organisms, including model organisms and human ancestors; and *(iii)* learning how to interpret all the sequence elements within the genome, not just the codons. Even now, over a dozen years after "completion" of the human genome sequence, a complete, accurate, and single contiguous stretch representing a reference human haploid genome is still being constructed, and updated versions of the genome sequence continue to be released. As described below, the greatest challenges to completing the human genome sequence are posed by regions that contain segmental duplications of nearly the same sequence.[1]

HUMAN VARIATION

The first publicly available human genome sequence was constructed from a small number of individuals. Nonetheless, the "reference" human genome sequence available at public websites is a consensus, composite haploid sequence and not the sequence of any one individual. It is neither a "normal" genome nor a "control" genome; instead, it serves as a reference, providing a universally available sequence against which the genomes of other individuals, as well as other species, can be compared and any differences ("variants") determined. Even before the original human genome sequence was completed, the need to discover as broad a range of human variations as possible in populations from around the world was clearly recognized to be essential if we were to begin to understand how variations in the genome lead to differences in phenotypic variations in traits and disease susceptibilities. The first concerted efforts to catalogue human genetic diversity after completion of the Human Genome Project was the dbSNP (database of single nucleotide polymorphisms) project, followed by the 1000 Genomes project. These initial catalogues that surveyed the extent of variation in both the coding and the noncoding portions of the genome have been supplemented enormously by more comprehensive sequencing efforts, which produced the NHLBI GO Exome Sequencing Project (ESP) and Exome Aggregation Consortium (EXaC). These projects focus on variation within exomes of hundreds of thousands of individuals and have made a vast number of variants and their frequency publicly available to researchers and clinicians worldwide.

Variants can be classified as rare or common (polymorphic) according to their *frequency* in a population under study. Any variant present in >2% of the population (*i.e.*, constituting >1% of the alleles at any locus in that population) is arbitrarily designated as polymorphic. With the advent of comprehensive DNA sequencing of larger and larger numbers of people from across the globe, it has become clear that most variants (85%) at any one base pair (bp) or segment of the genome have allele frequencies substantially below the 1% cutoff for being a polymorphism and are, instead, considered rare, sometimes restricted to a single ethnic group or even a single kindred.[2]

DNA variation can also be classified according to the *type* of DNA change seen in the variant (Table 33.1). Single nucleotide variants (SNVs), insertion/deletion variants (indels), copy number variants (CNVs), and structural variants (SVs) can have different

TABLE 33.1 **Types of DNA Variation**

Type	Description	Ability to Detect	Frequency in an Individual
Single nucleotide variants (SNVs)	A sequence change where, compared with a reference sequence, one nucleotide is substituted for another nucleotide	High	3–4 000 000
Deletion or duplication variants (Del/Dups)	A sequence change involving between 2 and ~1000 base pairs (bp) in which reference nucleotides are either not present (deleted) or are duplicated and inserted directly 3′ of the original copy of the sequence	Moderate	1 000 000?
Insertion/deletion variants (Indels)	A sequence change involving between 2 and ~1000 bp in which one or more reference nucleotides are replaced by one or more other nucleotides and is not an SNV or SV	Moderate	
Copy number variants (CNVs)	Del/Dup or Indel arbitrarily set as larger than ~1000 bp	Difficult	
Structural Variants (SVs)		Very difficult	

From Human Variome Society, Sequence Variant Nomenclature. 2016. Available at: http://varnomen.hgvs.org.

consequences, depending on their location and the number and identity of nucleotides affected.

The simplest and most common of all variants are SNVs, in which one nucleotide in the reference sequence is substituted by another. A locus characterized by an SNV usually has only two alleles, corresponding to the more common (major allele) and less common (minor allele) bases found at that particular location in the genome, although, theoretically, four alleles at any one base position are possible. SNVs are observed on average once every 1000 bp in the genome but are not distributed evenly throughout and are most often not equally frequent throughout all populations. Most SNVs are found in the 98% of genomic sequence that is not within exons, including the ~20% of the genome inside genes, in introns, with the remaining ~80% between genes. Nonetheless, many SNVs are in coding portions of genes and other known functional elements in the genome. Just under half of these do not alter the predicted amino acid sequence of the encoded protein and are thus termed *synonymous*, whereas the remainder do alter the amino acid sequence and are said to be *nonsynonymous*. Other SNVs introduce or change a stop codon, and yet others alter a known splice site; such SNVs are expected to have significant functional consequences.

A second general class of variation is the result of insertion and/or deletion of reference sequence ranging from 1 up to an arbitrary cutoff of ~300–1000 bp, that is, a variant size that can be detected by the most commonly used next-generation sequencing (NGS) technology. When reference nucleotides are simply deleted or duplicated, the variant is referred to as a "del/dup." When reference sequence has some nucleotides deleted and *replaced* by other inserted sequence, the variant is referred to as an "indel." Approximately half of all del/dups are referred to as "simple" because they have only two alleles—that is, the presence or absence of the inserted or deleted segment. Each individual is known to carry hundreds of thousands of indels, but this estimate is suspected to be too low, and a corrected estimate for indels per individual may be upward of a million.[3] Up to 30–40% of actual indels are likely to be missing from catalogues because of the technical difficulties of distinguishing many indels from sequencing errors using the current sequencing technologies.[4]

Some del/dup variants are multiallelic because of variable numbers of the identical segment of DNA inserted in tandem at a particular location, thereby constituting what is referred to as a *microsatellite* (also referred to as a *short tandem repeat* [STR]). Microsatellites are segments of DNA composed of units of 2, 3, or 4 nucleotides, such as TGTG …TG, CAACAA …CAA, or AAATAAAT … AAAT. The units are repeated between two and a few dozen times at a particular site in the genome. Some variation in microsatellite length is common enough to constitute a *polymorphism*, defined as an allele or alleles other than the reference sequence, found in ≥2% of the population. Many tens of thousands of polymorphic microsatellites are known to exist throughout the human genome. There are, however, STR DNA segments present within exons or splice junctions of >90% of genes associated with human disease. They rarely, but famously, can expand to become hundreds or thousands of nucleotides long, thereby causing such human disorders as fragile X syndrome or Huntington disease. Even without expansion, STRs within exons, some of which may be as small as 9–25 bp, have an outsized impact on the frequency of human disease, since they confer a five- to sixfold increase in the frequency of rare disease-causing indel mutations compared with neighboring exon sequences that do not contain an STR.[5]

One subclass of indel variants arises from mobile elements. Nearly half of the human genome sequence consists of families of repetitive elements dispersed throughout the genome, of which the two most common are the *Alu* (a short interspersed nuclear element [or SINE]-usually about 300 bp) and LINE (long interspersed nuclear element) families of repeats. Although most of the copies of these repeats are stationary, some of them contribute to human genetic diversity through *retrotransposition*, a process that involves insertion of a DNA segment generated through transcription of an *Alu* or LINE element into an RNA that is then reverse-transcribed into a DNA sequence that is inserted into the genomic DNA. Each mobile element indel consists of two alleles, one with and one without the inserted mobile element. Mobile element polymorphisms are found on all human chromosomes; although most are found in nongenic regions of the genome, a small proportion exist within genes. At least 5000 of them are known to be frequent enough to be polymorphisms, and insertion frequencies of >10% occur in various populations; other mobile elements are rare and have been implicated in causing insertional mutations in human disease genes. As with other indels, difficulties with sequencing, in particular the challenges posed to NGS by repetitive DNA, result in underestimation of *Alu* and LINE indels throughout the genome.[3]

Another important type of human variation includes CNVs. These consist of variations in the number of copies of segments of the genome that are larger than those involved in indels and range in size from 1000 bp to many hundreds of kilobase (kb) pairs. Variants larger than 500 kb are found in 5–10% of individuals in the general population, whereas variants encompassing more than a million bases (1 Mb) are found in 1–2%.[6] Smaller CNVs, in particular, may have only two alleles (*i.e.,* the presence or absence of a segment), similar to indels. Larger CNVs tend to have multiple alleles because of the presence of different numbers of copies of a segment of DNA in tandem. The content of any two human genomes can differ by as much as 50–100 Mb as a result of copy number differences at CNV loci, compared with some 3–4 Mb of differences because of SNVs and indels. Thus simply on the basis of the number of nucleotides involved in the variation, CNVs represent vastly more human variation than do SNVs or indels.

Not only are CNVs a very significant contributor to human variation and disease,[7] the areas of the genome where they are found are often the sites of segmental duplications,[8] some of the most difficult regions in which to develop an accurate reference sequence. Segmental duplications are ubiquitous throughout the genome. It is difficult to construct an accurate contiguous genome sequence in regions of segmental duplications because the most widely used massively parallel sequencing technologies generate only short reads of a few hundred base pairs and therefore may not be able to differentiate between the sequences of nearly identical DNA segments using unique DNA that flanks them. The assembly algorithms, which always try to create the simplest assembly possible, may consider two nearly identical segments of DNA that have been duplicated in the genome, but differ at a few locations in their sequence, to be the same sequence and collapse them into a single segment of DNA. The fixed sequence differences between the duplicated segments are incorrectly interpreted as polymorphic differences between two alleles of a nonduplicated locus instead of being real sequence differences between different copies of duplicated DNA. The fact that there is a segmental duplication is therefore overlooked. Conversely, the sequence of a unique DNA segment with many SNVs compared with the reference may be incorrectly interpreted as having come from distinct but highly similar duplicated segments, leading to the supposition that there are paralogues or duplicated segments. As a result, segmental duplications are liable to both false-negative and false-positive representations in the human genome assembly. CNVs are particularly difficult to detect and interpret if the number of segments in the reference is incorrect or ambiguous.

Many CNVs include genes, from one up to several dozens; thus CNVs are frequently implicated in diseases that result from altered gene dosage.[7] One well-known human immunological multisystem disorder, DiGeorge syndrome, is caused by a deletion CNV that occurs *de novo* in about 1 in 5000 individuals. When a CNV is more frequent, it represents a background of common polymorphism that must be understood if to properly interpret alterations in copy number observed in patients. As with all DNA variations, the significance of different CNV alleles in health and in disease susceptibility is the subject of intensive investigation.

The most common SVs in the human genome sequence are inversions, which differ in size from a few base pairs to large regions of the genome (up to several Mb) that can be present in either of two orientations in the genomes of different individuals.[8] As with CNVs, inversions are usually flanked by regions of sequence homology, which indicates that recombination of these homologous regions is likely involved in the origin of the inversions.[7,8] Most inversions do not involve gain or loss of DNA; in this case, the two alleles corresponding to the two orientations can achieve substantial frequencies in the general population. Such inversions can, however, cause significant gains or losses of DNA in the offspring of inversion carriers because of aberrant recombination during meiosis, leading to serious syndromes brought about by chromosomal imbalance. Furthermore, if inversions interfere with normal gene expression by disrupting a gene or altering the physical relationship between a gene and its regulatory elements, it may be detrimental to health.[9,10]

CLINICAL IMPACT OF HUMAN VARIATION

One of the greatest challenges facing human geneticists is linking variation to phenotype.[11] The significance for the health of the vast majority of variants of any type is unknown, and yet this knowledge is essential if we are to apply genomics to clinical care. The impact of variants ranges from completely benign to highly pathogenic, the latter causing devastating disorders of the immune system that may occur as new mutation dominants or as autosomal recessively inherited syndromes. Even common polymorphic variants may affect health or longevity, although their being common means that they are likely to produce a relatively subtle alteration of disease susceptibility rather than directly cause a serious illness.[12] Working out the functional impact of human variation will occupy genomics researchers for many years to come. An essential component of this work is to make databases of genetic variants and their impact on human health available to the research and clinical communities, as is being done with the ClinVar database hosted by the National Library of Medicine in the United States and the Leiden Open-source Variation Database (LOVD).[13-15]

COMPARATIVE GENOMICS

Evolution at work is nowhere better illustrated than in the field of comparative genomics, which deals with similarities in the sequence, structure, and chromosomal location of genes between different species whose evolutionary paths diverged from a few years ago to many hundreds of millions of years ago. Direct sequence comparison has revealed that an enormous number of human proteins have orthologues (genes derived from a common ancestor) in other organisms, ranging from 87% in chimps to 79% in mice, 63% in zebrafish, 39% in the fruit fly (*Drosophila melanogaster*), and 31% in the nematode (*Caenorhabditis elegans*) (Fig. 33.1). The study of the human genome and the genomic basis for human disease has benefited from studies done in other organisms, particularly mice, in which many decades of inbreeding and gene manipulation have permitted study of the genes and underlying mechanisms by which mutations responsible for immune defects lead to phenotypes. Mouse models of immunity have been highly relevant and informative for humans.

However, orthologous genes may serve different functions in different species; therefore one cannot assume that a disease-causing mutation in humans will cause a similar defect when its orthologue is similarly mutated in mice, and vice versa. For example, the V(D)J recombination activating genes *RAG1* and *RAG2* in humans and *Rag1* and *Rag2* in mice appear to have identical functions, with knockout mice showing the same inability to recombine T- and B-lymphocyte antigen receptor genes as humans with RAG1 or RAG2 deficient severe combined immunodeficiency (SCID; Ch. 35).[16] As a result, the lymphocyte profile in both species is absent T and B cells with normal natural killer (NK) cells, referred to as $T^-B^-NK^+$ SCID. A contrasting situation occurs in other SCID genotypes. For example, humans lacking the common γ chain (γc) of receptors for interleukin-2 (IL-2) and other cytokines, caused by mutations in the X-linked gene *IL2RG* have SCID in which T and NK cells are absent but nonfunctional B cells are present in normal to high numbers, $T^-B^+NK^-$ SCID. Mice with the orthologous gene *Il2rg* mutated or removed can make T cells but have no B cells, which gives them a $T^+B^-NK^-$ phenotype.[17]

Genes other than *IL2RG* alone must be responsible for this difference between species. Such genes, known as *modifiers,* have not yet been identified. Notably, different strains of mice can also have important phenotypic differences in the presence of a single gene mutation under study; some strains, for example, nonobese diabetic (NOD) and Murphy Roths Large (MRL) mice are highly prone to developing autoimmunity, while others such as C57BL/6 (B6) are resistant.[18]

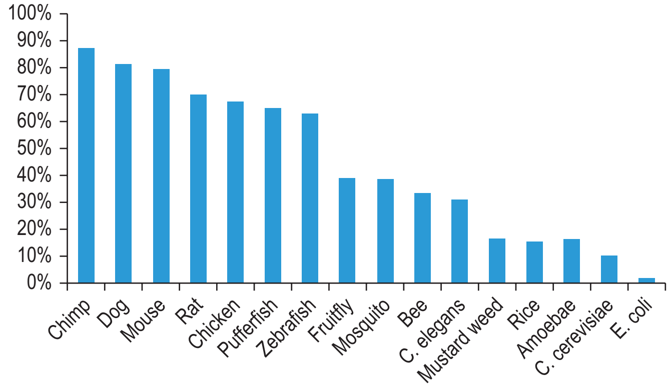


FIG 33.1 Homology Between Human Genes and Genes in Multiple Other Organisms. Comparisons were made by calculating gene homology or similarity between reference protein sequences from human and another species, with homology considered present when the probability of an interspecies match by chance alone was $<10^{-30}$. (Reproduced with permission from Nussbaum R. Human genomics and development. In: Epstein CJ, Erickson RP, Wynshaw-Boris A, editors. Inborn errors of development. New York: Oxford University Press; 2008).

FUNCTIONAL GENOMICS

Immediately after the Human Genome Project was declared to have been completed in 2003, an important follow-on project was launched to identify the functional segments of DNA, particularly portions in the 98–99% lying outside of the coding exons of genes, for which there was no simple sequence code that could be understood the way the triplet codon code can be read. This project, termed ENCODE (for "encyclopedia of DNA elements"), began as a pilot project to identify functional elements in 1% of the genome, but it was quickly expanded to include the entire human genome as well as the genomes of model organisms.[19] Before ENCODE, estimates were that 3–8% of the human genome had some role in function, given that this fraction of the genome appeared to be highly conserved among species with only very limited variation. This estimate was clearly too low, as it did not take into account rapidly evolving functional elements or those restricted to particular evolutionary lineages. This estimate also did not include segments of DNA that were too small to show conservation with statistical significance, nor did it include the functional elements in repetitive DNA that are not reliably scored as being evolutionarily conserved.

Since the same genome is present but functions differently in different cells of an individual, ENCODE used a number of different tissues for its studies (Fig. 33.2). A comprehensive catalogue of every segment of DNA that is transcribed into RNA in any tissue, including all splice isoforms, was required, including RNAs that code for a protein or are noncoding and function in gene regulation. ENCODE is analyzing not only total whole-cell RNAs but also those located in the nucleus or cytosol because subcellular localization of RNAs is important in how RNA is processed and functions. Other assays for functionality of segments of DNA include biochemical evidence, such as identifying

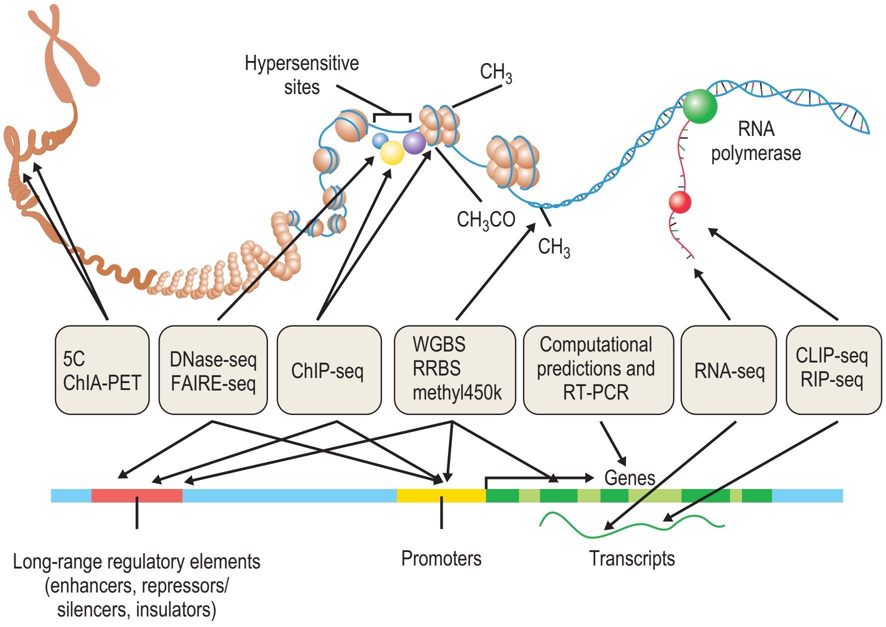


FIG 33.2 Diagram of Various Transcriptional and Chromatin Assays Being Used in the ENCODE Project. The assays can be used to identify regions of the genome involved in the regulation of gene expression. (With permission from https://www.encodeproject.org.)

(i) segments of DNA located in chromatin loops that allow chromatin–chromatin interaction; *(ii)* regions of open chromatin, which are accessible to transcription; *(iii)* motifs that bind transcription factors; *(iv)* regions that are associated with histones that have been modified to either promote or suppress transcription; and *(v)* regions with differential methylation of cytosine residues in different tissues, with methylation being associated with inactivity.

Now, approximately 15 years into the project, ENCODE has created a large database of more than 4 million functional regions of the genome. Mostly on the basis of biochemical evidence, it has been estimated that 80% of the genome *may* be of functional significance,[20] but this estimate is probably too high and has been strongly criticized. For example, evidence of transcription of a segment does not necessarily mean that the transcript plays any functional role, nor is it likely to be true that the millions of vestigial repetitive DNA sequences that are undergoing mutation without any apparent selection could have important functional roles.[21] A more stringent threshold for ascribing a functional role to DNA segments (*i.e.,* direct effect on gene expression and phenotype of at least one human cell type) reduces that fraction to the range of 10–20%. Much more research will be needed before the ENCODE project delivers its final assessment of the fraction of the human genome that plays a role in gene regulation—that is, in how different cells use their genomes.[22]

APPLYING HUMAN GENOMICS TO UNDERSTANDING DISORDERS OF THE HUMAN IMMUNE SYSTEM

Genome research has provided geneticists with a catalogue of all known human genes, knowledge of their location and structure, and an ever-growing list of variants in DNA sequences found among individuals in different populations. In the past, geneticists followed two approaches to identifying the genetic basis for human disorders. The first approach, linkage analysis, is *family based* (Fig. 33.3). Linkage analysis takes advantage of family pedigrees to follow the inheritance of a disease among family members and test a few hundred DNA variants distributed throughout the genome for consistent, repeated coinheritance, or segregation, with the disease. A demonstration of significant coinheritance with a *variant* or *variants* located in a particular region of the genome indicates that the disease-causing mutation is also located in a gene within or near this region. The variants showing coinheritance with the disease are usually not the variants responsible for the disorder. Marker variants need, however, to be located close enough that recombination between the marker and the gene mutation responsible for the disease is sufficiently rare that cosegregation is observed over a few generations.

The second approach, genome-wide association analysis (GWAS), is *population based.* A sample of affected individuals, or "cases," taken from the population, is chosen along with a matched set of unaffected "control" individuals from the same population. Then, a large number, in the order of a million or more, variants are examined individually for an increased or decreased frequency of cases compared with controls. The alleles used to *test* for association need not be the actual variants functionally *responsible* for the disease association—in fact, it is highly unlikely this would be the case. Instead, GWAS, like linkage, depends on the vast number of marker alleles that can be tested for association to be located close enough to the alleles functionally responsible for association to have been carried along, on the same chromosome in a haplotype block, through many generations (Fig. 33.4). Association analysis does not require pedigrees and is particularly useful for complex diseases that do not show strict mendelian inheritance.

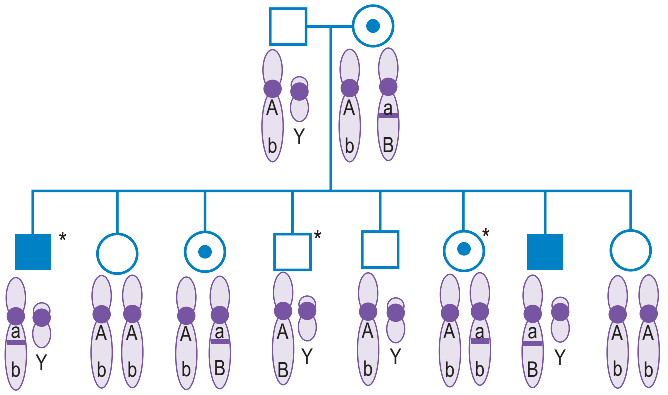


FIG 33.3 An Example of Family-Based Linkage Analysis as a Means to Identify a Disease Gene. The pedigree of a family with X-linked severe combined immunodeficiency (SCID) caused by an *IL2RG* mutation in the proximal long arm of the X chromosome *(marked with black bar)* is shown. Affected males have filled in square; female carriers are designated by dot in the center of their symbols. Two loci, one with alleles A and a, and the other with alleles B and b, are shown in the proximal and distal long arm of the X (Xq). There are no recombinations during female meiosis between allele a and the *IL2RG* mutation in any of the eight children, while three children *(marked with asterisk)* show a recombination with the locus on the distal Xq. The father's X chromosome is passed on without any recombinations to each of his daughters and is shown on the left of each pair of X chromosomes in the daughters.

Linkage analysis and association studies have limitations in investigating the genetic basis for human immunological disorders. Linkage analysis is problematic if the disorder is a rare autosomal recessive condition such that there are not enough families with two carrier parents to enable such a study; nonetheless, the increased frequency of consanguineous matings in some populations has been utilized to overcome this limitation.[23,24] Another challenge with linkage analysis is if the disorder is genetically lethal so that it is never inherited and always occurs sporadically as a result of a new mutation. Detecting an association in a case-control study is also a problem when the frequency of any particular allele associated with the disease is too low among the cases to give a detectable association. For example, if the disorder arises from different, independent mutations and if these mutations are found on many different haplotypes in affected individuals, it may be very difficult to establish a significant association with any one variant.

When linkage and association studies are not possible, as in very rare mendelian disorders, a third approach is now available—genome sequencing. Vastly improved methods of DNA sequencing have cut the cost of sequencing six orders of magnitude over what was spent generating the Human Genome Project's reference sequence, opening up new possibilities to discover the genes and mutations responsible for rare mendelian disorders. One can generate a whole-genome sequence (WGS) or, in what has often proven a cost-effective compromise, a sequence of only about

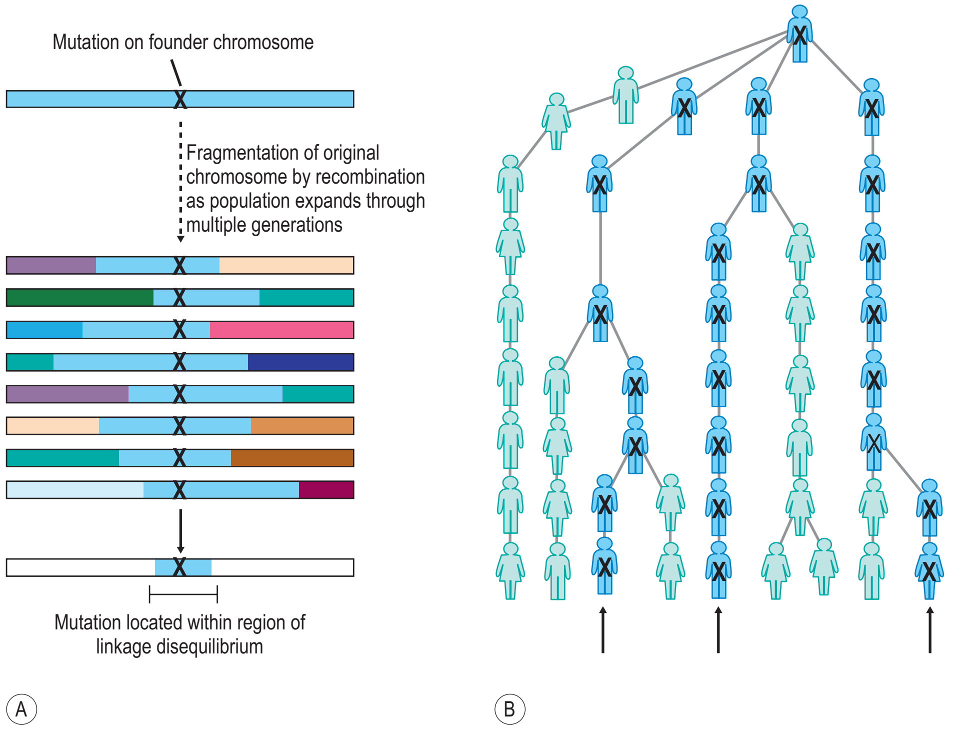

Mutation on founder chromosome
X
Fragmentation of original chromosome by recombination as population expands through multiple generations
Mutation located within region of linkage disequilibrium
A
B

FIG 33.4 An Example of a Population-Based Genome-Wide Association Analysis——As a Means to Identify a Disease Gene. (A) A mutation (X) that predisposes to a disease first occurs on a chromosome with a certain set of alleles at polymorphic loci along that chromosome *(symbolized by the blue color).* With each generation, meiotic recombination exchanges the alleles that were initially present at polymorphic loci on the "blue chromosome" for other alleles present on homologous chromosomes *(symbolized by other colors).* Over many generations, the only alleles that remain associated with the mutation are those at loci so close to the mutant locus that recombination between them is very rare. These alleles constitute a disease-associated haplotype. (B) Affected individuals in the current generation (*arrows*) carry the mutation (**X**) are enriched for the disease-associated haplotype (*individuals in blue*) compared with unaffected individuals. Depending on the age of the mutation and other population genetic factors, a disease-associated haplotype ordinarily spans a region of DNA of a few kilobases to a few hundred kilobases. (From Nussbaum R, McInnes RR, Willard HF. Thompson and Thompson Genetics in medicine. 8th ed. Toronto, Canada: Elsevier Canada; 2016: 177, Fig. 10.8, with permission.)

2% of the genome, the part containing the exons of genes, referred to as a whole exome sequence (WES).

As an example of what is now possible, suppose that there is a family "trio" consisting of a child affected with a rare immunodeficiency and his or her parents. WES is carried out on the child and the parents, yielding typically over 4 million SNV, indel, and CNV differences in the child compared with the human genome reference sequence. Which of these variants would be responsible for the disease? The extraction of useful information from such a massive amount of data relies on creating a variant filtering scheme based on a variety of reasonable assumptions about likely responsible explanations for the disease.

One example of a filtering scheme that can be used to sort through these variants is shown in Fig. 33.5, in which exome sequence was performed for two parents and their offspring, two affected and one unaffected.

1. *Location with respect to protein-coding genes.* Keep variants that are within or near exons of protein-coding genes and discard variants deep within introns or intergenic regions. It is possible, of course, that the responsible mutation might lie in a noncoding RNA gene (*e.g., RMRP,* the gene mutated in the immunodeficiency and dwarfing syndrome cartilage hair hypoplasia). However, these are currently more difficult to assess, and thus, as a simplifying assumption, it is reasonable to focus initially on protein-coding genes.
2. *Population frequency.* Keep rare variants from Step 1 and discard common variants with allele frequencies <0.05 (or some other arbitrary number between 0.01 and 0.1) because common variants are highly unlikely to be responsible for a disease whose population prevalence is much less than the q^2 predicted by the Hardy- Weinberg equilibrium.
3. *Deleterious nature of the mutation.* Keep variants from Step 2 that cause nonsense or nonsynonymous changes in codons within exons, cause frame-shift mutations, or alter highly conserved splice sites. Discard synonymous changes that have no predicted effect on protein function (unless there is reason to suspect that they influence splicing or expression, such as the last nucleotide of an exon, which is typically "G").
4. *Consistency with likely inheritance pattern.* If the disorder is considered most likely to be autosomal recessive, keep any variants from Step 3 for which an affected child has 2 variants in the same gene and each parent has one of the variants.

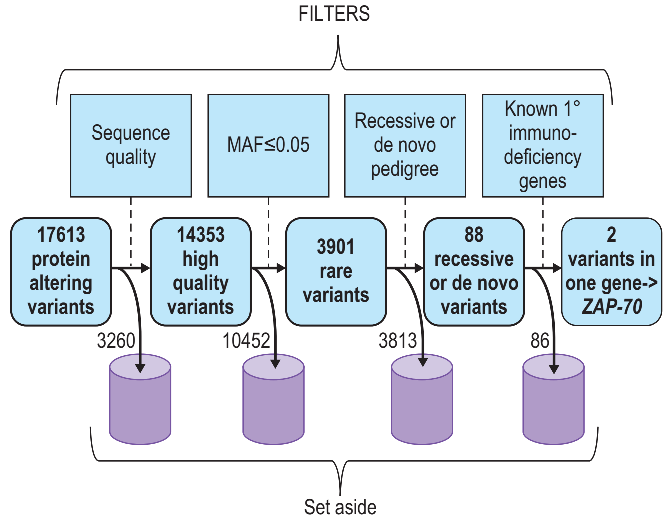

FILTERS
Sequence quality
MAF≤0.05
Recessive or de novo pedigree
Known 1° immuno-deficiency genes
17613 protein altering variants
14353 high quality variants
3901 rare variants
88 recessive or de novo variants
2 variants in one gene-> ZAP-70
3260
10452
3813
86
Set aside

FIG 33.5 **Variant Filtering Can Be Used to Sort Through Whole-Exome Sequence Analysis of a Familial Disease.** Protein-altering variants detected in the affected children, their unaffected sibling, and parents were first filtered by sequencing quality and allele frequency to yield a set of 88 rare, damaging or nonsynonymous variants that fit an autosomal recessive inheritance pattern and were then candidates for causing the autoimmune disorder. This group was then further analyzed for genes involved in immune function, which brought two variants in *ZAP-70,* for which the affected children were both compound heterozygotes, to the top of the list as possible causative mutations. (Adapted from Chan et al. A novel human autoimmune syndrome caused by combined hypomorphic and activating mutations in ZAP-70. J Exp Med 2016;213:155–65, with permission.)

The child need not be homozygous for the same deleterious variant but could be a compound heterozygote for two different deleterious mutations in the same gene.

In the example in Fig. 33.5, the two affected children were compound heterozygous for variants, with each parent and the unaffected sister carrying one of the variants. If there were consanguinity in the parents, the candidate genes and variants might be further filtered by requiring that the child be a true homozygote for the same mutation derived from a single common ancestor. If the autosomal recessive mode of inheritance is correct, then the parents should both be heterozygous for the variants. The disease-causing variant in a male could also be X-linked, in which case any variant found in an X-linked gene for which the mother is a heterozygote would be a candidate.

For either the autosomal or X-linked case, the disorder could be the result of a new dominant mutation; in this case, keep variants from Step 3 that are *de novo* changes in the child and are not present in either parent.

In the end, millions of variants can often be filtered down to a handful of SNVs, indels, or CNVs affecting a reasonable number of genes. Once the filtering reduces the number of genes and alleles to a manageable number, they can be assessed individually for other characteristics. Do any of the genes have a known function or tissue expression pattern that would be expected if it were the potential disease gene? Is the gene already implicated in other disease phenotypes, or does it have a role in pathways in which mutations are known to cause similar or different phenotypes? Finally, is the same gene mutated in other patients or in an animal model with the disease? Finding mutations in one of the candidate genes also in other patients would confirm that this gene was the responsible gene trio under study. In some cases, one gene remaining on the list in Step 4 may become a leading candidate because its involvement makes biological or genetic sense or it is known to be mutated in other affected individuals. In other cases, however, the gene responsible may turn out to be entirely unanticipated on biological grounds or may not be mutated in other affected individuals because of locus heterogeneity (*i.e.,* mutations in other as yet undiscovered genes can cause a similar disease).

Such variant assessments require extensive use of public genomic databases and software tools. These include the human genome reference sequence, databases of variants with their allelic frequencies, software that assesses how deleterious an amino acid substitution might be to gene function, collections of known disease-causing mutations, and databases of functional networks and biological pathways. The enormous expansion of this information over the past few years has played a crucial role in facilitating gene discovery of rare mendelian disorders.

In a recent report of the clinical use of WES in rare disorders,[25] 2000 individuals with a variety of disorders that had escaped diagnosis despite thorough conventional clinical evaluations underwent WES. A likely causative mutation or pair of mutations was found in 504 (25.2%). Of these mutations (of which many were severe developmental defects), half were *de novo,* not present in either parent. Also of interest was that 23 (4.6%) of these 504 patients in whom a diagnosis was made through WES had two different genetic disorders, resulting in a combined phenotype, most likely obscuring the diagnosis of either of the individual disorders had they been present alone.

In some cases, the difficulty in making a diagnosis was the result of there being a mutation in a gene, previously unknown to cause human disease, but implicated by finding mutations that fit the inheritance pattern (*de novo* dominant or autosomal recessive), and then later confirmed by finding other mutations in phenotypically similar patients. In others, the gene involved had been previously described, but not in association with the phenotype for which the patient was undergoing diagnostic exome analysis. For example, two children affected with an undiagnosed autoimmune syndrome consisting of early-onset bullous pemphigoid, nephritis, autoimmune antifactor VIII hemophilia, and inflammatory bowel disease. WES was performed for both children, their unaffected sister, and both unaffected parents.[26] Eighty-eight genes were found in which the two affected children were compound heterozygotes for two different rare mutations, whereas the unaffected sister and both parents were carriers for only one of the two mutations found in the children. Of these, only one gene was known to have any immunological function: *ZAP70,* previously known to be mutated in combined immunodeficiency but not with strictly autoimmune disease of this type (see Fig. 33.5). Both affected children were compound heterozygotes for two missense mutations, p.R192W and p.R360P, in *ZAP70.* The mother carried the R192W allele, whereas the father and unaffected sister carried the R360P allele. Neither mutation had been reported previously, nor were they near mutations in the same gene that had been previously associated with human disease. Functional studies revealed the p.R192W had reduced activity, whereas the R360P encoded a modestly hyperactive protein because of disruption of an autoinhibitory mechanism. The combination of these alleles, as proven in laboratory studies, fully explained the autoimmune phenotype.

ON THE HORIZON

- As the costs of deep sequencing fall, examination of whole-genome sequence should allow for examination of noncoding regulatory regions.
- Data storage and computer speed will need to be improved to handle 50-fold more data from each individual.
- New analytical paradigms for the analysis of variation in noncoding DNA and the interpretation of DNA variants that are not clearly associated with deleterious effects will be required.
- Proving causality of new gene variants in primary immunodeficiencies will depend on future cataloguing of deleterious mutations in each gene and on molecular studies of the direct effects of the variants.

Since the application of WGS or WES to rare mendelian disorders was first described in 2009, hundreds of such disorders have been studied and the causative mutations found in nearly 500 previously unrecognized disease genes.[27] The pace of discovery will only increase, and we can anticipate that the number of known mendelian disorders, including those characterized by primary immunodeficiency and/or autoimmunity, will grow over coming years. It has been predicted that the genetic basis for every human single gene disorder will be elucidated in the near future. However, the genetic basis for complex disorders that show increased familial occurrence, and yet are not single gene disorders and do not follow mendelian inheritance, remains elusive and will require another leap in technology and genetic analytical tools to be fully unraveled. The data being generated through human and model organism genomics will certainly aid in the elucidation of these more challenging disorders of the human immune system. New tools available to immunologists and geneticists will open great opportunities for discovery and understanding of genetic disorders of immunity which, in turn, will deepen our knowledge of immune system networks.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Steinberg KM, et al. Single haplotype assembly of the human genome from a hydatidiform mole. Genome Res 2014;24:2066–76.
2. Gonzaga-Jauregui C, Lupski JR, Gibbs RA. Human genome sequencing in health and disease. Annu Rev Med 2012;63:35–61.
3. Jiang Y, Turinsky AL, Brudno M. The missing indels: an estimate of indel variation in a human genome and analysis of factors that impede detection. Nucleic Acids Res 2015;43:7217–28.
4. Huddleston J, Eichler EE. An incomplete understanding of human genetic variation. Genetics 2016;202:1251–4.
5. Madsen BE, Villesen P, Wiuf C. Short tandem repeats in human exons: a target for disease mutations. BMC Genomics 2008;9:410.
6. McCarroll SA, et al. Common deletion polymorphisms in the human genome. Nat Genet 2006;38:86–92.
7. Stankiewicz P, Lupski JR. Structural variation in the human genome and its role in disease. Annu Rev Med 2010;61:437–55.
8. Carvalho CM, Lupski JR. Mechanisms underlying structural variant formation in genomic disorders. Nat Rev Genet 2016;17:224–38.
9. Lakich D, et al. Inversions disrupting the factor VIII gene are a common cause of severe haemophilia A. Nat Genet 1993;5:236–41.
10. Puig M, et al. Human inversions and their functional consequences. Brief Funct Genomics 2015;14:369–79.
11. Amendola LM, et al. Performance of ACMG-AMP Variant-interpretation guidelines among nine laboratories in the Clinical Sequencing Exploratory Research Consortium. Am J Hum Genet 2016;99:247.
12. Tryka KA, et al. NCBI's Database of Genotypes and Phenotypes: dbGaP. Nucleic Acids Res 2014;42:D975–9.
13. Landrum MJ, et al. ClinVar: public archive of interpretations of clinically relevant variants. Nucleic Acids Res 2016;44:D862–8.
14. Harrison SM, et al. Using ClinVar as a resource to support variant interpretation. Curr Protoc Hum Genet 2016;89:8.16.1–23.
15. Fokkema IF, et al. LOVD v.2.0: the next generation in gene variant databases. Hum Mutat 2011;32:557–63.
16. Revy P, et al. The repair of DNA damages/modifications during the maturation of the immune system: lessons from human primary immunodeficiency disorders and animal models. Adv Immunol 2005;87:237–95.
17. Kovanen PE, Leonard WJ. Cytokines and immunodeficiency diseases: critical roles of the gamma(c)-dependent cytokines interleukins 2, 4, 7, 9, 15, and 21, and their signaling pathways. Immunol Rev 2004;202: 67–83.
18. Choi Y, Simon-Stoos K, Puck JM. Hypo-active variant of IL-2 and associated decreased T cell activation contribute to impaired apoptosis in autoimmune prone MRL mice. Eur J Immunol 2002;32:677–85.
19. Diehl AG, Boyle AP. Deciphering ENCODE. Trends Genet 2016;32: 238–49.
20. ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. Nature 2012;489:57–74.
21. Doolittle WF. Is junk DNA bunk? A critique of ENCODE. Proc Natl Acad Sci USA 2013;110:5294–300.
22. Eddy SR. The ENCODE project: missteps overshadowing a success. Curr Biol 2013;23:R259–61.
23. Al-Herz W. Primary immunodeficiency disorders in Kuwait: first report from Kuwait National Primary Immunodeficiency Registry (2004—2006). J Clin Immunol 2008;28:186–93.
24. Al-Herz W, et al. Parental consanguinity and the risk of primary immunodeficiency disorders: report from the Kuwait National Primary Immunodeficiency Disorders Registry. Int Arch Allergy Immunol 2011;154:76–80.
25. Yang Y, et al. Molecular findings among patients referred for clinical whole-exome sequencing. JAMA 2014;312:1870–9.
26. Chan AY, et al. A novel human autoimmune syndrome caused by combined hypomorphic and activating mutations in ZAP-70. J Exp Med 2016;213:155–65.
27. Chong JX, et al. The genetic basis of mendelian phenotypes: discoveries, challenges, and opportunities. Am J Hum Genet 2015;97:199–215.

MULTIPLE-CHOICE QUESTIONS

1. Which of the following variations in a gene expressed in T lymphocytes is MOST likely to be deleterious, capable of causing a human immune disorder if found in homozygosity:
   - **A.** A single nucleotide substitution AAA (lysine) ≥ TAA (stop) in an early exon of a gene required for T-cell receptor rearrangement
   - **B.** A single nucleotide substitution CTC (leucine) ≥ ATC (isoleucine) in a gene encoding a cytokine receptor
   - **C.** Deletion of 6 amino acids (CAACAA) in a short tandem repeat (STR) consisting of $(CAA)_{10}$ in the 3' untranslated portion of a gene expressed in all lymphocytes

2. True or False: A human inherited immunodeficiency caused by mutation of single gene can reliably be reproduced in the mouse by knockout of the orthologous murine gene.

3. DiGeorge syndrome is most commonly caused by:
   - **A.** A single nucleotide variation
   - **B.** A copy number variation
   - **C.** A retrotransposition
   - **D.** An indel mutation

34

Primary Antibody Deficiencies

Tracy Hwangpo, Harry W. Schroeder, Jr.

Primary antibody deficiency diseases are characterized by an inability of the humoral immune system to produce sufficient quantities of protective antibodies to properly protect the host from hazardous antigens.[1] This inability may be evident from birth, or it may manifest at a later age. In some cases, the deficiency may either resolve or worsen with time.[2] Many of these diseases are caused by mutations that alter the function of genes that regulate B-cell development or homeostasis. Others reflect mutations in the immunoglobulin (Ig) genes themselves. In some of the most common conditions, a genetic predisposition has been well documented, but the underlying defect remains unclear. The typical patient will present with a history of recurrent upper or lower respiratory infections and reduced serum concentrations of one or more classes of immunoglobulin (IgM, IgG, or IgA). However, patients with normal serum Ig levels may exhibit specific deficits in their ability to mount a protective response against certain antigens, and some patients with agammaglobulinemia can be remarkably asymptomatic. A classification of primary antibody deficiency diseases can be found in Table 34.1.

Primary immune deficiency disorders are the consequence of specific defects in B-cell development (Chapter 7). B-cell production begins in the fetal liver and shifts to bone marrow in the latter stages of fetal life (Fig. 34.1). As B cells mature, they leave the liver and bone marrow and migrate via blood into secondary lymphoid organs, including spleen, lymph nodes, and other peripheral and mucosal tissues. Contact with a polymeric cognate antigen, such as a polysaccharide, can activate the B cell and allow it to differentiate into an antibody-producing plasma cell. In contrast, the response to protein antigens, including toxins and viral proteins, requires T-cell help. In the germinal centers, B cells can replace an upstream heavy (H) chain constant domain with a downstream one (*e.g.*, μ to γ1), altering effector function (Chapter 15). They can also introduce mutations at a high level into the variable (V) domains, tailoring the antibody to the antigen, a process termed *affinity maturation.*

CLINICAL MANIFESTATIONS

Patients with antibody deficiencies most commonly present with a history of recurrent sinusitis, bronchitis, and pneumonia. Patients can also present with recurrent cellulitis, boils, gastrointestinal (GI) discomfort, myalgias, arthralgias, fatigue, and depression. The respiratory infections typically involve encapsulated bacterial pathogens, such as *Streptococcus pneumoniae* and *Haemophilus influenzae.* Protection against these bacteria requires production of antipolysaccharide antibodies, which does not require T-cell help. Because a similar susceptibility for infection is seen among patients that are deficient in neutrophil function or in the pivotal third component of complement (C3), all three of these arms of the host defense system should be evaluated in patients who suffer with recurrent bacterial infections.

CLINICAL PEARLS

Clinical Manifestations of Antibody Deficiency

- Recurrent bacterial infections
  - Early in untreated disease, infections are primarily caused by encapsulated pyogenic bacteria (*e.g., Streptococcus pneumoniae* and *Haemophilus influenzae* type b).
  - Later in untreated disease, damage to mucosal surfaces engenders susceptibility to staphylococci, nontypable *H. influenzae,* and gram-negative rods, as well.
- Recurrent viral infections
  - In most cases, viral infections are cleared normally, but protective immunity does not develop. For example, recurrent shingles can be a common symptom in untreated patients.
  - In some cases, patients may continue to excrete virus after resolution of their clinical symptoms.
- Increased prevalence of other immunological disorders
  - Paradoxical increased risk of antibody-mediated autoimmune disorders, such as idiopathic thrombocytopenia, autoimmune thyroiditis, systemic lupus erythematosus, and celiac disease.
  - Lymphoid hypertrophy
  - Increased risk of allergic disorders, especially among patients with IgA deficiency.

The clinical course of uncomplicated primary infections with viruses, such as varicella zoster or mumps virus, does not differ significantly from that of the normal host. However, patients with antibody deficiency have difficulty generating long-lasting immunity; thus chickenpox can repeatedly recur as shingles. The general rule is that T cells typically control established viral infections, whereas antibodies limit initial viral dissemination and cell entry, thereby preventing reinfection. As with all rules, there are exceptions. Patients with hypogammaglobulinemia can have difficulty clearing hepatitis B virus from the circulation, poliovirus from the gut, and enterovirus from the brain, leading to progressive and sometimes fatal outcomes.

Because sinopulmonary infections are also commonly seen in normal infants and children, in individuals with allergies, in smokers, and in patients with other diseases, such as cystic fibrosis, the threshold for an extensive evaluation for immunodeficiency can be a matter of clinical judgment. However, two or more episodes of bacterial pneumonia within a 5-year period, unexplained bronchiectasis, *H. influenzae* meningitis in an older child or adult, chronic otitis media in an adult, recurrent intestinal

TABLE 34.1 Primary Antibody Deficiencies

Disorder	Gene or Locus	Chromosome
IgA deficiency / Common variable immunodeficiency (CVID)	MHC	6p21.3
CVID1: ICOS deficiency (AR)	ICOS	2q33
CVID2/IgAD2: TACI deficiency (AD/AR)	TNFRSF13B	17p11.2
CVID3: CD19 deficiency (AR)	CD19	16p11.2
CVID4: BAFF-R (AR)	TNFRSF13C	22q13.1-q13.31
CVID5: CD20 deficiency (AR)	CD20	11q13
CVID6: CD81 deficiency (AR)	CD81	11p
CVID7: CD21 (AR)	CD21	1q32
CVID8: LRBA (AR)	LRBA	4q31.3
CVID9: PKCδ (AR)	PRKCD	3p21.31
CVID10: NF-κB2 (AD)	NFkB2	10q24
CVID11: IL-21 (AR)	IL21	4q27
CVID12: NF-κB1 (AD)	NFkB1	4q24
CTLA-4 (AD)	CTLA-4	2q33
TWEAK (AD)	TNFSF12	17p13
P13K (AD)	PIK3CD	1p36.2
PI3K regulatory subunit (AD)	PIK3R1	5q13.1
BLK	BLK	8p23-p22
IRF2BP2	IRF2BP2	1q42.3
IKAROS (AD)	IKZF1	7p12.2
Kabuki syndrome (AD)	KMT2D, KDM6A	12q13.12, Xp11.2
Transient hypogammaglobulinemia of infancy (THI)	Unknown	Unknown
X-linked agammaglobulinemia (XL)	BTK	Xq21.3-q22
X-linked agammaglobulinemia with growth hormone deficiency (XL)	BTK	Xq21.3-q22
Hyper-IgM syndrome		
HIGM1: X-linked hyper-IgM syndrome (XHM) (XL)	CD154 or CD40L	Xq26
HIGM2: Activation-induced cytidine deaminase deficiency (AR/AD)	AID	12p13
HIGM3: CD40 deficiency (AR)	CD40	20q12-q13.2
HIGM5: Uracil-DNA glycosylase (UNG) deficiency (AR)	UNG	12q23-q24.1
HIGM6: XHM with ectodermal dysplasia (XHM-ED) (XL)	NEMO	Xq28
Autosomal Agammaglobulinemia (AGM)		
AGM1: Immunoglobulin μ H-chain deficiency (AR)	IGHG1	14q32.33
AGM2: Surrogate light-chain deficiency (AR)	IGLL1/CD179B	22q11.21
AGM3: Ig-associated α(Igα) deficiency (AR)	CD79A	19q13.2
AGM4: BLNK deficiency (AR)	BLNK	10q23.2
AGM5: LRRC8 truncation (AD)	LRRC8	
AGM6: Ig-associated β (Igβ) deficiency (AR)	CD79B	17q23
AGM7: PI3K regulatory subunit (AR)	PIK3R1	5q13.1
AGM8: E47 transcription factor deficiency (AD)	TCF3	
Myelodysplasia with hypogammaglobulinemia		Monosomy 7
Thymoma with immunodeficiency (Good syndrome)		Trisomy 8
Selective IgG Subclass Deficiencies		
Immunoglobulin γ H-chain deficiencies (AR)	IGHG1	14q32.33
Selective κ light-chain deficiency (AR)	IGKC	2p12
Selective κ light-chain deficiency (AR)	IGLC	22q11.2
Antibody deficiency with normal serum immunoglobulin levels		
Vκ A2 deficiency (AR)	IGKV2D-29	2p12

Autosomal Dominant (AD), Autosomal Recessive (AR), X-linked (XL)

infections and diarrhea caused by *Giardia lamblia,* or a family history of immunodeficiency all warrant evaluation by an immunologist.

The purest forms of antibody deficiency result from mutations that prevent the expression or function of the pre–B-cell receptor (Chapter 4). For example, a function-loss mutation of μ heavy chain or components of the surrogate light chain [VpreB, λ14.1 (λ5)] have been found to affect only the B-cell lineage. However, most of the diseases associated with primary antibody deficiency involve more than one cell lineage. For example, *X-linked agammaglobulinemia* (XLA) is the product of loss-of-function mutations in Bruton tyrosine kinase (BTK). Patients with *X-linked hyper-IgM syndrome* (HIGM1) may also exhibit T-cell dysfunction, placing them at risk for infection with *Pneumocystis jiroveci.* Immune deficiency also appears to place patients at risk for autoimmunity, which is increased among patients with *IgA deficiency* (IgAD), *common variable immunodeficiency* (CVID), and HIGM1.

Clinical manifestations of the primary immunodeficiency can also be heavily influenced by the patient's past medical history. A delayed diagnosis and failure to treat infections aggressively can lead to permanent damage to the respiratory or GI mucosa, thus creating susceptibility to nontypable *H. influenzae,* staphylococci, *Pseudomonas* spp., and enteric bacteria.

PRINCIPLES OF DIAGNOSIS AND TREATMENT

Diagnostic Tests and Their Interpretation

Testing for immune deficiency should be performed when the patient has a history of repeated infections that exceeds

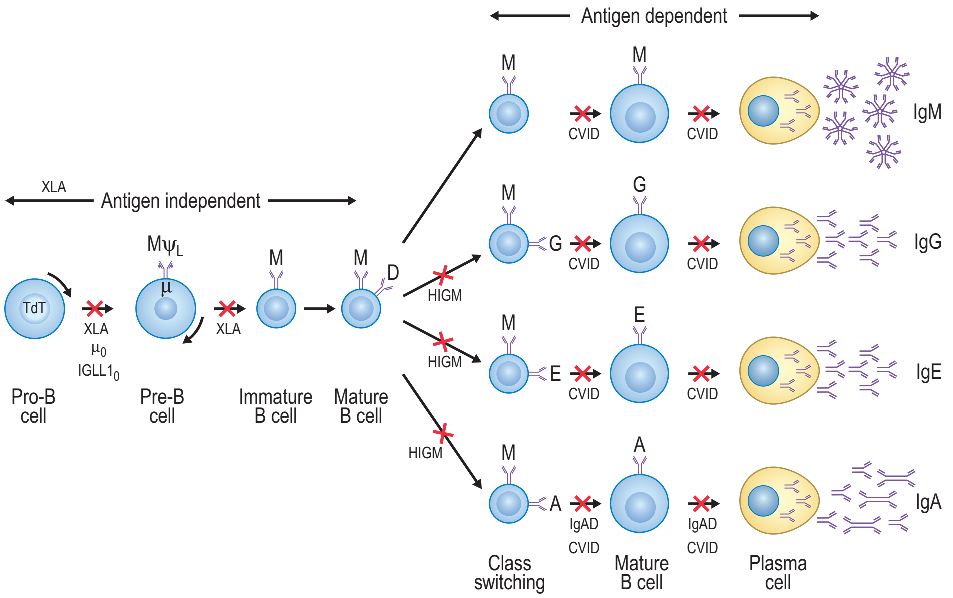


FIG 34.1 Defects in B-Cell Development Can Lead to Humoral Immune Deficiency. X-linked agammaglobulinemia (XLA) and autosomal agammaglobulinemias (AGMs) result from either a failure to generate a functional antibody receptor signaling complex or a failure to transmit signals from this complex. These failures obstruct production of immature B cells. Hyper-IgM syndrome (HIGM) results from either a failure to engage in proper cognate interactions with T cells or disruptions in the genes that permit class-switch recombination or somatic hypermutation. These failures inhibit class switching to IgG, IgA, and IgE, as well as limiting affinity maturation. Common variable immune deficiency (CVID), IgG subclass deficiencies (IgSD), selective IgA deficiency (IgAD), and transient hypogammaglobulinemia of infancy (THI) reflect a selective or generalized failure to progress from the immature B-cell stage to the plasma cell stage.

KEY CONCEPTS

Hypogammaglobulinemia and Antibody Deficiency

- The genetic mutations that underlie most primary antibody deficiencies tend to affect genes that play key roles in the regulation of lymphocyte development and homeostasis. Clinically significant antibody deficiency is not synonymous with hypogammaglobulinemia.
  - Serum immunoglobulin (Ig) concentrations vary widely with age.
  - Serum immunoglobulin levels vary with exposure to drugs (*e.g.*, steroids), infectious agents, and other environmental stressors.
  - A complete absence of IgG subclasses resulting from homozygous deletions of heavy-chain genes has been observed in healthy individuals.
  - Absence of a specific Vκ gene segment has been associated with an increased risk of *H. influenzae* infection in spite of normal serum Ig levels.

expectations for normal individuals, when an opportunistic pathogen or one of a low virulence is responsible for an infection, when a diagnosis of a disorder frequently associated with immunodeficiency has been made, or when there is a family history of primary immunodeficiency. Table 34.2 illustrates four levels of testing complexity and when such testing might commonly be undertaken.

Level I testing is both revealing and cost effective. It includes measuring serum Ig (IgM, IgG, and IgA), complement (50% hemolytic power of serum [CH_{50}] and complement components C3, C4, and mannose-binding lectin protein [MBL]), a complete blood count with differential (CBC/diff), and an erythrocyte sedimentation rate (ESR). Lymphopenia is found most often in disorders that affect the production or function of T cells (Chapter 35) but can also occur in patients with CVID. Congenital absence of an individual complement component will result in total absence of measurable complement-mediated hemolysis (Chapter 21). MBL deficiency increases susceptibility to respiratory infections.[3] The ESR is often, although not always, elevated in individuals with inflammatory disorders and can be useful in the evaluation of patients with a questionable or unclear history of recurrent or chronic infection.

KEY CONCEPTS

Testing for Immune Function

Testing for immune function should be performed:
- When the patient's history suggests a rate or severity of infection that exceeds normal expectations
- When the organism responsible for infection is of low virulence or is considered to be an opportunistic pathogen
- When there is a diagnosis of a genetic syndrome or disorder associated with immune deficiency either in the patient or in the patient's family

Interpretation of the significance of the quantitative Ig determinations requires appreciation of age-related changes in Ig concentrations (Fig. 34.2).[4,5] At the end of the second trimester of pregnancy, active transport of IgG across the placental barrier

TABLE 34.2 Laboratory Diagnosis of Primary Antibody Deficiency

Level	Test	Application(s)
I	CBC with differential Complement (CH$_{50}$, C3, C4, mannose-binding lectin [MBL] protein) Erythrocyte sedimentation rate (ESR) Quantitative serum immunoglobulin (Ig)M, IgG, and IgA levels	Primary screening tests
Ia	Urinalysis, 24-hour urine for protein Stool for α-1-antitrypsin	Symptoms suggest protein loss through kidneys or gastrointestinal (GI) tract
II	B-cell functional evaluation Quantitative IgG subclasses, IgE Natural or commonly acquired antibodies (isohemagglutinins, rubella, rubeola, tetanus) Response to immunization T cell–dependent antigens (tetanus) T cell–independent antigens (unconjugated pneumococcal vaccine, unconjugated *H. influenzae* B vaccine)	Level I normal but history suggests Antibody deficiency Better definition of a level I defect
III	Quantification of blood T-and B-cell subpopulations by immunofluorescence assays using monoclonal antibody markers T cells: CD3, CD4, CD8 B cells: CD19, CD20, CD21, Ig (μ, δ, κ, λ),	Panhypogammaglobulinemia or severely low IgM and IgA
IV	Disease-specific analysis Gene expression Gene sequencing	Gene-specific diagnosis Genetic counseling

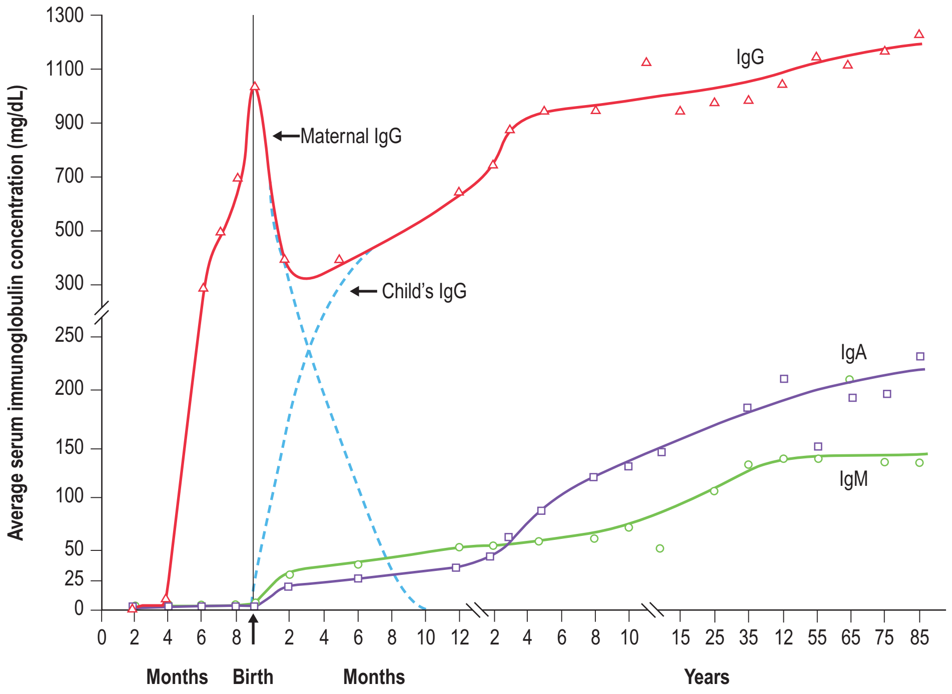


FIG 34.2 Age-Related Changes in the Serum Concentration of Immunoglobulins (Ig). Shown are average serum immunoglobulin concentrations of the major isotypes as a function of age.[4,5]

begins and, at birth, the infant's serum IgG concentration is typically 20–25% higher than that of the mother. Catabolism of maternal IgG coupled with the slow development of endogenous antibody function leads to a physiological nadir of serum IgG in infants of 4–6 months of age. In normal infants, this loss of maternal protection is often associated with the first appearance of otitis media or bronchitis. Thus the onset of sinopulmonary infections within the first 3 months of age should also raise the index of suspicion for immunodeficiency in the mother. After age 6 months, maternally-derived IgG has largely been lost, and IgG antibodies specific for diphtheria or tetanus become useful functional measures.

IgM is the first isotype to reach young adult levels, followed by total IgG and then IgA. This physiological delay in the production of serum IgA can complicate the diagnosis of IgA deficiency in infants and young children. Serum Ig concentrations in healthy adults tend to remain remarkably constant but can increase dramatically in response to infection and suffer a decline in response to immunosuppressive agents, such as corticosteroid administration. With increasing age, serum immunoglobulin

CLINICAL PEARLS

Interpreting Quantitative Immunoglobulins (Igs)

- Normal ranges of serum immunoglobulin levels vary with age (Appendix 2); hence evaluation should take the age of the patient into account.
  - As the mother's transplacental contribution of IgG is catabolized, total serum IgG concentrations reach a nadir at 4–6 months of age
  - IgG2 and IgG4 subclass levels rise more slowly than IgG1 and IgG3; hence reference to adult controls can lead to the false diagnosis of IgG subclass deficiency in young children.
  - Serum IgA concentrations typically do not achieve adult values until puberty. They are often the first to decline in many primary immunodeficiencies

may continue to rise.[4] The physiological significance of this increase in older adults is unclear, although in some cases it reflects an accumulation of expanded B-cell clones or monoclonal gammopathies (Chapter 80).

KEY CONCEPTS

Functional Tests of Specific Antibody Production

- Immunoglobulin M (IgM) T-independent responses can be assessed by measurement of serum isohemagglutinins (anti-A and anti-B titers) in patients who are not blood type AB.
- IgG T-independent responses can be assessed by measurement of antibodies produced in response to immunization with unconjugated purified pneumococcal polysaccharide vaccine.
- IgG T-dependent recall responses can be assessed by measurement of ≥4-fold rise in titer of antibodies to diphtheria or tetanus toxoid alter booster immunization.

The common laboratory practice of defining the lower range of normal for serum Ig levels as two standard deviations below the age-adjusted mean carries with it the risk of falsely labeling otherwise normal patients as having immunodeficiency. Ig levels vary widely with environmental exposure, and normal biological variation is much broader than that defined by the mean of the population. Patients with a combined deficiency of IgA and an IgG subclass may benefit from more aggressive therapy; thus quantitative measurements of all four IgG subclasses—IgG1, IgG2, IgG3, and IgG4—can be useful in fully defining the extent of humoral immune deficiency.

Among patients with borderline serum IgG levels, tests to evaluate the host's ability to produce functional specific antibody should be performed before making a decision to institute a more aggressive therapy, especially among patients receiving corticosteroids, which can lower total serum IgG while preserving function. The most commonly employed tests include measurement of isohemagglutinins (naturally occurring IgM antibodies to the polysaccharide antigens that define the ABO blood type system on red blood cells [RBCs]) and postimmunization responses to polysaccharide antigens (*e.g.,* Pneumovax23 or unconjugated *H. influenzae B* vaccine) and protein antigens (*e.g.,* tetanus or diphtheria toxoids). IgM is made by the newborn, and most infants can generate isohemagglutinins, making determination of anti-A and anti-B titers a useful measurement of B-cell function even in infants. In older children and adults, isohemagglutinin titers <1 : 8 are considered significantly decreased.[6]

Serum for specific antibody titers should be obtained before and 4–6 weeks after immunization. Optimally, the paired sera should be assayed simultaneously to avoid confusion that may result from single-tube dilution differences at the time the assay is performed. As a general rule, a high baseline titer or ≥4-fold rise in a specific titer in individuals with a low baseline titer confirms that a specific humoral response is intact. The development of conjugated polysaccharide vaccines complicates analysis for those who have received such a vaccine, although information can still be gleaned from study of responses to polysaccharide antigens present only in multivalent unconjugated vaccines.

Because of its relatively low molecular weight and slow turnover compared with other isotypes, protein loss through the kidneys or GI tract can result in the selective loss of IgG. Should symptoms warrant, 24-hour urine for protein or a stool sample for alpha-1-antitrypsin level can document protein loss.

An elevated IgE level may support a suspicion of allergy as an underlying explanation of sinopulmonary symptoms. Serum IgE concentrations are often elevated in patients with IgA deficiency. Extreme elevations of IgE suggest the *hyper-IgE syndrome.*

Enumeration of B cells and of T cells should be performed for any individual who has severe panhypogammaglobulinemia. The most widely used method of demonstrating B cells relies on immunofluorescent labeling of surface CD19, which is restricted in expression to mature B cells. Because an infant may have serum IgG of maternal origin for the first several months of life, the determination of the number of circulating B cells is the single most useful test in making the presumptive diagnosis of XLA, a disorder in which pre-B cells in bone marrow fail to develop to cells of mature phenotype. Absence of circulating B cells also characterizes the *immunodeficiency associated with thymoma* in adults, whereas adults with *chronic lymphocytic leukemia* may have hypogammaglobulinemia with an overabundance of circulating B cells that typically express CD5.

HIGM1 represents the product of a loss-of-function mutation of the *CD154* (*CD40L*) gene. CD154, a surface antigen found on activated T cells, binds CD40 on B cells to facilitate class switching, survival, and proliferation (Chapter 8). A fluorescent-labeled CD40 fusion protein can be used to evaluate the expression of functional CD154 on T cells by flow cytometry. Confirmation of the diagnosis, carrier detection, and prenatal diagnosis often depend on molecular or sequence analysis of the gene.

Replacement Therapy With Human Immunoglobulin

There are a number of commercial preparations of human Ig that are approved by the US Food and Drug Administration (FDA) and available in the United States (Chapter 84). No commercial preparations in this country are available to supplement IgM or IgA exclusively, but some of the preparations contain minute amounts. All commercial human Ig preparations are effective in treating patients with immunodeficiency disorders. Clinically relevant differences relate to the route of administration, which can be by intravenous or subcutaneous means, to the method of stabilization and storage, and to quantities of contaminating serum IgA in the preparations. Low IgA content is a concern for those rare individuals with immunodeficiency and absent IgA, who manufacture IgG or IgE antibodies directed against IgA and have a history of anaphylactic reactions upon infusion of IgA-containing blood products.[7]

Ig replacement therapy is not indicated for patients whose immune deficiency is limited to the selective absence of IgA because of the risk of anaphylaxis upon receipt of IgA-containing products, even though such reactions are extremely rare.[7] However, Ig replacement therapy has been found to be beneficial in patients with a combined deficit of IgA and IgG subclasses who exhibit impaired antibody responses to carbohydrate antigens.

The goal of Ig replacement is to provide sufficient concentrations of functional antibodies to prevent disease, not to achieve a target IgG level in the serum. Specific approaches to and protocols for Ig replacement therapy are described in detail in Chapter 84. Note: *The only indication for immunoglobulin replacement in a patient with immunodeficiency is severe impairment of the ability to produce functional antibody.* Such impairment exists in primary immunodeficiency diseases associated with low levels of all five isotypes of immunoglobulin, such as XLA, CVID, HIGM, and *severe combined immunodeficiency* (SCID) (Chapter 35). Patients with a documented inability to produce specific antibodies after immunization with a history of significant morbidity from infections are also candidates for intravenous immunoglobulin (IVIG) therapy even though they may present with normal or near-normal levels of IgG.[8] This includes certain cases of IgG subclass deficiency, such as those associated with compound IgA, IgG2, and IgG4 deficiency, boys with Wiskott-Aldrich syndrome, and patients with ataxia–telangiectasia. Since most patients with transient hypogammaglobulinemia of infancy can produce normal amounts of specific antibodies after immunization despite having a low total serum IgG, they usually are not candidates for Ig replacement although those patients with a history of significant infections may benefit.

X-LINKED AGAMMAGLOBULINEMIA

Diagnosis

XLA, also known as Bruton agammaglobulinemia, is the prototypic humoral immunodeficiency.[9] Function-loss mutations in *BTK* lead to a block in B-cell maturation, a near total absence of B cells in the periphery, and panhypogammaglobulinemia. As a result of the transplacental transfer of maternal Ig, affected boys typically do not begin to suffer recurrent pyogenic infections until after age 6 months. The normal delay in endogenous Ig production and the presence of maternal IgG requires that testing of infants known or suspected to have XLA should begin with examination of the number of B cells in blood. Deficient expression of BTK protein can be detected by flow cytometry, a technique that can also be used for carrier detection. For those cases where protein is present but the phenotype suggests XLA, analysis of the *BTK* gene at the nucleotide level remains the definitive diagnostic procedure. As with most X-linked lethal diseases, approximately one-third of sporadic cases are caused by *de novo* mutations, and diagnosis may require individual mutation analysis. There can be significant variation in the manifestations of the disease in any given family member; thus a paucity of symptoms should not prevent diagnostic evaluation even in adults.[10] Some cases of XLA have been reported with mild mutations that cause infections late in life.[11]

Clinical Manifestations

Although patients begin to suffer recurrent infections by age 1 year, with antibiotics and good hygiene, it is not uncommon to delay suspicion of the diagnosis well into mid-childhood. Indeed, diagnoses have been made in older adults, including aged male relatives of affected probands.[10,11] Recurrent upper and lower respiratory tract infections are common. Untreated, these infections may lead to bronchiectasis (Fig. 34.3), pulmonary failure, and death at an early age. The infections are typically caused by pyogenic encapsulated bacteria. Diarrhea caused by *G. lamblia* is also common, although less so than in CVID. Systemic infections include bacterial sepsis, meningitis, osteomyelitis, and septic arthritis. *Mycoplasma* and *Chlamydia* infections of the urogenital tract may lead to epididymitis, prostatitis, and urethral strictures. Skin infections include cellulitis, boils, and impetigo.

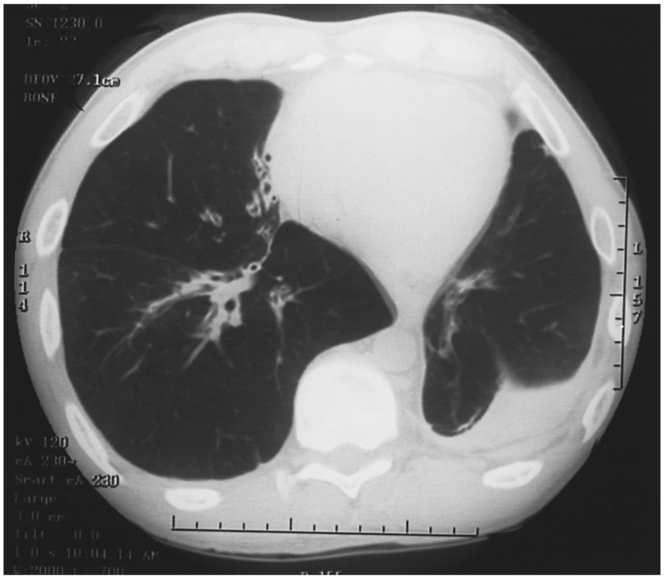


FIG 34.3 A Computed Tomography (CT) Scan, With Contrast, Demonstrates Bronchiectasis, Bronchitis, and Emphysema in the Lungs of a 36-Year-Old Man With X-Linked Agammaglobulinemia (XLA). Because of a left lower-lobe lobectomy, the mediastinum has shifted to the left. As a result of bronchiectatic scarring, the diameter of the bronchi in the right lung is greater than the diameter of the corresponding blood vessels, and the bronchi remain dilated in the periphery. Bronchial plugs can be seen filling some of the bronchi on the right. Finally, as a result of emphysema, the right upper lobe demonstrates greater radiolucency. In addition to suffering from XLA, this patient has a 30 pack-year history of smoking, which has exacerbated his clinical condition.

Although in patients with XLA most viral infections can be resolved, these patients are unusually sensitive to infections with enteroviruses. Patients with XLA can develop paralytic poliomyelitis after vaccination with live virus. Echovirus and Coxsackie virus infections may involve multiple organs with the patients going on to develop chronic meningoencephalitis, dermatomyositis, and/or hepatitis.

Untreated patients often complain of arthritis affecting the large joints. This appears to have an infectious etiology because the arthritis typically resolves with Ig replacement therapy. Enterovirus and mycoplasma have been identified in the affected joints of these patients.

Infections with opportunistic organisms, such as *Mycobacterium tuberculosis, Histoplasma,* and *P. jiroveci,* and malignancies are rare, likely reflecting intact cell-mediated immunity.

Origin and Pathogenesis

BTK belongs to a subfamily of the Src cytoplasmic protein-tyrosine kinases. BTK is phosphorylated following activation of the B-cell receptor (BCR). It plays a critical role in the proliferation, development, differentiation, survival, and apoptosis of

B-lineage cells. Individuals with XLA begin with normal numbers of early B-lineage progenitors in their bone marrow. These B-cell progenitors express the expected markers of B-cell differentiation, including terminal deoxynucleotidyl transferase (TdT), CD19, and CD10. There is, however, a relative deficiency of cells containing cytoplasmic μ heavy chains in bone marrow. Development of cells beyond the pre–B stage is even more severely impaired. Those cells that make it through the gauntlet can produce antigen-specific antibodies. Although present in low numbers, these B cells in lymphoid tissues enable XLA patients to express endogenous Ig, class switch, and even suffer allergic or autoantibody-mediated reactions.

Patients have been described with an X-linked recessive form of agammaglobulinemia that is associated with growth hormone deficiency. Genetic analysis of the *BTK* gene in one such patient identified a frameshift mutation leading to a premature stop codon and the loss of carboxy-terminal amino acids.[12]

Treatment and Prognosis

The primary goal of therapy is to prevent damage to the lungs. Human Ig replacement therapy should be started as soon as the diagnosis is made (Chapter 84). Patients treated with sufficient quantities (0.4–0.6 g/kg every 3–4 weeks for IVIG or 100–150 mg/kg every week for subcutaneous immunoglobulin [SQIG]) suffer few lower respiratory tract infections. However, these patients remain at risk for viral infections, including enteroviral meningoencephalitis. Since mucosal Ig cannot be replaced, the patients also remain at risk for recurrent upper respiratory infections, which may require prophylactic antibiotic therapy. Ig-treated patients may lead normal lives without concern about exposure to infectious agents in childcare settings or classrooms.[13] Immunizations of any type are unnecessary because the monthly replacement therapy will provide passive protection. Since patients are unable to mount antibody responses, vaccines, especially live vaccines, carry some risk of untoward side effects, so these vaccines are relatively contraindicated.

A patient with XLA who develops symptoms of enteroviral central nervous system (CNS) or neuromuscular infection should have appropriate culture of the involved organ system. For patients with agammaglobulinemia who have chronic enteroviral infections, Ig therapy should be given at higher doses and maintained until symptoms cease and the virus can no longer be detected.

AUTOSOMAL AGAMMAGLOBULINEMIA

Origin and Pathogenesis

The Pre–B-Cell Receptor and Signal Transduction Axis

Expression of the pre-BCR is a key step in the maturation of the pre–B cell[14] (Chapter 7). Function-loss mutations in any one of the genes that code for components of the pre-BCR and its associated signaling complex can inhibit pre–B-cell development, leading to an absence of mature B cells. This phenotype is seen in patients with biallelic function-loss mutations of the μheavy chain region (μ_0, AGM1), the λ-like surrogate light chain (IGLL1, AGM2), the Ig-associated α (Igα, CD79A, AGM3) and -β(Igβ, CD79B, AGM6) chains, and the adaptor B-cell linker protein (BLNK, AGM4), which is a key cytoplasmic component of the pre-BCR signaling pathway.

TCF3

The *TCF3* gene, also called *E2A*, encodes 2 basic helix-loop-helix (bHLH) transcription factors, E12 and E47, through alternative splicing. E12 and E47 are involved in regulation of immunoglobulin gene expression. A mutation in the DNA-binding region of E47 has been shown to result in agammaglobulinemia as a result of a dominant negative early block in B-cell development.[15]

LRRC8

An absence of B cells has also been reported in a young female patient with a truncation of leucine-rich repeat containing 8 (*LRRC8*, AGM5),[16] a gene of unknown function that is expressed in patients with progenitor B cells.

PIK3R1

A homozygous nonsense mutation in the *P1K3R1* gene, encoding the p85a subunit of phosphoinositide 3-kinase (PI3K), was found in a 19-year-old female with agammaglobulinemia, absent B cells, and inflammatory bowel disease by exome sequencing. The loss of p85a resulted in decreased pro–B cells, a contrast to the other agammaglobulinemia defect, which is localized to the pre–B-cell stage.[17]

Diagnosis and Treatment

Diagnosis in each of these cases requires gene mutation analysis. Treatment follows the same guidelines given for XLA.

HYPER-IGM SYNDROME

Diagnosis

Patients with the HIGM syndrome exhibit markedly reduced serum concentrations of IgG, IgA, and IgE with normal to elevated levels of IgM and normal numbers of circulating B cells.[18] The altered distribution of Ig isotypes reflects a block in the ability of B lymphocytes to switch from IgM to the other isotypes. Increased IgM reflects polyclonal expansion of IgM synthesis in response to infection. Patients with HIGM suffer from the same infections with encapsulated bacteria common to all patients with antibody deficiency. The HIGM phenotype can be inherited as an X-linked, autosomal recessive, or autosomal dominant trait. The phenotype can also be acquired in association with neoplasia or congenital rubella.

HIGM Syndrome Type 1: CD40L (CD154) Deficiency

Class switch recombination is a multistep process that requires exquisite coordination between the B cell and its cognate T-helper (Th) cell. A key step in the initiation of the process is the binding of constitutively expressed CD40 on the B cell to its ligand, CD40L (CD154), which is expressed on activated T cells. The most common form of the disease, X-HIGM1, results from loss-of-function mutations in *CD154* (Xq26).

HIGM Syndrome Type 2: AID Dysfunction

Activation-induced cytidine deaminase (AID; 12p13), a member of the cytidine deaminase family, is required for class switch recombination between Ig H chain constant domains and for somatic hypermutation of the Ig V domains (Chapter 4). The hyper-IgM phenotype (HIGM2) can result from either biallelic *AID* function-loss mutations or from a dominant negative mutation on only one of the *AID* alleles.

HIGM Syndrome Type 3: CD40 Deficiency

The cognate partner for CD40L (CD154) is CD40, the gene for which (*CD40*) is located on an autosome (20q12-q13.2). Patients with HIGM3 with function-loss mutations on both alleles of the *CD40* gene present with a phenotype indistinguishable from HIGM1.

HIGM Syndrome Type 4: As yet Unknown Causes

Patients presenting with a HIGM-like phenotype, but lacking demonstrable mutations in genes previously associated with HIGM, such as *CD40, CD154, NEMO, AID,* and *UNG,* have been grouped by some immunologists into a category termed *HIGM4.* B cells from patients with HIGM4 demonstrate defective class switch recombination, but normal somatic hypermutation. A genetic cause has yet to be identified, and spontaneous recovery has been reported.[19]

HIGM Syndrome Type 5: UNG Deficiency

Activation-induced cytidine deaminase (AID) acts by deaminating cytidine nucleotides in DNA, leaving a uracil nucleotide in its place. Uracil-DNA glycosylase (UNG; 12q23-24.1) can remove the uracil, permitting normal or error-prone repair. Patients with function-loss mutations on both alleles of the *UNG* gene (HIGM5) have presented with a history of bacterial infections, hyperplasia, increased serum IgM levels, and low IgG and IgA.

NEMO

The nuclear factor κB (NF-κB) essential modulator (NEMO) plays a key role in the NF-κB pathway and consequently in the CD40 signal transduction pathway. NEMO acts as a scaffold for two kinases important for the activation of NF-κB. The *IKBKG* gene encodes for NEMO and is located on the X-chromosome (Xq28). Hypomorphic mutations in the gene causes a syndrome with affected individuals having ectodermal dysplasia. This results in conical teeth, an absence of eccrine sweat glands, and a paucity of hair follicles. Patients with NEMO are considered to have a defect in innate immunity and cell-mediated immunity as a result of defective NF-κB activation, which is important in Toll-like receptor (TLR) signaling. Laboratory abnormalities include impaired natural killer (NK) cell function, impaired pneumococcal responses, hypogammaglobulinemia, antigen-specific T-cell proliferative abnormalities, and increased serum IgA levels. Less than 20% of patients with NEMO have HIGM and, thus in the past, have been grouped with the HIGM group.[20]

Clinical Manifestations

CD40-CD154 Axis (HIGM1, HIGM3, NEMO)

The majority of patients with HIGM have inherited disruptions in the CD40-CD154 axis. Lymph nodes and spleen are deprived of germinal centers. Recurrent upper and lower respiratory tract infections are the most common clinical complaint. Patients may also exhibit recurrent neutropenia with oral ulcers, perirectal abscesses, and opportunistic infections with *P. jiroveci, Toxoplasma gondii,* or *Cryptosporidium* cholangitis. Autoimmunity is observed in approximately 20% of patients.

Without prophylaxis, one third of patients develop *P. jiroveci* pneumonia, which can be the presenting problem in affected infants. These patients are also at risk for serious infections with cytomegalovirus (CMV), adenovirus, *Cryptococcus neoformans,* or mycobacteria. These features signal a compromise in cell-mediated immunity as well as the characteristic flaw in antibody production, putting them in the spectrum of a combined immunodeficiency.

Chronic diarrhea occurs in more than half of the patients. Organisms include *Cryptosporidium, G. lamblia, Salmonella,* and *Entamoeba histolytica.* One quarter may require total parenteral nutrition because of the presence of diarrhea or perirectal abscesses. Oral ulcers, gingivitis, and perirectal ulcers are associated with neutropenia, which may occur chronically or intermittently in up to two-thirds of the patients. One-fifth of the patients develop sclerosing cholangitis that can lead to hepatic failure. Cryptosporidiosis is present in half of these patients.

Approximately one-quarter of patients with NEMO have autoimmunity or autoinflammation. An intestinal inflammatory disorder may be the presenting problem with chronic diarrhea and abdominal pain, with a few having steroid dependence.

Although originally distinguished by the high level of serum IgM, IgM levels are often normal in affected individuals. IgG is low in all patients. IgA and IgE are usually low but can be normal or even elevated in one-tenth of the population. B- and T-cell counts are within the normal range in more than 90% of the patients and depressed in the rest.

Lymphoid hyperplasia is a common finding in CD40–CD154 axis patients with active infections. Individual nodes may become extremely large, and some patients develop splenomegaly. Hilar adenopathy causes a diagnostic dilemma, as the risk of lymphoma is increased in HIGM. Although the lymphoid tissue is usually histologically abnormal, reactive processes are far more common than malignancy. Plasma cells can be abundant or sparse. Primary follicles are poorly developed. The most characteristic abnormality is the absence of germinal centers.

AID–UNG Axis (HIGM2 and HIGM5)

Infected patients with AID–UNG deficiency may present with giant germinal centers filled with highly proliferating B cells, presumably as a result of intense antigen stimulation. Approximately one-fourth of patients with HIGM2, but not HIGM5, present with evidence of autoimmunity. Manifestations include hemolytic anemia, thrombocytopenia, and autoimmune hepatitis. Autoantibodies in these patients are of the IgM isotype.

Origin and Pathogenesis

CD40–CD154 Axis (HIGM1, HIGM3, and NEMO)

CD154, a member of the tumor necrosis factor (TNF) family, is a type II transmembrane protein that is predominantly expressed on mature, activated CD4 T cells. Its expression peaks at 6–8 hours after activation and then falls to resting levels by 24–48 hours. CD154 is also expressed on CD4 thymocytes, activated CD8 T cells, NK cells, monocytes, basophils, mast cells, activated eosinophils, and activated platelets. Newborn T cells are deficient in CD154 expression, although they can be induced to express the antigen if strongly stimulated.

CD40 is a member of the TNF receptor superfamily. It is constitutively expressed on pro-B, pre-B, and mature B cells, as well as on interdigitating cells, follicular dendritic cells (DCs), thymic epithelial cells, monocytes, platelets, and some carcinomas.

Engagement of B-cell CD40 with CD154 on an activated T cell that also expresses Fas ligand (FasL or CD95L) leads to the upregulation of Fas (CD95) on the B cell. NEMO is a part of the signaling pathway. If the B cell has concomitantly bound its cognate antigen and engaged the BCR-signaling pathway, it becomes resistant to Fas-mediated apoptosis and expresses CD80/CD86 on the cell surface. The activated B cell can then engage CD28 on the T-cell surface and trigger the T cell to secrete its cytokines. If the B cell fails to engage its BCR, the Fas pathway predominates, and the B cell is eliminated. With proper activation of the CD40–CD154 pathway, exposure to interleukin-2 (IL-2) and IL-10 induces production of IgM, IgG1, and IgA; and exposure to IL-4 induces production of IgG4 and IgE. This change in Ig isotype reflects both induction of switching and the enhanced survival and proliferation of the B cell. In the absence of CD154, B cells can express IgM but have difficulty switching and are likely to undergo apoptosis, rather than proliferate in response to antigen.

CD40–CD154 interactions between $CD154^+$ T cells and $CD40^+$ macrophages lead to enhanced production of IL-12, which then stimulates T cells to release interferon-γ (IFN-γ). Activation of this pathway appears necessary for the defense against *P. jiroveci* and other opportunistic organisms.

Treatment and Prognosis

The availability of human Ig replacement therapy has greatly improved the quality of life in patients with HIGM. Adequate replacement can result in the reduction of serum IgM concentrations, prevention of infections with encapsulated bacteria, resumption of growth, and the gradual resolution of splenomegaly and lymphoid hyperplasia. Autoimmune and lymphoproliferative complications may respond to anti-CD20 therapy (rituximab).[21]

Unfortunately, in spite of the improvement gained by Ig replacement, the prognosis of patients with defects in the CD40–CD154 axis remains guarded. Death at a young age continues to be common. It is primarily the result of opportunistic infections, including *Pneumocystis* pneumonia, cholangitis, CMV infection, mycobacterial infections, and cirrhosis secondary to hepatitis. Prophylaxis with trimethoprim–sulfamethoxazole can significantly reduce the risk of *Pneumocystis* pneumonia and is indicated in those with CD40L and CD40 deficiency. Regular monitoring of GI manifestations and management of neutropenia is mandatory. Neutropenia should be treated with a trial of granulocyte macrophage–colony-stimulating factor (GM-CSF), since some have responded to this therapy. Bone marrow transplantation is a viable option for patients that fail to respond to supportive therapy.

SELECTIVE IGA DEFICIENCY

Selective IgA deficiency (IgAD), selective IgG subclass deficiencies, CVID, and a syndrome of recurrent sinopulmonary infections (RESPIs) with normal serum Ig levels appear to share an overlapping set of gene defects.[22] Clinically, these disorders are marked by an increased susceptibility to upper and lower respiratory infections with encapsulated bacteria. IgAD and CVID feature similar B-cell differentiation arrests but differ in the extent of Ig deficits. The correlation between serum Ig levels and severity of infection is not absolute.[23]

Diagnosis

Approximately 1 in 600 individuals of European ancestry are unable to produce detectable quantities of IgA1 and IgA2, making selective IgAD the most frequently recognized primary immunodeficiency in the Americas, Australia, and Europe. The diagnosis is dependent on the sensitivity of the laboratory measurement. The clinical laboratory typically reports serum IgA levels of less than 7 mg/dL, the concentration below which nephelometry becomes unreliable.

Patients with uncomplicated IgAD have normal serum levels of IgM, have normal or elevated levels of IgG, and demonstrate normal cell-mediated immunity. A minority of patients may demonstrate additional evidence of immune dysfunction, with inability to generate appropriate IgG2 anticarbohydrate antibodies, frank IgG subclass deficiencies, or evidence of impairment of T-cell function. Patients with IgA serum levels that fall >2 standard deviations (SDs) below the mean serum level for their age are considered to have partial IgA deficiency. These patients can also suffer from recurrent infections.

Clinical Manifestations

The likelihood that an individual with IgA deficiency that was identified serendipitously will require medical attention is difficult to assess because most studies in the literature reflect patients who were ascertained to have the condition as a result of clinical symptoms. Among patients with IgAD referred to immunology clinics, more than 85% present with recurrent infections, typically with encapsulated bacteria. Among affected children, symptoms may begin in the first year of life, although the physiological lag in serum IgA may delay the diagnosis until after the age of 2 years. In some patients, respiratory infections disappear with maturity. In others, infections may persist throughout adult life. Rarely, patients with IgAD may experience recurrent bronchitis, pneumonia, and even bronchiectasis. These more severely afflicted patients often exhibit concurrent IgG2 and IgG4 subclass deficiencies. Some symptomatic patients have elevated IgE levels and manifest allergic or asthmatic components to respiratory dysfunction. The rise in IgE has been explained as a compensatory response to the absence of IgA. This appears to be a double-edged sword because up to 20% of these patients complain of allergic rhinitis, conjunctivitis, urticaria, and atopic eczema. Allergic reactions may be enhanced because of the lack of IgA-blocking antibodies in serum, and unusually severe asthma has also been associated with IgAD.

Among those less common IgA-deficient patients that are truly devoid of IgA, as many as three-fifths produce IgG or IgE anti-IgA antibodies.[24] These uncommon patients are at an uncertain risk for adverse reactions following transfusion with blood products (as mentioned previously), plasma from normal donors, or from some preparations of Ig replacement therapy, which, of course, contain IgA. Patients with high anti-IgA levels (>1 : 1 000) typically have potent antibodies directed against all IgA. These patients are at risk for severe anaphylaxis. Patients with low anti-IgA antibody titers (<1 : 256) are often multiparous women or those who had received multiple transfusions. These patients rarely demonstrate severe anaphylaxis after infusion with plasma or blood products but do present with hives and rashes.

Patients with IgAD often develop autoimmune diseases. GI disorders include pernicious anemia, inflammatory bowel disease, intestinal disaccharidase deficiency, lactase deficiency, pancreatic insufficiency, and celiac disease. The last, in particular, can be difficult to diagnose without biopsy, since serological diagnosis often relies on detection of antitissue transglutaminase and antiendomysial or antigliaden IgA antibodies.[25] Hepatobiliary disorders include chronic active hepatitis, cholelithiasis, lupoid hepatitis, and primary biliary cirrhosis. Skin disorders include pyoderma gangrenosum, paronychia, and vitiligo. It is unclear whether this autoimmune diathesis is the end result of recurrent infections or is the product of recurrent insult by antigens that would otherwise be cleared by IgA, or whether the underlying deficit that leads to IgAD also increases the risk of developing an autoimmune disorder. For example, autoimmune disorders, such as insulin-dependent diabetes mellitus and celiac disease, are associated with the same major histocompatibility complex (MHC) haplotypes (Chapter 5) as IgAD and CVID.

IgAD is associated with an increased risk for the development of malignant epithelial tumors, such as gastric and colonic adenocarcinoma, Hodgkin disease, and acute lymphoblastic leukemia. Patients with chronic GI infections may demonstrate a nodular lymphoid hyperplasia of the small intestine that can lead to intestinal obstruction. Histological evaluation reveals active B-lymphocyte proliferation in the germinal centers of the Peyer patches. These "constipated" lymph nodes have been mistaken for lymphoma. In others, the simultaneous presence of IgAD and malignancy may simply reflect the high prevalence of IgAD in the Caucasian population.

Origin and Pathogenesis

IgAD, selective IgG subclass deficiencies, and CVID are diseases that are defined by a quantitative phenotype, a paucity of serum Igs of a given isotype in spite of the presence in the blood of B lymphocytes bearing the missing isotypes. By definition, the fundamental defect involves the failure of B lymphocytes bearing a given isotype to differentiate into plasma cells. These diseases appear to represent a common endpoint for multiple pathogenic processes. All three phenotypes may be acquired and many of the recognized precipitating causes, such as phenytoin, are the same (Table 34.3).

TABLE 34.3 Other Conditions Associated With Humoral Immunodeficiency

Genetic Disorders	
Monogenic diseases	Ataxia–telangiectasia Autosomal forms of severe combined immunodeficiency (SCID) Transcobalamin II deficiency and hypogammaglobulinemia Wiskott-Aldrich syndrome X-linked lymphoproliferative disorder (Epstein-Barr virus [EBV] associated) X-linked SCID
Chromosomal anomalies	Chromosome 18q- syndrome Monosomy 22 Trisomy 8 Trisomy 21
Systemic Disorders	
Malignancy	Chronic lymphocytic leukemia Immunodeficiency with thymoma T-cell lymphoma
Metabolic or physical loss	Immunodeficiency caused by hypercatabolism of immunoglobulin Immunodeficiency caused by excessive loss of immunoglobulins and lymphocytes
Environmental Exposures	
Drug-induced	Antimalarial agents Captopril Carbamazepine Glucocorticoids Fenclofenac Gold salts Imatinib Levetiracetam Penicillamine Phenytoin Sulfasalazine Zonisamide
Infectious diseases	Congenital rubella Congenital infection with cytomegalovirus (CMV) Congenital infection with *Toxoplasma gondii* EBV Human immunodeficiency virus (HIV)

IgA deficiency is associated with MHC haplotypes (6p21.3) that are more common in European populations than in the peoples of Sub-Saharan African and East Asia. In the United States, the prevalence of IgAD among African Americans is one-twentieth of that observed among Americans of European descent, and in Japan, the incidence is approximately 1 in 18 500. IgA deficiency has also been observed in family members of patients with CVID with altered function of the transmembrane activator and CAML interactor (TACI, 17p11.2), which is a receptor for B cell–activating factor (BAFF).

Treatment and Prognosis

Most individuals with IgAD suffer respiratory infections no more frequently than the average individual and thus require no special treatment. All individuals with IgA deficiency should be warned of the risk of serious transfusion reactions caused by antibodies to IgA. Wearing a medical alert bracelet is recommended. Should transfusion be necessary, the ideal donors are other individuals with IgAD. Washed erythrocytes are safer than whole blood.

Patients with selective IgA deficiency who suffer from clinically significant, recurrent upper respiratory infections often respond to prophylactic antibiotics with potency against encapsulated bacteria. Treatment of allergy in those patients with a compensatory increase in IgE is helpful. Patients who present with combined IgA and IgG subclass deficiencies and have a poor pneumococcal antibody response may require IG replacement therapy.

COMMON VARIABLE IMMUNODEFICIENCY AND CVID-LIKE DISORDERS

Diagnosis

The diagnostic category of CVID includes a heterogeneous group of patients older than age 4 years and exhibit deficient production of more than one major antibody class and whose antibody response to vaccination is significantly depressed or absent. These patients tend to have normal numbers of clonally diverse B lymphocytes in their blood. These B cells can recognize antigens and respond with proliferation, but their ability to develop into memory B cells or mature plasma cells appears quantitatively impaired. In the presence of infection, abortive differentiation can lead to massive B-lymphocyte hyperplasia, splenomegaly, and intestinal lymphoid hyperplasia.

With an estimated prevalence of 1 in 25 000, CVID is the most prevalent human primary immunodeficiency requiring medical attention.[26] Both sexes are equally affected. As with IgAD, the prevalence among African Americans is one-twentieth that of Americans of European descent. Some patients present during childhood, but most are diagnosed after the third decade of life. The typical patient reports a normal pattern of recurrent otitis media as an infant and toddler that resolved in childhood. During adolescence, respiratory infections appear and steadily increase in frequency and duration. Recurrent pneumonia as a young or middle-aged adult is often the precipitating complaint that brings the patient to the attention of the clinical immunologist. Although CVID appears to be an acquired disorder, family studies have clearly documented that susceptibility for the disease can be inherited and the manifestations of the disorder may change with time. Transitions within the spectrum of normal serum Ig concentrations to IgA deficiency to IgA deficiency with IgG subclass deficits to frank CVID have been documented in both sporadic and familial cases.[2]

CVID is a diagnostic category of primary immunodeficiencies that includes a number of immune disorders. Most patients of Northern European descent with CVID exhibit a distinctive phenotype characterized by a broad deficiency of Ig isotypes in spite of the presence of normal numbers of surface Ig bearing B-cell precursors in peripheral blood. Almost all of these patients are IgA deficient and, by definition, demonstrate total serum IgG levels of <500 mg/dL. Some IgG subclasses are more affected than others, with the sequential order of involvement being IgG4 > IgG2 > IgG1 > IgG3. Most patients are also deficient in IgM and IgE.

Patients with uncomplicated conditions demonstrate normal cell-mediated immunity, although a minority of patients may have evidence of T-cell dysfunction as well as other hematopoietic cell types. In some cases, B-cell numbers are reduced, although not to the extent exhibited by disorders of pre-BCR formation or BTK signaling.

IgAD and CVID have been associated with congenital infection with rubella virus, CMV, and *T. gondii.* The administration of certain drugs has also been linked to a depression in serum Ig levels (see Table 34.1). Several medications used to treat epilepsy have been associated with the development of antibody deficiencies. For example, up to 20% of patients treated with phenytoin develop a mild decrease in serum IgA levels, and a minority may progress to a CVID-like phenotype. Medications used for the treatment of rheumatoid arthritis, inflammatory bowel disease, and chronic myelogenous leukemia can also decrease production of antibody. Persistence of antibody deficiency usually requires choosing between discontinuation of the drug and persisting infection with the virus or parasite. Recovery of Ig production may take months to years.

Clinical Manifestations

Although some patients with CVID have reduced numbers of circulating B cells, the majority has normal quantities of IgA, IgG, and IgM-bearing B-cell precursors in the blood. Defects in B-cell survival, number of circulating $CD27^+$ memory B cells (including IgM^+CD27^+ B cells), B-cell activation after antigen receptor cross-linking, T-cell signaling, and cytokine expression have been observed. Both increases and decreases in the relative numbers of CD4 to CD8 T cells are common, and cutaneous anergy is a frequent finding.

The clinical manifestations of CVID are similar but more severe than the ones seen in IgAD. Respiratory symptoms often begin with recurrent sinusitis, otitis media, and mild bronchitis. The frequency and severity of the upper respiratory infections worsen in the young adult, and lower respiratory infections, such as pneumonia, become common. Apparently asymptomatic, untreated patients may suffer recurrent subclinical pulmonary infections that can lead to irreversible chronic lung damage with bronchiectasis, unilateral hyperlucent lung, emphysema, and cor pulmonale. Recurrent cellulitis, boils, folliculitis, impetigo, or erythroderma can be presenting complaints.

Intermittent or chronic diarrhea due to *G. lamblia* is a common complaint. Patients can develop a malabsorption syndrome that resembles celiac sprue but is unresponsive to gluten avoidance (Fig. 34.4). Untreated patients often complain of asymmetrical, oligoarticular arthralgias or frank arthritis, which in some cases reflect infections with encapsulated organisms or with *Mycoplasma* species and thus require antibiotic therapy. Paradoxically, antigen-specific IgE can be produced in sufficient quantities to enable anaphylactic reactions.

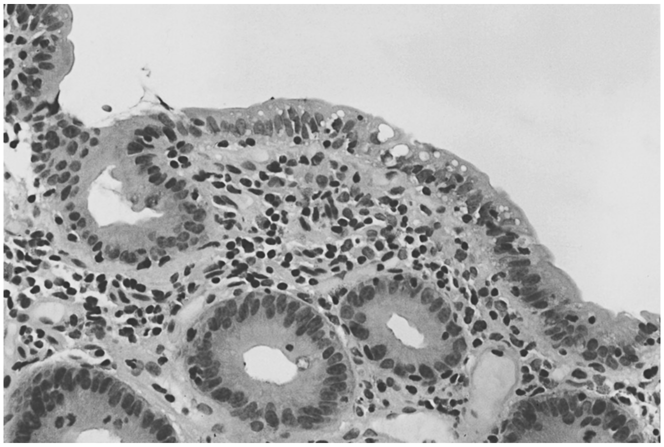


FIG 34.4 Hypogammaglobulinemic Sprue in a 41-Year-Old White Male With Common Variable Immune Deficiency (CVID) and Insulin-Dependent Diabetes Mellitus. The patient suffered from intractable diarrhea. Shown is a hematoxylin and eosin (H&E) stain of a duodenal biopsy obtained by endoscopy. The villi are blunted, and there is a marked increase in intraepithelial lymphocytes. However, unlike typical celiac disease, the villi are not completely blunted, and few plasma cells are seen. The patient is homozygous for the HLA-DQ2, -DR17(3), -B8 haplotype. Although the patient failed to respond to a gluten-free diet, the diarrhea resolved with corticosteroid therapy.

Patients with CVID are often anergic, but only a minority suffers infections characteristic of cell-mediated immune dysfunction, including mycobacteria, *P. jiroveci,* and fungi. CD8 T-cell numbers may be depressed in such patients. Most viral infections are cleared normally. Exceptions include enteroviral infections, including meningoencephalitis, as well as hepatitis B and C, which can progress to chronic active hepatitis. Lack of humoral immunity enhances susceptibility to viral reactivation. Untreated patients often complain of recurrent herpes zoster (shingles).

Autoimmune diseases are common in CVID. Coombs-positive hemolytic anemia with idiopathic thrombocytopenic purpura, a combination known as *Evans syndrome,* may predate the diagnosis of CVID.

Noncaseating granulomas in the lungs, lymph nodes, skin, bone marrow, and liver reminiscent of or indistinguishable from sarcoid-like syndrome is more common in African Americans but can be seen in up to one-fifth of all patients. Occasionally the granulomas result from mycobacterial and fungal infections. In the majority of cases, the cause remains unclear, and the granulomas resolve spontaneously.

There is an increased risk for the development of GI and lymphoid malignancies, especially non-Hodgkin lymphomas. Confounding the diagnosis of malignancy is the patients' propensity to develop benign lymphoproliferative disorders. Lymphadenopathy, splenomegaly, or a combination of both is common in untreated patients.

Origin and Pathogenesis

The typical presenting manifestation of CVID is hypogammaglobulinemia, not agammaglobulinemia, suggesting a partial or varying block in B-cell maturation. Careful analysis of B cells in patients has also revealed a spectrum of immune deficiency ranging from the nearly complete absence of memory B cells to a less severe disorder. All of these findings serve to underline the complex etiology for the disorder, and many details remain to be elucidated. The MHC represents the most common genetic susceptibility locus for CVID.[22] Because of linkage disequilibrium, the gene, or genes, within this locus that create susceptibility have yet to be identified with certainty. Non–MHC-associated single-gene defects have been identified, although they represent only a minority of patients with CVID. These include function-loss mutations in genes involved in late-stage B cell–T cell communication, late-stage B-cell growth factors, and B-cell and T-cell signaling and activation pathways (see Table 34.1). These include the genes for ICOS (CVID1), an immune costimulator molecule used by T cells to activate B cells in germinal centers; BAFFR (CVID4) and TACI (CVID2), the receptors for BAFF, CD19 (CVID3), CD21 (CVID7), and CD81 (CVID6), components of the B-cell costimulatory receptor; CD20 (CVID5), an important marker of B-cell differentiation; and LRBA, Cytotoxic T lymphocyte antigen-4 (CTLA-4), PKCδ, Tweak, PIK3CD, PIK3R1, NF-κB2, and KMT2D, which are involved in B-cell and T-cell signaling pathways. TLR7 and TLR9 activation can be deficient in these patients, although the genes are intact.[27]

The Major Histocompatibility Complex

A large array of genes that play important roles in the control of the immune response are located in the MHC on chromosome 6 (Chapter 5). Studies at the University of Alabama at Birmingham have shown that in the Southeastern United States the majority of patients with IgAD and CVID share parts or all of one of three extended MHC haplotypes marked by either HLA-DQ2, -DR17(3), and -B8 or HLA-DQ2, -DR7, and -B44. The combined results of three studies of patients from Alabama, New England, and Australia indicated a 13% prevalence of immunodeficiency in individuals homozygous for HLA-DQ2, -DR17(3), and -B8. These MHC alleles are also more common in patients who suffer from diabetes mellitus, pernicious anemia, celiac disease, autoimmune thyroid disease, and myasthenia gravis. Some individuals with TACI mutations inherited MHC haplotypes associated with the disease,[28] and the combination of specific MHC and KIR alleles also increases susceptibility.[29] This suggests that epistatic interactions between different susceptibility alleles may influence the penetrance of the disorder.

The CD19 (CVID3), CD81 (CVID6), CD21 (CVID7) B-Cell Coreceptor Complex

CD21 (complement component C3d/Epstein-Barr virus receptor 2) binds to membrane- IgM-bound antigen when complement C3d is also attached to that antigen (Chapter 21). In association with CD81 and CD19, this coreceptor complex enhances the antigen-binding signal, promoting B-cell activation. Patients with mutations in CD19, CD21, and CD81 have been reported.[30]

The BAFF–BAFFR (CVID4)–TACI (CVID2) Axis

The TNF family members B-cell activating factor of the TNF family (BAFF) and a proliferation-inducing ligand (APRIL) bind to two receptors, B-cell maturation antigen (BCMA) and transmembrane activator and calcium-modulator and cyclophilin ligand interactor (TACI), both members of the TNF-R family. BCMA is expressed exclusively on B cells, whereas TACI is expressed on activated T cells as well. A third receptor, BAFF-R, which is unique for BAFF, is expressed on B cells and on resting T cells. The BAFF/APRIL system plays a key role in mature B-cell homeostasis and development. BAFF and APRIL can also induce class switching in naïve human B cells. Loss-of-function (autosomal recessive) or altered-function (autosomal dominant) TACI alleles have been found in approximately 10% of patients with CVID in many series.[31] Two polymorphic alleles, A181E and C104R, are prominent in patients with CVID who have TACI alterations. These alleles are also present in approximately 2% of the normal population, suggesting that the presence of these altered alleles functions as a susceptibility factor for the development of the disease. Family members may have IgA deficiency or may have no evidence of immune dysfunction. However, patients with CVID who have these altered alleles have a higher prevalence of complications from CVID, including lymphoproliferation, splenomegaly, and autoimmune phenomena.[32] BAFF-R deficiency has been reported in four patients with CVID to date.[30]

CD20 (CVID5)

CD20 encodes a B-cell membrane–spanning molecule important in B-cell proliferation and differentiation. It is expressed from early pre-B until mature B-cell stage but is lost on differentiation to plasma cell. One case of a female consanguineous patient with CD20 deficiency with low IgG and normal IgA and IgM levels and impaired antibody responses to pneumococcal polysaccharides, has been reported.[30,33]

ICOS (CVID1)

ICOS is a T-cell surface receptor that is important for germinal center formation, terminal B-cell differentiation, effector T-cell responses, and immune tolerance. Patients with ICOS deficiency have low to absent B cells, and some have varying degrees of defective T-cell signaling. They present with recurrent respiratory tract infections and autoimmune complications.[30,33]

The LRBA (Cvid8)–CTLA-4 Axis

Lipopolysaccharide-responsive beige-like anchor protein (LRBA) is a cytosolic protein that functions in vesicle trafficking, autophagy, and cell survival. It is expressed by almost all cell types with higher expression in immune effector cells. CTLA-4 is an inhibitory T-cell receptor (TCR) that competes with the costimulatory protein CD28 for binding CD80/86, thereby preventing excessive T-cell activation and maintaining immune tolerance. LRBA plays a role in CTLA-4 surface expression.

Patients with LRBA deficiency display early-onset hypogammaglobulinemia with autoimmunity and inflammatory bowel disease. They show reduced levels of at least two Ig isotypes (IgM, IgG, or IgA) and suffer from recurrent infections, autoimmunity, and chronic pulmonary and GI disorders.

Patients with haploinsufficiency of CTLA-4 presented with autoimmunity, recurrent infections, benign lymphoproliferation, and varying levels of Igs and B-cells, and T-cell defects.[30,33]

PKCδ Deficiency

PKCδ plays a key role in BCR-mediated signaling downstream of BTK and is important in B-cell proliferation, apoptosis, and tolerance. PKCδ deficiency presents with a variable phenotype, with one affected patient having a CVID-like characteristics (hypogammaglobulinemia and severe infections) and others having lupus or autoimmune lymphoproliferative syndrome (ALPS)–like disease.[30]

TWEAK Deficiency

A patient with TNF-like weak inducer of apoptosis (TWEAK) deficiency displayed low to normal IgG, low IgM, and low IgA and had a history of pneumococcal meningitis, osteomyelitis, thrombocytopenia, and neutropenia. The patient had an autosomal dominant mutation in TNF superfamily member 12 (*TNFSF12*), which encodes TWEAK.[30]

NF-κB1 (CVID12) and NF-κB2 (CVID10) Deficiency

The NF-κB1 and NF-κB2 (noncanonical) pathways are important in B cell signaling, with NF-κB2 having a more limited set of involved receptors, such as ICOS, TACI, BAFR-R, and BCMA, whereas NF-κB1 affects T-cell and TLR signaling as well.

Heterozygous mutations in NF-κB2 were identified in at least 10 patients who presented with early-onset CVID with autoimmunity and recurrent sinopulmonary infections. Patients with NF-κB2-deficiency display panhypogammaglobulinemia, aberrant B-cell subsets, and some degree of T-cell and NK-cell dysfunction. Half of these also suffer from pituitary hormone deficiencies.

Patients with *NK-KB1* autosomal dominant mutations that create an unstable and rapidly degrading protein have recurrent infections, autoimmunity, benign lymphoproliferative disease, and lymphoma.[30]

PI3K Mutations

Heterozygous mutations in *P1K3CD,* which encodes the P13K catalytic subunit p110δ, have been reported in over 50 patients with CVID-like disease. The mutation leads to P13K signaling pathway overactivity. Such patients suffer from respiratory tract infections, skin infections, autoimmunity, and lymphoma. The phenotype associated with dominant gain of function *P1K3CD* mutations is currently referred to as activated phosphatidylinositol 3-kinase δ syndrome (APDS).

PIK3R1 encodes the P13K regulatory subunit p85α. Splice-site mutation of p85α results in a loss of an exon of p85α, which is important in inhibiting the catalytic activity of p110d. This also results in autosomal dominant overactive P13K signaling. A single patient with a homozygous *P1K3R1* mutation causing complete loss of p85α expression has been found with agammaglobulinemia and absence of B cells, suggesting complete absence of p85α reduces inhibiting P13K signaling.[30]

Other Genes: *BLK, IRF2BP2, IKAROS*

A heterozygous loss of function mutation in *BLK* has been noted in related patients with CVID. These patients have respiratory tract infections and bacterial skin infection with panhypogammaglobulinemia.

A gain of function mutation in *IRF2BP2* has been found in members of a family with CVID. These patients have autoimmune disease and respiratory tract infections.

Heterozygous mutation in the gene *IKZF1,* encoding the transcription factor IKAROS, was recently found in patients with CVID and low B-cell numbers. These patients have progressive loss of B cells and serum Igs.[30]

Kabuki Syndrome

Kabuki syndrome (KS) is a rare but recognizable condition that consists of a characteristic face, short stature, cardiac anomalies, a variable degree of intellectual disability, and immunological defects. Mutations in *KMT2D* and *KDM6A* have been identified as the main cause for KS. Similar to CVID, patients with KS have recurrent infections, reduced Ig levels, and autoimmunity.[34]

Treatment and Prognosis

Therapy in CVID begins with the aggressive treatment of ongoing infections and the institution of prophylactic measures to prevent or ameliorate future infections. Patients suffering from moderate upper respiratory tract infections and bronchitis will likely benefit from empiric therapy with agents effective against encapsulated organisms. The concomitant inheritance of MBL protein deficiency appears to further predispose to the development of bronchopulmonary complications, such as bronchiectasis, lung fibrosis, and respiratory insufficiency.[3] The course of treatment for patients with immunodeficiency is often prolonged, and intravenous administration of antibiotics may be required.

The most effective therapy for hypogammaglobulinemia is Ig replacement. A number of studies have demonstrated a steadily decreasing incidence of infection with increasing frequency of Ig administration. At higher doses, even patients with bronchiectasis may experience improvement in pulmonary function. Each patient may demonstrate his or her own individual response to therapy, exhibiting dramatic differences in the frequency and severity of infections, with moderate changes in the replacement dose. Patients suffering from a serious acute infection often benefit from one-time booster doses of Ig. Ultimately, replacement dosage must be individualized based on the response of the patient. Adverse reactions occur most frequently at the time of the first administration of Ig, likely because of concurrent infection increasing the potential for generation of immune complexes.

Some patients with CVID can sustain severe anaphylaxis when given IVIG or other blood products that contain serum or plasma. These patients may possess anti-IgA antibodies, including IgE anti-A antibodies.[7] For patients with a history of severe adverse reactions, it is advisable to try batches of IVIG with the lowest IgA possible and to test the patient with the different batches in an intensive care unit. Once the clinician has identified a batch that can be tolerated, the patient may receive therapy under more relaxed conditions.

Serum Ig concentrations in patients with CVID may change over time,[2] with rare patients regaining normal serum IgG levels and no longer requiring Ig therapy. Careful review of the clinical history of these patients may reveal evidence of exposure to pharmacological agents associated with the development of hypogammaglobulinemia (*e.g.,* phenytoin). However, the overwhelming majority of patients require replacement therapy for life.

Although IgG can be replaced, at present IgM and IgA cannot be provided to patients. The absence of these multimeric proteins may help explain why even patients on high-dose replacement therapy may continue to suffer from recurrent sinusitis or GI discomfort.[35] Recurrent sinusitis can be ameliorated with continued prophylactic therapy with antibiotics. Patients with CVID also are at risk of infection by *G. lamblia* as well as other enteric pathogens. Some patients develop lactose intolerance or gluten-sensitive enteropathy. Gluten avoidance ameliorates symptoms in only a minority of cases. A majority responds to corticosteroids or anti-TNF agents. The use of these agents can be a double-edged sword, however, since resistance to infection will decrease in a patient who already has immune deficiency. Other patients develop a malabsorption syndrome that can lead to hypoalbuminemia and hypocalcemia (as a result of malabsorption of vitamin D), and decreased levels of vitamin A and carotene.[36] The cause of diarrhea and malabsorption in this latter patient subset remains unclear, and treatment is limited to supportive measures, with vitamin and mineral replacement as indicated.

Patients with bronchiectasis should be treated aggressively with replacement therapy. In severe cases, aggressive pulmonary toilet will benefit the patient, including bronchodilator therapy, position and postural drainage, or other physical therapies. The use of corticosteroids should be avoided.

Mothers with IgA deficiency may fail to secrete IgA in their colostrum. Although colostral IgM levels may be elevated in an attempt to compensate for the lack of maternal IgA, the newborn remains relatively unprotected against intestinal pathogens. Of greater concern is the fact that the babies of mothers with untreated CVID are born in a state of humoral immunodeficiency and thus are at risk for life-threatening sinopulmonary infection. To compensate for the loss of IgG across the placenta and to provide the infant with the passive immunity that will be required, the dose of replacement gammaglobulin therapy should be increased by 50% during the third trimester of pregnancy.

Splenomegaly is common in untreated patients. Hypersplenism in most patients responds to aggressive therapy with antibiotics and IVIG. The presumption is that the hypersplenism is secondary to reactive hyperplasia of lymphoid follicles within the spleen attempting to respond to infection. Development of esophageal varices or other hematological manifestations of hypersplenism (refractory thrombocytopenia, anemia, neutropenia, and lymphopenia) may require splenectomy as the therapy of last resort. The outcome for most such patients has been good, with resolution of symptoms, although patients with altered TACI alleles tend to do less well.

The development of a constellation of pulmonary abnormalities that include granulomatous and lymphoproliferative histopathologic patterns (lymphocytic interstitial pneumonia [LIP], follicular bronchiolitis, and lymphoid hyperplasia), termed granulomatous–lymphocytic interstitial lung disease (GLILD), can be an ominous sign. These patients appear more likely to develop granulomatous liver disease, autoimmune hemolytic anemia, lymphoproliferative disease, and progressive pulmonary disease.[37]

THERAPEUTIC PRINCIPLES

- The primary goal of treatment is to keep the patient infection free.
- In patients whose respiratory mucosa is intact, intravenous or subcutaneous replacement immunoglobulin G (IgG) therapy is generally effective in protecting the patient from pulmonary infections.
- For those patients who have developed bronchiectasis or who continue to subject themselves to environmental toxins (*e.g.,* smoking), replacement IgG will ameliorate but may not prevent all such infections.
- Because mucosal Ig cannot be replaced, even patients on adequate IgG replacement therapy remain at risk for sinus or gastrointestinal infections.
- Prophylactic antibiotics that are effective against encapsulated organisms can significantly reduce the frequency of upper respiratory tract infections in patients who continue to suffer in spite of replacement therapy with intravenous immunoglobulin (IVIG).
- Prolonged diarrhea in patients with hypogammaglobulinemia is often caused by *Giardia lamblia* and responds well to metronidazole therapy.
- Patients with primary antibody deficiencies should not receive live vaccines.

SELECTIVE IGG SUBCLASS DEFICIENCIES

Diagnosis

A diagnosis of clinical immunodeficiency should be supported by clear evidence of functional impairment. Most individuals with modest reductions in serum IgG subclass levels are functionally normal. Indeed, individuals with deletions of the heavy chain Ig gene locus, some of whom completely lack IgG1, IgG2, IgG4, and IgA[38] have been reported to be asymptomatic.

The diagnosis of a functional IgG subclass deficiency can thus be made with confidence only when there is both a significant decrease in the serum concentration of a specific isotype and there is clear evidence of abnormal specific antibody production. Up to 10% of normal males and 1% of normal females have IgG4 deficiency, which makes a diagnosis of immunodeficiency as a result of an isolated IgG4 subclass deficiency problematic. Among patients with a deficiency of IgG1 or IgG3, documentation of the ability to produce protective titers of antitetanus toxin and antidiphtheria toxin antibodies following standard tetanus toxoid and diphtheria immunizations is a strong indication that replacement gammaglobulin therapy is likely unwarranted. Similarly, documentation of a strong antipneumococcal polysaccharide response in patients with an apparent IgG2 deficiency would suggest gammaglobulin replacement is likely not required. IgG2 levels normally begin to rise in childhood later than other subclasses. Conversely, the lack of a response to vaccination calls for appropriate prophylactic antibiotic therapy before a trial of IVIG is attempted.

Clinical Manifestations

The clinical spectrum of isolated IgG subclass deficiency is quite variable, and deficiencies of each of the four IgG subclasses have been described. Some individuals are referred to the clinical immunologist with only a mild reduction of total IgG, but most symptomatic patients have marked deficiencies of one or more IgG subclass despite normal total IgG concentrations. Since IgG1 makes up the majority of serum IgG in most patients, a deficiency of IgG1 tends to correlate with depressed serum levels of total IgG.

Determination of IgG subclasses is rarely performed on asymptomatic individuals; thus most patients with an isolated IgG2 deficiency come to medical attention as a result of recurrent sinusitis, otitis media, or pulmonary infections. Individuals may have few residual symptoms between infections, but some have severe chronic inflammation with refractory sinusitis, pulmonary fibrosis, or bronchiectasis. Because protective antibodies directed against carbohydrate antigens are usually of the IgG2 subclass, many affected patients exhibit an impairment of their ability to mount specific protective responses to encapsulated pathogens. However, normal responses have also been described.[39] Many clinicians would agree, however, that patients with IgG2 deficiency, who suffer from recurrent sinopulmonary infections and who respond to less than half of the polysaccharide antigens with which they have been challenged, meet the standard for functional immune deficiency and thus warrant aggressive prophylactic therapy up to and including Ig replacement should the infections be severe.

IgG3 deficiency can occur alone or in association with IgG1 deficiency. Recurrent infection of the respiratory tract with chronic lung disease has been reported. With a serum half-life only 2 weeks, IgG3 levels may be consumed rapidly during the course of an active infection in an otherwise normal individual.[40]

Before making the diagnosis of IgG3 deficiency, serum levels of IgG3 should be rechecked when the individual is asymptomatic.

Compared with the serum, IgG4 is overrepresented in secretions, and IgG4-committed B cells are present at mucosal sites, suggesting a role in mucosal immunity. Since IgG4 is normally present in the serum in very low concentrations, the significance of a low serum level in a patient with recurrent infection remains unclear.

Origin and Pathogenesis

The origin of IgG subclass deficiency is unknown. Homozygous deletions of portions of the Ig heavy chain constant locus associated with total absence of IgG2, IgG3, and IgG4 or combinations of these isotypes have been described in healthy individuals. IgG2 deficiency is often found in association with selective IgA deficiency with or without IgG4 deficiency, and patients with selective IgG subclass deficiencies have been shown to have inherited the same MHC haplotypes as those who suffer from IgAD and CVID. These observations suggest that patients with recurrent infections have a more complex defect than the mere elimination of one or more IgG isotype. In some instances, subclass deficiency is associated with a T-cell defect, as in chronic mucocutaneous candidiasis and ataxia–telangiectasia. IgG subclass deficiency can be acquired. Acute infections, medications, chemotherapy, irradiation, surgery, and human immunodeficiency virus (HIV) infection have all been temporally linked to the development of a deficiency in one or more IgG subclass.[41]

Treatment and Prognosis

The natural history of IgG subclass deficiency plus or minus IgA deficiency, especially in children, is not constant.[42] Associated allergic rhinosinusitis and asthma must be aggressively treated with conventional therapy for these disorders, as these conditions increase the risk of purulent sinusitis and pneumonia. Causes of anatomical obstruction should be sought when persistent infection of a sinus or pulmonary segment is the presenting complaint; the role of surgical therapy for anatomical obstruction should not be overlooked.

Many patients with IgG subclass deficiency do well on prophylactic antibiotics and will never need Ig supplementation. However, Ig replacement therapy can be beneficial in patients with severe, recurrent infections. Patients who begin therapy should improve within the first 2 months, but to avoid the placebo effect, a full 6-month trial is recommended.

ANTIBODY DEFICIENCY WITH NORMAL SERUM IMMUNOGLOBULIN LEVELS

Occasional patients may present with normal serum Ig concentrations and a selective inability to respond to infections with pyogenic organisms. Diagnosis requires documentation of an inability to respond to antigenic challenge. These patients may respond to replacement Ig therapy. The antibody response to specific polysaccharide antigens can be very selective. In human, most protective anti-*H. influenzae* type b (anti-Hib) antibodies utilize the rare *VκA2* gene. The Navajo population in the Southwestern United States suffers a 5- to 10-fold increased incidence of Hib disease. This population also exhibits a high prevalence of an A2 allele with a defective recombination signal sequence, preventing use of germline-encoded antibodies that can generate protective antigen-binding sites.[43]

Analysis of a group of well-characterized patients, mostly female, with a history of RESPIs and normal serum Ig levels revealed a high prevalence of the same MHC haplotypes observed in IgAD, selective IgG subclass deficits, and CVID.[23] These patients tend to respond to aggressive antibiotic therapy, including prophylaxis.

SELECTIVE LIGHT-CHAIN DEFICIENCY

Selective deficiencies of either κ or λ light chains have been reported.[44-46] In one such case, the patient was the offspring of a consanguineous (uncle–niece) union; and in the second, a molecular analysis demonstrated different loss of function mutations in the patient's Cκ alleles. The parents of these children had no health difficulties, but each of the patients required medical attention for recurrent sinopulmonary infections and diarrhea. Two of the patients with κ deficiency exhibited IgA deficiency, and the remaining patients with κκ deficiency and λ deficiency had panhypogammaglobulinemia.

TRANSIENT HYPOGAMMAGLOBULINEMIA OF INFANCY

Diagnosis

As infants make the transition from dependence on maternal Ig to reliance on endogenously produced antibodies, they suffer a physiological nadir of serum Ig at 4–6 months of age, a period associated with susceptibility to mild upper respiratory infections and otitis media (see Fig. 34.2). Children who *(i)* exhibit serum concentrations of one or more of the three major Ig classes that fall below the 95% confidence interval (CI) for age on ≥2 occasions during infancy, *(ii)* demonstrate a rise in these values to or toward normal over time, and *(iii)* lack features consistent with other forms of primary immunodeficiency fall within the catch-all diagnosis of transient hypogammaglobulinemia of infancy (THI).[47,48] By definition, the diagnosis of THI can be made with certainty only in retrospect.

Clinical Manifestations

Ig concentrations are rarely measured in infants unless there is some reason to suspect an immunodeficiency. Most patients with this diagnosis come to medical attention because of either recurrent infections or as a result of routine screening studies of relatives of other patients with immunodeficiency. Yet bearing in mind that 2.5% of normal infants will fall below the 95% CI range at any one time, the diagnosis of THI is remarkably rare. Among two major centers, one in the United States and one in Germany, the diagnosis was given to only 16 of 18 000 children in whom the index of suspicion warranted Ig determinations.[49,50]

Patients with THI typically are able to synthesize specific antibodies in response to immunization with T-dependent antigens (*e.g.,* tetanus and diphtheria toxoids).[51] They may have difficulty, however, responding to polysaccharide antigens (*e.g.,* isohemagglutinins and vaccination with Pneumovax23). Some will fail to sustain protective antibody responses to antigens. Most patients with THI, especially those ascertained as a result of family studies or mild upper respiratory infections alone, exhibit fewer infections over time. The great majority of infants with THI will normalize their serum Ig levels within 24 months of age. However, a minority fails to normalize IgG, continues to suffer with recurrent infections, and may develop evidence of autoimmune disease. These patients often become part of the hypogammaglobulinemia syndrome complex that includes CVID and may ultimately require long-term therapy with gammaglobulin, prophylactic antibiotics, or both.

Treatment and Prognosis

Children with suspected THI should be monitored with serial determination of serum Igs and isohemagglutinin titers to confirm acquisition of normal immune function. Some children will not achieve normal levels of IgG for several years, and some will remain IgG subclass or IgA deficient. Treatment of THI with IVIG is generally not warranted unless the child suffers from persistent, recurrent, invasive infections, including pneumonia.

ON THE HORIZON

- Elucidation of the molecular basis of selective defects in humoral responses to pathogens, in part through the use of high-throughput sequencing to characterize the precise molecular composition of antibody responses in immune deficiencies
- Further elucidation of the molecular basis of common variable immune deficiency, hypogammaglobulinemia, and immunoglobulin A (IgA) deficiency

FRONTIERS IN RESEARCH

Bruton reported the first case of agammaglobulinemia in 1952, as well as the first successful therapy for this classic primary antibody deficiency. There has been remarkable progress in the identification of single-gene disorders since that time. However, for the majority of patients the underlying pathogenesis of the most common manifestation of primary antibody deficiency in our population, that is, hypogammaglobulinemia in the adult, still remains unclear. It seems increasingly likely that this disorder is multifactorial in nature, dependent on the inheritance of one or more susceptibility loci in association with either environmental influences or random chance. In rare cases, patients with hypogammaglobulinemia have shown resolution of their symptoms, suggesting that a better understanding of pathogenesis might yield therapies of remission. The molecular basis of selective deficiencies in the response to pathogens in the presence of normal serum Ig levels also remains unclear. The availability of whole exome sequencing and whole-genome sequencing has helped identify the genetic cause of an increasing number of patients with presumed CVID. More genes and likely presumptive alleles will be found in the next decade.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Durandy A, Kracker S, Fischer A. Primary antibody deficiencies. Nat Rev Immunol 2013;13:519–33.
2. Johnson ML, Keeton LG, Zhu ZB, et al. Age-related changes in serum immunoglobulins in patients with familial IgA deficiency and common variable immunodeficiency (CVID). Clin Exper Immunol 1997;108:477–83.
3. Litzman J, Freiberger T, Grimbacher B, et al. Mannose-binding lectin gene polymorphic variants predispose to the development of bronchopulmonary complications but have no influence on other clinical and laboratory symptoms or signs of common variable immunodeficiency. Clin Exper Immunol 2008;153:324–30.
4. De Greef GE, van Tol MJ, Van Den Berg JW, et al. Serum immunoglobulin class and IgG subclass levels and the occurrence of homogeneous immunoglobulins during the course of ageing in humans. Mech Ageing Dev 1992;66:29–44.
5. Stiehm ER, Fudenberg HH. Serum levels of immune globulins in health and disease: a survey. Pediatrics 1966;37:715–27.
6. Soothill JF, Hayward AR, Wood CB. Pediatric immunology. Oxford, UK: Blackwell Scientific Publications; 1983:463–4.
7. Gharib A, Caperton C, Gupta S. Anaphylaxis to IGIV in immunoglobulin-naive common variable immunodeficiency patient in the absence of IgG anti-IgA antibodies: successful administration of low IgA-containing immunoglobulin. Allergy Asthma Clin Immunol 2016;12:23.
8. Carr TF, Koterba AP, Chandra R, et al. Characterization of specific antibody deficiency in adults with medically refractory chronic rhinosinusitis. Am J Rhinol Allergy 2011;25:241–4.
9. Lopez-Herrera G, Vargas-Hernandez A, Gonzalez-Serrano ME, et al. Bruton's tyrosine kinase—an integral protein of B cell development that also has an essential role in the innate immune system. J Leukoc Biol 2014;95:243–50.
10. Nelson KS, Lewis DB. Adult-onset presentations of genetic immunodeficiencies: genes can throw slow curves. Curr Opin Infect Dis 2010;23:359–64.
11. Sigmon JR, Kasasbeh E, Krishnaswamy G. X-linked agammaglobulinemia diagnosed late in life: case report and review of the literature. Clin Mol Allergy 2008;6:5.
12. Stewart DM, Tian L, Notarangelo LD, et al. X-linked hypogammaglobulinemia and isolated growth hormone deficiency: an update. Immunol Res 2008;40:262–70.
13. Winkelstein JA, Conley ME, James C, et al. Adults with X-linked agammaglobulinemia: impact of disease on daily lives, quality of life, educational and socioeconomic status, knowledge of inheritance, and reproductive attitudes. Medicine (Baltimore) 2008;87:253–8.
14. Conley ME, Dobbs AK, Farmer DM, et al. Primary B cell immunodeficiencies: comparisons and contrasts. Annu Rev Immunol 2009;27:199–227.
15. Boisson B, Wang YD, Bosompem A, et al. A recurrent dominant negative E47 mutation causes agammaglobulinemia and BCR(−) B cells. J Clin Invest 2013;123:4781–5.
16. Sawada A, Takihara Y, Kim JY, et al. A congenital mutation of the novel gene LRRC8 causes agammaglobulinemia in humans. J Clin Invest 2003;112:1707–13.
17. Conley ME, Dobbs AK, Quintana AM, et al. Agammaglobulinemia and absent B lineage cells in a patient lacking the p85alpha subunit of PI3K. J Exp Med 2012;209:463–70.
18. Durandy A, Kracker S. Immunoglobulin class-switch recombination deficiencies. Arthritis Res Ther 2012;14:218.
19. Karaca NE, Durandy A, Gulez N, et al. Study of patients with Hyper-IgM type IV phenotype who recovered spontaneously during late childhood and review of the literature. Eur J Pediatr 2011;170:1039–47.
20. Kawai T, Nishikomori R, Heike T. Diagnosis and treatment in anhidrotic ectodermal dysplasia with immunodeficiency. Allergol Int 2012;61:207–17.
21. Hennig C, Baumann U, Ilginus C, et al. Successful treatment of autoimmune and lymphoproliferative complications of patients with intrinsic B-cell immunodeficiencies with Rituximab. Br J Haematol 2010;148:445–8.
22. Schroeder HW Jr, Schroeder HW III, Sheikh SM. The complex genetics of common variable immunodeficiency. J Investig Med 2004;52:90–103.
23. Johnston DT, Mehaffey G, Thomas J, et al. Increased frequency of HLA -B44 in Recurrent Sino-Pulmonary Infections (RESPI). Clin Immunol 2006;119:346–50.
24. Nielsen LK, Dziegiel MH. Recombinant human immunoglobulin (Ig)A1 and IgA2 anti-D used for detection of IgA deficiency and anti-IgA. Transfusion 2008;48:1892–7.
25. Borrelli M, Maglio M, Agnese M, et al. High density of intraepithelial gammadelta lymphocytes and deposits of immunoglobulin (Ig)M anti-tissue transglutaminase antibodies in the jejunum of coeliac patients with IgA deficiency. Clin Exper Immunol 2010;160:199–206.

26. Bonilla FA, Barlan I, Chapel H, et al. International Consensus Document (ICON): common variable immunodeficiency disorders. J Allergy Clin Immunol Pract 2016;4:38–59.
27. Yu JE, Knight AK, Radigan L, et al. Toll-like receptor 7 and 9 defects in common variable immunodeficiency. J Allergy Clin Immunol 2009;124:349–56, 56 e1–3.
28. Salzer U, Chapel HM, Webster ADB, et al. Mutations in TNFRSF13B encoding TACI are associated with common variable immunodeficiency in humans. Nat Genet 2005;37:820–8.
29. Wang Y, Hwangpo TA, Martin MP, et al. Killer cell immunoglobulin-like receptors are associated with common variable immune deficiency pathogenesis. J Allergy Clin Immunol 2016;138:1495–8.
30. Bogaert DJ, Dullaers M, Lambrecht BN, et al. Genes associated with common variable immunodeficiency: one diagnosis to rule them all? J Med Genet 2016;53:575–90.
31. Sathkumara HD, De Silva NR, Handunnetti S, et al. Genetics of common variable immunodeficiency: role of transmembrane activator and calcium modulator and cyclophilin ligand interactor. Int J Immunogenet 2015;42:239–53.
32. Zhang L, Radigan L, Salzer U, et al. Transmembrane activator and calcium-modulating cyclophilin ligand interactor mutations in common variable immunodeficiency: clinical and immunologic outcomes in heterozygotes. J Allergy Clin Immunol 2007;120:1178–85.
33. Yong PF, Thaventhiran JE, Grimbacher B. "A rose is a rose is a rose," but CVID is Not CVID common variable immune deficiency (CVID), what do we know in 2011? Adv Immunol 2011;111:47–107.
34. Cheon CK, Ko JM. Kabuki syndrome: clinical and molecular characteristics. Korean J Pediatr 2015;58:317–24.
35. Malamut G, Verkarre V, Suarez F, et al. The enteropathy associated with common variable immunodeficiency: the delineated frontiers with celiac disease. Am J Gastroenterol 2010;105:2262–75.
36. Sneller MC, Strober W, Eisenstein E, et al. NIH conference. New insights into common variable immunodeficiency. Ann Intern Med 1993;118:720–30.
37. Chase NM, Verbsky JW, Hintermeyer MK, et al. Use of combination chemotherapy for treatment of granulomatous and lymphocytic interstitial lung disease (GLILD) in patients with common variable immunodeficiency (CVID). J Clin Immunol 2013;33:30–9.
38. Lefranc MP, Hammarstrom L, Smith CI, et al. Gene deletions in the human immunoglobulin heavy chain constant region locus: molecular and immunological analysis. Immunodefic Rev 1991;2:265–81.
39. Shackelford PG, Granoff DM, Polmar SH, et al. Subnormal serum concentrations of IgG2 in children with frequent infections associated with varied patterns of immunologic dysfunction. J Pediatr 1990;116:529–38.
40. Tabata N, Azuma E, Masuda S, et al. Transient low level of IgG3 induced by sepsis. Acta Paediatr Jpn 1995;37:201–2.
41. Morell A. IgG subclass deficiency: a personal viewpoint. Pediatr Infect Dis J 1990;9:S4–8.
42. Kutukculer N, Karaca NE, Demircioglu O, et al. Increases in serum immunoglobulins to age-related normal levels in children with IgA and/or IgG subclass deficiency. Pediatr Allergy Immunol 2007;18:167–73.
43. Feeney AJ, Atkinson MJ, Cowan MJ, et al. A defective Vkappa A2 allele in Navajos which may play a role in increased susceptibility to haemophilus influenzae type b disease. J Clin Invest 1996;97:2277–82.
44. Bernier GM, Gunderman JR, Ruymann FB. Kappa-chain deficiency. Blood 1972;40:795–805.
45. Barandun S, Morell A, Skvaril F, et al. Deficiency of kappa- or lambda-type immunoglobulins. Blood 1976;47:79–89.
46. Stavnezer-Nordgren J, Kekish O, Zegers BJ. Molecular defects in a human immunoglobulin kappa chain deficiency. Science 1985;230:458–61.
47. Stiehm ER. The four most common pediatric immunodeficiencies. J Immunotoxicol 2008;5:227–34.
48. Moschese V, Graziani S, Avanzini MA, et al. A prospective study on children with initial diagnosis of transient hypogammaglobulinemia of infancy: results from the Italian Primary Immunodeficiency Network. Int J Immunopathol Pharmacol 2008;21:343–52.
49. Tiller TL Jr, Buckley RH. Transient hypogammaglobulinemia of infancy: review of the literature, clinical and immunologic features of 11 new cases, and long-term follow-up. J Pediatr 1978;92:347–53.
50. Dressler F, Peter HH, Muller W, et al. Transient hypogammaglobulinemia of infancy: five new cases, review of the literature and redefinition. Acta Paediatr Scand 1989;78:767–74.
51. Tiller TL Jr, Buckley RH. Transient hypogammaglobulinemia of infancy: review of the literature, clinical and immunologic features of 11 new cases, and long-term follow-up. J Pediatr 1978;92:347–53.

MULTIPLE-CHOICE QUESTIONS

1. A 50-year-old man presents with a history of two pneumonias in the last 2 years. He has suffered occasionally with sinusitis and bronchitis since childhood, but these infections have readily responded to antibiotic therapy, and he denies dyspnea on exertion. A grandson is receiving therapy for an unknown immune deficiency. Serum immunoglobulin (Ig) levels are drawn. IgM is <20 g/dL. IgG is 100 mg/dL. IgA is <7 mg/dL. Flow cytometry reveals rare B cells. He has no response to pneumococcal polysaccharide vaccination. What is the most likely diagnosis?
   - **A.** Hyper-IgM syndrome
   - **B.** Common variable immune deficiency
   - **C.** X-linked agammaglobulinemia
   - **D.** Autosomal agammaglobulinemia

2. A 35-year-old woman presents with a 10-year history of recurrent sinopulmonary infections. She has suffered with allergic rhinitis since childhood and was diagnosed with celiac disease at age 15 years. Serum immunoglobulin (Ig) levels are drawn. IgM is 100 g/dL. IgG is 650 mg/dL. IgA is <7 mg/dL. Flow cytometry reveals normal B-cell numbers. After vaccination with Pneumovax23, she responds to only five of 23 pneumococcal polysaccharide antigens. She has never smoked, and testing for cystic fibrosis is negative. She complains of dyspnea on climbing more than one flight of stairs and suffers with fatigue. She suffered pneumonia once as a child. What would be the best approach?
   - **A.** The patient has selective IgA deficiency. She should be treated with antibiotics, as needed, for infections.
   - **B.** The patient may have an IgG subclass deficiency in addition to IgA deficiency. She should be treated with prophylactic antibiotics, since gammaglobulin replacement is contraindicated for patients with IgA deficiency.
   - **C.** The patient may have an IgG subclass deficiency in addition to IgA deficiency. She should be treated with gammaglobulin replacement therapy.

3. A 2-month-old baby girl presents with her second episode of bacterial pneumonia. Serum immunoglobulin (Ig) levels are drawn. IgM is 100 g/dL. IgG is 100 mg/dL. IgA is <7 mg/dL. Flow cytometry reveals normal B-cell numbers. What is the most likely diagnosis?
   - **A.** Selective IgA deficiency
   - **B.** Transient hypogammaglobulinemia of infancy
   - **C.** Common variable immune deficiency
   - **D.** Maternal common variable immune deficiency

35

Primary T-Cell Immunodeficiencies

Chaim M. Roifman

This chapter is dedicated to the diagnosis and management of inherited immune defects associated predominantly with T-cell dysfunction. One of the central roles of T cells is the coordination of both the humoral and cell-mediated arms of the immune system; therefore T-cell deficiencies are often accompanied by dysfunction of other cell types, such as B and natural killer (NK) cells. The focus of this chapter is on primary immunodeficiencies that are profound enough to require hematopoietic stem cell transplantation (HSCT) or other modalities of management and treatment directed at T-cell absence or dysfunction.

T-CELL IMMUNODEFICIENCIES

KEY CONCEPTS

T-Cell Immunodeficiency

- A group of disorders are caused by defects in genes critical to the growth, maturation, and survival of T lymphocytes.
- Some defects may also affect other arms of the immune system, such as B cells, natural killer (NK) cells, or components of innate immunity.
- The immune defects may be one feature of a multisystem syndrome.
- The age of presentation and diagnosis varies widely from early infancy (or at birth when detected through newborn screening) to adulthood.
- The spectrum of clinical manifestations is wide and encompasses presentation at early infancy with extreme susceptibility to opportunistic infections (severe combined immunodeficiency [SCID]) or disorders featuring predominately autoimmunity or predisposition to cancer (combined immunodeficiency [CID]).

SEVERE COMBINED IMMUNODEFICIENCY

KEY CONCEPTS

Severe Combined Immunodeficiency

- Infants present typically at 4–6 months of age with severe persistent infections, oral thrush, chronic diarrhea, and failure to thrive. Siblings of affected patients and cases detected through newborn screening are diagnosed before symptoms occur.
- By definition, affected infants have a permanent profound T-cell lymphopenia of <300–500 CD3 cells/μL, with or without reduced numbers of circulating B cells and natural killer (NK) cells.
- Invariably, *in vitro* responses to T-cell mitogens are depressed and T-cell receptor excision circle (TREC) levels are low or undetectable.
- Ideally, patients with severe combined immunodeficiency (SCID) should be placed in protective isolation and offered hematopoietic stem cell transplantation (HSCT) (or, in some cases, gene therapy).

The most clinically severe and studied group of patients are those traditionally designated as having severe combined immunodeficiency (SCID). The clinical hallmarks of SCID include consistent clinical presentation in the first year of life, with repeated life-threatening bacterial, viral, or fungal infections, chronic diarrhea, and failure to thrive.[1] Evaluation of the immune system typically reveals profound T-cell lymphopenia of <500 CD3 cells,[2,3] or according to more stringent definitions, <300 T cells. Invariably, *in vitro* responses to mitogens are severely depressed and T-cell receptor excision circle (TREC) numbers are markedly reduced or below detection.[4] The identification of these cases by TREC-based newborn screening (NBS) is the earliest alert of T-cell deficiency.[5]

COMBINED IMMUNODEFICIENCY

The second large group of disorders with significant T-cell deficiency is loosely designated as combined immunodeficiency, which is arbitrarily defined by the presence of >300–500 autologous CD3 cells/μL in the circulation.[3,4] The omission of "severe" from this category of patients might be misleading, as many of these patients present with signs and symptoms indistinguishable from SCID. This category of heterogeneous deficiencies can be further subdivided into subgroups that are discussed in detail in this chapter. They include T-cell deficiencies that invariably or frequently present with normal or close to normal numbers of circulating lymphocytes, such as zeta-associated protein-70 (ZAP-70) deficiency. Other disorders are caused by hypomorphic mutations in genes associated with SCID[3] and are therefore frequently referred to as "leaky SCIDs." A third subgroup encompasses patients with multisystem syndromes that include variable degrees of T-cell deficiency, such as cartilage hair hypoplasia and DiGeorge syndrome. Finally, a fourth group involves more selective T-cell defects, such as FOXP3 deficiency, which predispose patients to a limited array of infections or are associated with autoimmunity or lymphoid malignancies. There may be a wide range of clinical manifestations, from typical SCID presentation to mild or delayed presentation. Laboratory features are also extremely variable but universally reflect T-cell dysfunction.[6,7]

OMENN SYNDROME

Omenn syndrome (T^+B^+ SCID; OMIM #603554) is an aberrant inflammatory process associated with "leaky SCID" phenotypes, which characteristically allow for thymic escape and propagation of poorly controlled T-cell clones.[8] T-cell development is severely

KEY CONCEPTS

Combined Immunodeficiency

- Arbitrarily defined as a T-cell deficiency with >300–500 autologous CD3 circulating cells/µL, which may be reduced or dysfunctional.
- Circulating autologous T cells may be the typical feature of the defect or may be caused by genetic variants leading to an incomplete defect (leaky severe combined immunodeficiency [SCID]).
- Clinical manifestations may be indistinguishable from SCID.
- Others may present with SCID and aberrant inflammatory syndrome (Omenn syndrome) consisting of erythroderma, lymphadenopathy, and eosinophilia.
- Frequently, the predominant presenting manifestation can be autoimmunity or lymphoid malignancy.
- *In vitro* responses to mitogens and T-cell receptor excision circle (TREC) levels may be depressed or normal. The T-cell repertoire is often abnormal, as are T-cell subpopulations (regulatory T cells [Tregs] and memory cells). Immunoglobulins and specific antibodies are variable.

restricted because of inefficient deletion and suppression of self-reactive T cells and dysregulated lymphostromal cross-talk in the thymus. The condition is associated with hypomorphic mutations of various SCID-causing genes, including *RAG1/2, Artemis, DNA ligase IV, RMRP, ZAP-70, IL2RG, IL7Rα, CD3δ,* and *ADA* (Table 35.1). In addition to typical SCID manifestations, patients present with features including generalized erythroderma, which may be accompanied by alopecia, loss of eyebrows and eyelashes. Marked enlargement of lymph nodes, liver, and spleen are frequently present.

The hallmarks of laboratory findings are elevated serum immunoglobulin E (IgE) levels and increased eosinophil counts. T-cell number may be normal or even elevated, but the T-cell repertoire is skewed with overrepresentation of a few T-cell clones and underexpression or absence of most other Vβ families[9] (Chapter 8). Skin biopsy reveals acanthosis and parakeratosis. Dyskeratosis and spongiosis are seen in the Malpighian layer, and vacuolization is often observed in the basal layer. The thymus is dysplastic, and expression of the transcription factor autoimmune regulator (AIRE) in medullary thymic epithelial cells is reduced.[10]

TABLE 35.1 T-Cell Immunodeficiency With Autologous T Cells Associated With Omenn Syndrome

Gene Defect	Immune Phenotype	Additional Features
RAG1	$T^+B^-NK^+$	None
RAG2	$T^+B^-NK^+$	None
Artemis	$T^+B^-NK^+$	Irradiation sensitivity
DNA ligase IV	$T^+B^-NK^+$	Irradiation sensitivity, microcephaly
Adenosine deaminase	$T^+B^{-/+}NK^{-/+}$	Multisystem involvement
ZAP-70	$T^+B^+NK^+$	None
RMRP	$T^+B^+NK^-$	Short stature, metaphyseal dysplasia
γc (IL-2Rγ)	$T^+B^+NK^-$	None
IL-7Rα	$T^+B^+NK^+$	None
22q11.2 microdeletion	$T^+B^+NK^+$	DiGeorge syndrome
CD3δ	$T^+B^+NK^+$	None

T^+ = >500 CD3 cells/µL

Regardless of the molecular cause, treatment with corticosteroids and cyclosporine are effective in controlling inflammation, improving skin lesions, and reducing lymphadenopathy and hepatosplenomegaly. Ultimately, a cure can be achieved only with HSCT.[11]

SEVERE COMBINED IMMUNE DEFICIENCY

SCID is characterized by near complete absence of circulating autologous T cells and impaired B-cell function. Defects in the differentiation and maturation of T-lineage cells in the thymus are a primary cause of T-cell deficiency (Fig. 35.1). In particular, mutations affecting T-cell activation or downstream signaling molecules can result in SCID or CID (Fig. 35.2).

SCID With T-Cell Lymphopenia (T^-B^+)

This T^-B^+ group of patients (Table 35.2) is the most frequently identified phenotype of SCID (28 to 50% of all cases) and is believed to occur at a frequency of 1:50,000 live births.[1]

γc (IL-2Rγ) Deficiency

Common gamma chain (γc) or interleukin-2 receptor gamma (IL-2Rγ) deficiency (OMIM #300400) is transmitted in an X-linked manner. Affected males have mutations in the IL-2Rγ chain, also known as the *common gamma chain* because of its critical role as a coreceptor in IL-4, IL-7, IL-9, IL-15, and IL-21 receptors. Multiple pathways downstream of these receptors are essential for differentiation and growth of T cells and NK cells, hence the abnormal immune profile and profound defect in T-cell and NK-cell maturation. Affected males typically present at 3–6 months of age, with failure to thrive and repeated microbial and fungal infections. Infections are fatal unless HSCT is administered.

Evaluation of the immune system typically reveals extremely low to absent T and NK cells, while sparing B cells ($T^-B^+NK^-$). *In vitro* responses to mitogens are profoundly depressed, TRECs are undetectable, and the thymus appears dysplastic with absence of Hassall corpuscles and disrupted architecture.

γc deficiency may have an atypical presentation because of hypomorphic mutations. Various degrees of autologous T-cell or NK-cell levels have been reported. Importantly, patients with the R222C mutation can present with SCID-like symptoms but have normal numbers of T cells and NK cells, near normal responses to mitogens, normal TREC levels, and structurally normal thymus gland.[12] Cure can be achieved by providing HSCT. Gene therapy has recently been attempted but remains an experimental procedure.

JAK3 Deficiency

Janus kinase-3 (JAK3) deficiency (OMIM #600802) is an inherited autosomal recessive mutation mimicking the presentation of γc deficiency. This is expected, as JAK3, a lymphoid specific tyrosine kinase, is a signaling molecule downstream of the γc receptor. Upon IL-2 binding, JAK3 phosphorylates signal transducer and activator of transcription (STAT) factors, which dimerize and translocate to the nucleus, where they bind regulatory elements, thus inducing target genes.

Patients typically present with SCID features, but similar to γc, patients with JAK3 deficiency may also present with combined immunodeficiencies and autologous T cells. Milder phenotypes of T-cell deficiency caused by hypomorphic mutations in JAK3 have also been described. Some patients have been diagnosed during

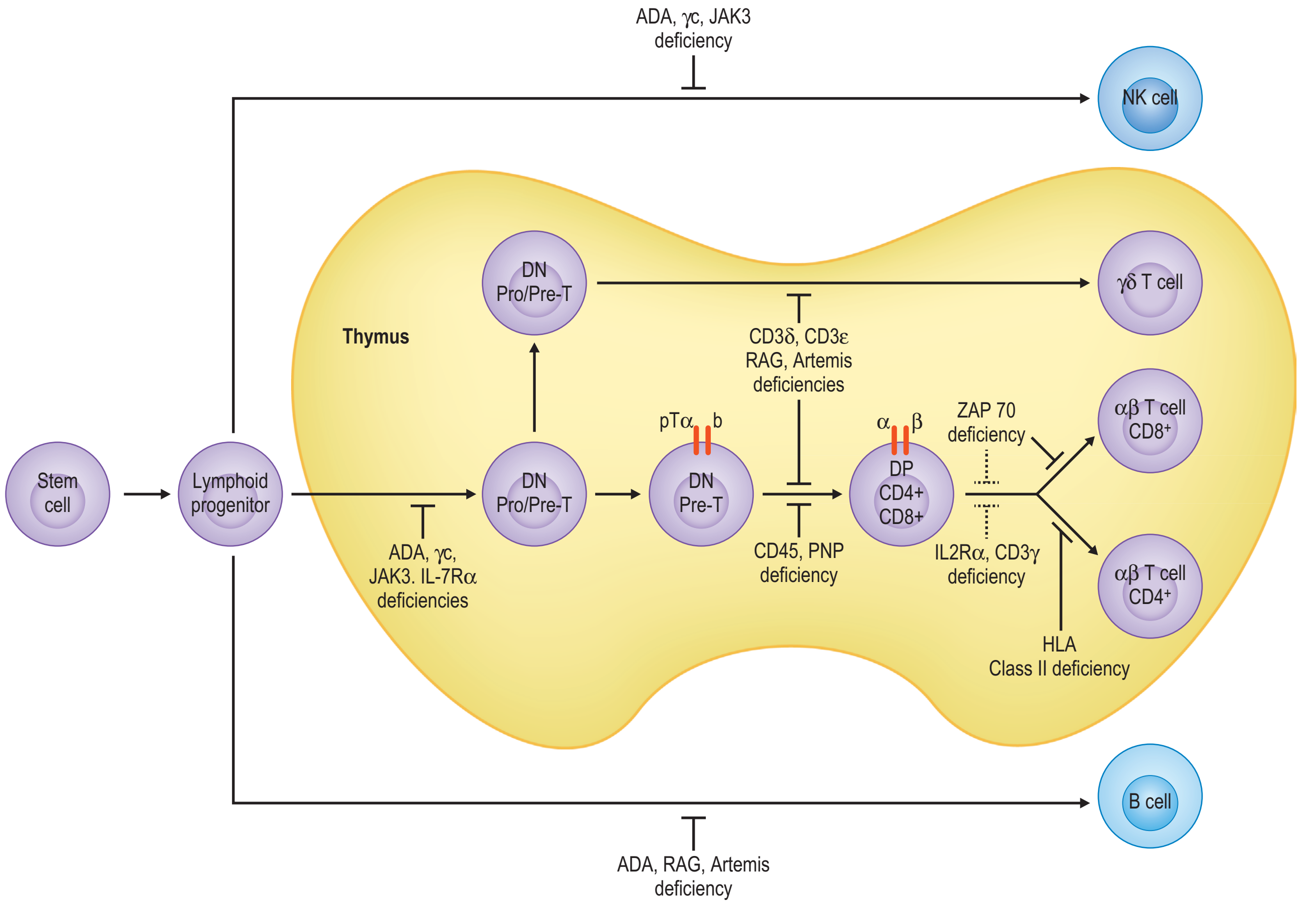

ADA, γc, JAK3
deficiency
NK cell
Thymus
DN
Pro/Pre-T
γδ T cell
CD3δ, CD3ε
RAG, Artemis
deficiencies
pTα
b
α
β
ZAP 70
deficiency
αβ T cell
CD8+
Stem
cell
Lymphoid
progenitor
DN
Pro/Pre-T
DN
Pre-T
DP
CD4+
CD8+
ADA, γc,
JAK3. IL-7Rα
deficiencies
CD45, PNP
deficiency
IL2Rα, CD3γ
deficiency
αβ T cell
CD4+
HLA
Class II deficiency
B cell
ADA, RAG, Artemis
deficiency

FIG 35.1 **Impaired Maturation of T-Lineage Cells in the Thymus Can Lead to Profound T-Cell Deficiency.** Lymphoid progenitor cells originating from bone marrow reach the thymus and mature into functional T cells. During differentiation, naïve thymocytes (DN, double-negative cells) undergo β chain rearrangement and pairing of the β chain with the pre-Tα chain, giving rise to cells expressing both CD4 and CD8 (DP, double-positive cells). After positive and negative selection, thymocytes downregulate the expression of either CD4 or CD8, resulting in single-positive cells. These cells emigrate from the thymus to blood and secondary lymphatic organs, where they encounter antigens and exert their biological functions. Various defects in this process can cause profound T-cell deficiencies. Blocked lines indicate specific stages impaired by known genetic abnormalities. Some defects may also interfere with maturation of B and natural killer (NK) cells. ADA, adenosine deaminase deficiency; γc- gamma common chain of the interleukin (IL)-2 receptor; JAK3, Janus activated kinase 3; IL7Rα, alpha chain of the IL7 receptor; DN, double-negative; RAG, recombination activating genes; PNP, purine nucleoside phosphorylase; ZAP-70, Zeta chain associated protein kinase 70; IL2Rα, alpha chain of the IL2 receptor.

late childhood (age 6 years), with some remaining asymptomatic. Additionally, patients may present with lymphoproliferative disorder, whereas others remain asymptomatic, despite carrying the same mutation.

IL-7Rα Deficiency

IL-7 receptor α (IL-7Rα) deficiency is transmitted in an autosomal recessive manner and is caused by mutations in the IL-7Rα chain. Together, IL-7Rα chain and γc forms the IL-7R (OMIM*146661), which is essential for the normal development of T cells. This receptor mediates survival and proliferative signals to thymocytes during the early stages of cell maturation. Patients carrying such mutations present in early infancy with typical features of SCID. Even mutations that preserve the production and expression of the faulty IL-7Rα chain result in a SCID phenotype. Immune evaluation reveals the absence of circulating T cells but preserved B and NK cells ($T^-B^+NK^+$). Cure can be attained with HSCT.

CD45 Deficiency

CD45 deficiency(OMIM*151460) is an extremely rare autosomal recessive form of SCID. CD45 is a receptor tyrosine phosphatase that regulates Src tyrosine kinases involved in T-cell receptor (TCR) and B-cell receptor (BCR) signaling. In the absence of functional CD45, thymocytes are prevented from transitioning from $CD4^+CD8^+$ double-positive cells to single-positive $CD4^+$ or $CD8^+$ expressing mature T cells. Patients present with a SCID phenotype and assessment of immunity shows low to absent T cells but normal numbers of B and NK cells ($T^-B^+NK^+$). Interestingly, while $TCR\alpha\beta^+$ T cells are absent, $TCR\gamma\delta^+$ cells are preserved.

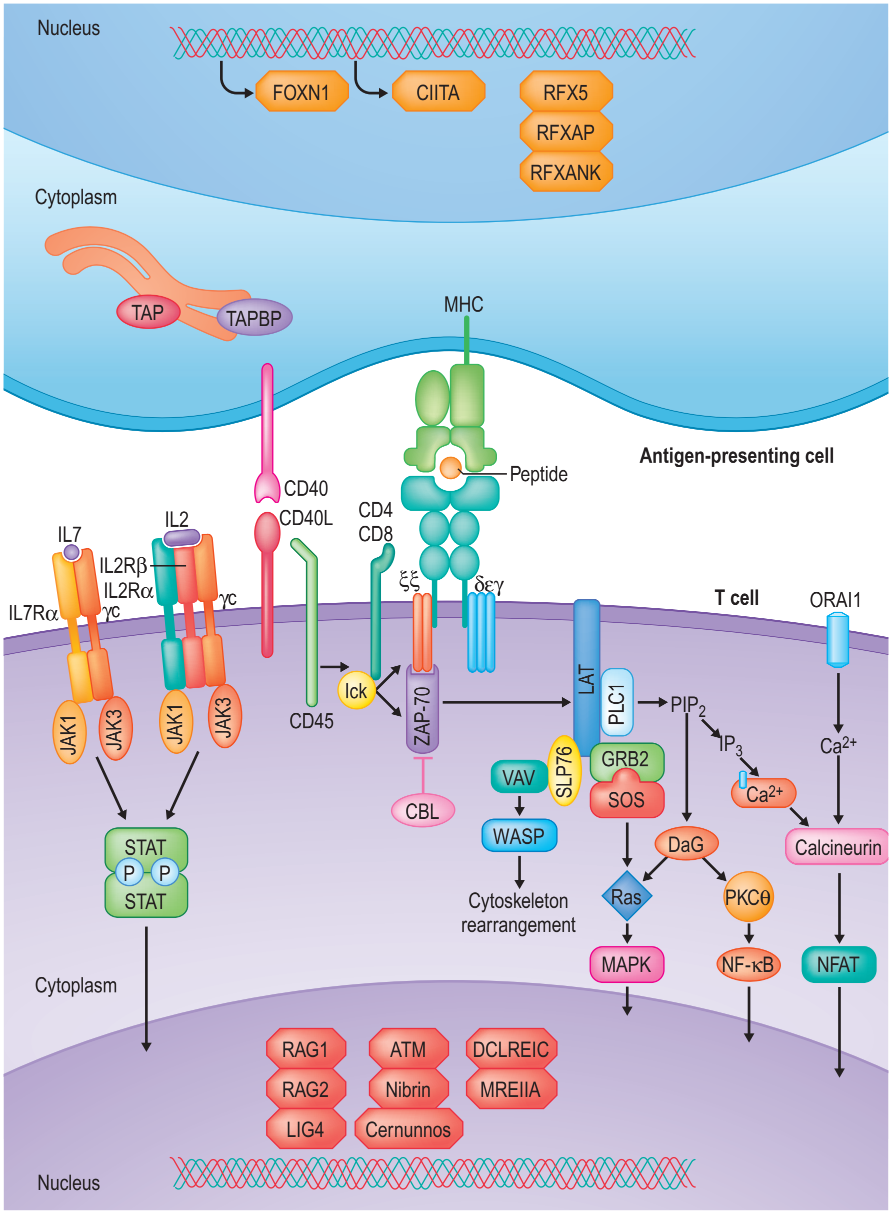

Nucleus
FOXN1
CIITA
RFX5
RFXAP
RFXANK
Cytoplasm
TAP
TAPBP
MHC
Peptide
Antigen-presenting cell
CD40
CD40L
CD4
CD8
IL7
IL2
IL2Rβ
IL2Rα
γc
IL7Rα
γc
ξξ
δεγ
T cell
ORAI1
JAK1
JAK3
JAK1
JAK3
CD45
lck
ZAP-70
LAT
PLC1
PIP2
IP3
Ca2+
Ca2+
SLP76
GRB2
SOS
VAV
WASP
CBL
STAT
P
P
STAT
DaG
Ras
PKCθ
Calcineurin
Cytoskeleton rearrangement
MAPK
NF-κB
NFAT
Cytoplasm
RAG1
RAG2
LIG4
ATM
Nibrin
Cernunnos
DCLREIC
MREIIA
Nucleus

FIG 35.2 T-Cell Signaling Cascade. Activation of T cells, such as following engagement of peptide-major histocompatibility complex (MHC) with the T-cell receptor (TCR), involves signaling through interleukin (IL) receptors, coreceptors, and calcium channels. Recruitment of additional intracellular molecules trigger changes in gene expression, cytokine secretion, and receptor expression. TAP, transporter associated with antigen presentation; TAPBP, TAP binding protein; MHC, major histocompatibility complex; CIITA, class II MHC transactivator; IL7, interleukin (IL)-7; IL7Rα, alpha chain of the IL7 receptor; CD40L, CD40 ligand; ZAP70, zeta chain–associated protein kinase 70; JAK, Janus activated kinase; γc- gamma common chain of the IL-2 receptor; LCK, lymphocyte-specific protein tyrosine kinase; STAT, signal transducer and activator of transcription; CBL, Casitas B-lineage lymphoma; WASP, Wiskott-Aldrich syndrome protein; SLP76, SH2 domain containing leukocyte protein of 76 kDa; LAT, linker for activation of T cells; PLC1, Phospholipase C 1; GRB2, growth factor receptor–bound protein 2; DAG, diacylglycerol; RAS, Rat sarcoma; MAPK, mitogen-activated protein kinase; PKC, protein kinase C; NF-κB, nuclear factor kappa-light-chain-enhancer of activated B cells; NFAT, nuclear factor of activated T-cells; RAG, recombination activating genes; LIG4, DNA ligase 4; ATM, ataxia telangiectasia mutated; DCLRE1C, DNA cross-link repair 1C; MRE11A, MRE11 meiotic recombination 11 homologue A; RMRP, RNA component of mitochondrial RNA-processing endoribonuclease.

TABLE 35.2 **Severe Combined Immunodeficiency (SCID) With T-Cell Lymphopenia (T^-B^+)**

Disease (Immune Phenotype)	Inheritance	Molecular Defect/Presumed Pathogenesis	Additional Features	Treatment
γc deficiency ($T^-B^+NK^-$)	X-linked	Abnormal signaling through interleukin (IL)-2R and other receptors containing γC (IL-4, IL-7, IL-9, IL-15, IL-21)	None	Hematopoietic stem cell transplantation (HSCT)
JAK3 deficiency ($T^-B^+NK^-$)	AR	Abnormal signaling downstream of γc	None	HSCT
IL-7Rα deficiency ($T^-B^+NK^+$)	AR	Abnormal IL7R signaling	None	HSCT
CD45 deficiency ($T^-B^+NK^+$)	AR	Impaired TCR and BCR signaling	None	HSCT
CD3δ deficiency ($T^-B^+NK^+$)	AR	Arrest of thymocytes differentiation at the $CD4^-CD8^-$ stage	Thymus size may be normal	HSCT
CD3ε deficiency ($T^-B^+NK^+$)	AR	Arrest of thymocytes differentiation at the $CD4^-CD8^-$ stage; absent γ/δ T cells	None	HSCT
CD3ζ deficiency ($T^-B^+NK^+$)	AR	Abnormal signaling	None	HSCT
Coronin-1A deficiency ($T^-B^+NK^+$)	AR	Abnormal T-cell egress from thymus and lymph nodes; defect in actin assembly	Normal thymus size; attention deficit disorder; growth retardation	HSCT
FOXN1 deficiency ($T^-B^+NK^+$)	AR	Impaired maturation of thymus and epithelial cells	Alopecia; nail dystrophy	HSCT, thymus transplantation

CD3–TCR Complex Defects

Biallelic mutations in the CD3δ (OMIM *186790), ε (OMIM *186830), and ζ (OMIM *186780) coreceptor chains can result in a typical SCID phenotype ($T^-B^+NK^+$). Together with CD3γ (OMIM *186740), these transmembrane proteins function as signaling molecules for the antigen recognizing TCR heterodimer, which is composed of either the αβ or γδ chains. CD3δ deficiency leads to a complete absence of both the αβ and γδ TCR on circulating CD3 cells because of blockade in thymocyte development at the CD4, CD8 double-negative to double-positive stage. The thymus appears normal in size but lacks Hassall corpuscles. Complete deficiencies (null mutations) in CD3ε and CD3ζ also result in a lack of circulating T cells. Atypical presentations are associated with leaky defects of CD3ε or CD3δ. Although clinical presentation could be indistinguishable from typical SCID, these patients may have a sizeable amount of autologous T cells ($T^+B^+NK^+$). Cure of CD3 deficiencies can be achieved with HSCT.[13]

SCID With T-Cell Lymphopenia and Syndromic Features

Coronin-1A Deficiency

Coronin-1A (OMIM *605000) is a leukocyte-specific signaling molecule, which is essential for TCR ligation–dependent generation of inositol-1,4,5-triphosphate (IP3). IP3 mediates Ca^{2+} release from intracellular stores and is critical for the proper function of downstream signaling enzymes, culminating in T-cell activation and survival. In addition, coronin-1A is essential for lymphocyte migration and lymphocyte homeostasis.

Patients with coronin-1A deficiency (OMIM #615401), an autosomal recessive condition, can present during infancy with a SCID phenotype ($T^-B^+NK^+$), including oral thrush, recurrent respiratory infections, BCG lymphadenitis, or Epstein-Barr virus (EBV)–associated lymphoproliferation. Others may present at a much older age (11 years) with EBV-associated lymphoproliferation and a $T^+B^+NK^+$ CID presentation. Evaluation of the immune system reveals a very low number of circulating T cells, absent naïve T cells, abnormal T cell repertoire and markedly reduced TREC levels. In-vitro responses to mitogens are variably reduced. Serum immunoglobulin levels are within normal range but production of specific antibodies in response to vaccination is impaired. Imaging can detect a thymus shadow. The associated syndromic features consist of developmental delay, growth retardation, low cognitive functions, and attention deficit disorder. Immune function can be restored with HSCT.

FOXN1 Deficiency (Combined Immunodeficiency With Alopecia Totalis)

Forkhead box protein N1 (FOXN1) deficiency (OMIM #601705) is an autosomal recessive syndrome encompassing alopecia, which affects the scalp, eyebrows, and eyelashes; nail dystrophy, which is noted at birth; and profound cellular immunodeficiency. The syndrome is caused by mutations in the *FOXN1* gene, a member of the forkhead/winged helix transcription factor family, which is mainly expressed in the thymus epithelium and in skin. *FOXN1* plays a critical role in the maturation of the thymus and skin epithelial cells. A deleterious mutation in this gene was identified in nude athymic mice. Patients can present at 2–4 months with Omenn syndrome, recurrent infections, and failure to thrive. Circulating lymphocytes may be low or normal with CD4 lymphopenia, and yet the TCR repertoire is restricted. Circulating T cells are phenotypically immature and double $CD4^-CD8^-$. While TRECs are low, *in vitro* responses to mitogens are variable. Both HSCT and thymus transplantation have been tried in these patients. HSCT led to full engraftment and infection-free clinical recovery. However, immune reconstitution and maturation of donor stem cells are still in question. Thymus transplantation appears successful, with recovery to lymphocyte repertoire and function.

SCID With T-Cell and B-Cell Lymphopenia (T^-B^-)

Up to 30% of all SCID cases appear to lack both circulating T cells as well as B cells (Table 35.3).

RAG1/2 Deficiency

This type of SCID (OMIM #601457) has an autosomal recessive pattern of inheritance and is caused by mutations in the recombination activating genes *RAG1* (OMIM *179615) and *RAG2* (OMIM *179616). Recognition of different antigens is mediated

TABLE 35.3 Severe Combined Immunodeficiency (SCID) With T- and B-Cell Lymphopenia (T^-B^-)

Disease (Immune Phenotype)	Inheritance	Molecular Defect/Presumed Pathogenesis	Additional Features	Treatment
RAG1/2 deficiency ($T^-B^-NK^+$)	AR	Defective V(D)J recombination	None	Hematopoietic stem cell transplantation (HSCT)
Artemis deficiency ($T^-B^-NK^+$)	AR	Defective V(D)J recombination, radiation sensitivity; DCLRE1C gene defects	None	HSCT
DNA-PK deficiency ($T^-B^-NK^+$)	AR	Defective V(D)J recombination	None	HSCT
DNA Ligase IV deficiency ($T^-B^-NK^+$)	AR	Defective V(D)J recombination, radiation sensitivity	Growth delay; microcephaly; bone marrow abnormalities; lymphoid malignancies	HSCT
Cernunnos deficiency ($T^-B^-NK^+$)	AR	Defective V(D)J recombination, radiation sensitivity	Growth delay; microcephaly; bird-like facies; bone defects	HSCT
AK2 deficiency/reticular dysgenesis ($T^-B^-NK^-$)	AR	Impaired mitochondrial energy metabolism and leukocyte differentiation	Deafness	gCSG HSCT
Adenosine deaminase deficiency ($T^-B^-NK^-$)	AR	Accumulation of toxic purine nucleosides	Neurological, hepatic, renal, lung, skeletal and bone marrow abnormalities	HSCT, PEG-ADA, gene therapy

by antigen-specific TCR or BCR. This requires the production of variable regions of the antigen receptors, a process known as *V(D)J recombination* (Chapter 4). Proper recognition of BCR and TCR are important for maturation of these lineages as well as clonal expansion. RAG1 and RAG2 are the chief enzymes responsible for recombination of the antigen receptors. Patients who carry null mutations present with typical SCID features at infancy; however, hypomorphic mutations can be associated with Omenn syndrome. Other leaky phenotypes have been reported with moderate lymphopenia and autoimmunity; EBV-associated lymphoproliferation or noninfectious granulomas of the skin, bones, lymphoid tissues, and the respiratory system; and CD4 lymphocytopenia.

Diagnosis is supported by flow cytometry showing marked reduction in circulating T and B cells in null mutations or by profoundly skewed T-cell repertoire in hypomorphic cases. Autologous cells fail to respond to mitogens or antigens and TREC levels are low. Immunoglobulins (Igs) and specific antibodies are lacking.[4] Cure can be achieved by administering HSCT, even in cases of Omenn syndrome.[11] Gene correction is currently being investigated as an alternative option to HSCT.[14]

Artemis Deficiency

Artemis deficiency is an autosomal recessive disorder caused by mutations in the DNA cross-link repair 1C (*DCLRE1C*) gene (OMIM *605988), encoding the protein Artemis. Similar to RAG1/2, *DCLRE1C* is obligatory for DNA recombination of antigen receptors, by opening RAG-generated coding hairpins in a DNA protein kinase-dependent manner. Artemis is critical for the repair of genomic lesions and is essential for genomic stability. Clinically, patients may present with typical SCID features. Hypomorphic mutations in the Artemis gene may present with Omenn syndrome, and EBV-associated lymphoma has also been reported. HSCT remains the sole modality for curing this condition. Because of the genomic instability caused by Artemis deficiency, these patients may be sensitive to ionizing radiation and possibly chemotherapy, and therefore reduced intensity conditioning is advised.

KEY CONCEPTS

Diagnosis of T-Cell Immunodeficiencies

- T-cell receptor excision circle (TREC)–based newborn screening detects all types of profoundly lymphopenic T-cell severe combined immunodeficiency (SCID) and may detect some, but not all, cases of combined immunodeficiency (CID).
- Immunophenotyping remains an indispensable tool in the diagnosis of all types of T-cell immunodeficiencies. When needed, expanding the panel of markers to detect regulatory T cells (Tregs) as well as naïve and memory T cells and B cells can be valuable.
- Genetic analysis has become the gold standard for diagnosis of T-cell immunodeficiency. Modern next-generation sequencing allows for analysis of SCID or CID panels. Whole exome sequencing enables identification of both known as well as novel defects in the immune system.
- Basic functional assays, such as responses to mitogens, are still widely used to aid in the diagnosis of SCID and CID. Analysis of cytokine secretion, cytokine autoantibodies, and detailed study of T-cell receptor repertoire all contribute to better defining immunodeficiency. Novel mutations should be assessed by studying their impact on protein translation and function.
- Thymus biopsies can be helpful when considering high-risk procedures, such as hematopoietic stem cell transplantation (HSCT). Although normal thymus morphology cannot exclude primary T-cell deficiency, abnormal thymus architecture and abnormal Hassall corpuscles are specific and sensitive predictors of primary T-cell deficiency. This invasive procedure should be restricted to cases where all other diagnostic tools remain inconclusive, and a life-saving procedure is considered. It should only be performed in centers with abundant experience with this procedure.

SCID With T-Cell and B-Cell Lymphopenia (T^-B^-) and Syndromic Features

DNA-PKcs Deficiency

DNA-dependent protein kinase, catalytic subunit (DNA-PKcs) deficiency is caused by mutations in the *PRKDC* gene (OMIM *600899) and is transmitted in an autosomal recessive manner.

The serine/threonine kinase recognizes open DNA ends and is part of a complex involved in the nonhomologous end joining (NHEJ) pathway essential for DNA recombination. Clinical presentation is indistinguishable from SCID, with virtually absent T cells and B cells but normal NK cells. These patients may present with severe neurological abnormalities.[15] HSCT may be a curative option, although caution must be exercised during the conditioning regimen because of the underlying defect in DNA repair.

DNA Ligase IV Deficiency

DNA ligase IV deficiency (OMIM #606593) is clinically pleomorphic with presentations ranging from intact immunity to a SCID phenotype. Other features include facial dysmorphia, various degrees of developmental delay, and microcephaly. These features are consistent with DNA ligase IV knock-out mice, which are characterized by defective lymphogenesis and neurogenesis. DNA ligase IV is a component of the NHEJ complex and participates in the repair of DNA double-strand breaks that arise during DNA damage, such as from ionizing radiation, or the endogenous recombination process. Hypomorphic mutations in DNA ligase IV can be associated with increased sensitivity to radiotherapy,[16] EBV-associated lymphoproliferation and T-cell leukemia. Diagnosis may be challenging, with immune evaluation revealing low to absent T cells and B cells during infancy in some patients, whereas in others diagnosis may be delayed to the second decade of life in individuals with various degrees of lymphopenia and hypogammaglobulinemia. Microcephaly and developmental delay can help in the diagnosis if present. HSCT was attempted in patients with DNA ligase IV deficiency with limited success because of increased toxicity of conditioning regimens as well as graft-versus-host disease (GvHD).

Cernunnos Deficiency

Patients with Cernunnos deficiency (OMIM #611291) present early in life with profound T-cell lymphopenia, progressive decrease in B cells (T^-B^- SCID), and microcephaly. The syndrome has an autosomal recessive pattern of inheritance. Cernunnos (OMIM *611290) forms a complex with DNA ligase II and XRCC4 and likely plays an important role in the NHEJ pathway. Disruption of Cernunnos, like other members of this pathway, is associated with radiosensitivity. Clinical presentation consists of recurrent infections, as in typical SCID, as well as developmental delay, microcephaly, urogenital and bone malformations, and facial dysmorphic features.

AK2 Deficiency (Reticular Dysgenesis)

Adenylate kinase-2 (AK2) deficiency, also known as *reticular dysgenesis* (OMIM #267500), is a profound inherited immune deficiency characterized by severe lymphopenia and marrow failure. The disorder has an autosomal recessive pattern of inheritance and is caused by mutations in the *AK2* gene.[17,18] Mitochondrial AK2 catalyzes the reversible phosphorylation between nucleoside triphosphates and monophosphates. AK2 defects are associated with impaired mitochondrial energy metabolism and leukocyte differentiation. Patients suffer repeated bacterial, viral, and fungal infections and have bilateral sensorineural deafness.[17] Patients present in infancy with profound neutropenia and lymphopenia, the thymus is dystrophic, and TRECs are undetectable. Diagnosis is confirmed by genetic analysis. Bone marrow reconstitution can be achieved with HSCT and remains the most effective modality of treatment.

Adenosine Deaminase Deficiency

Adenosine deaminase (ADA) deficiency (OMIM #102700) is an autosomal metabolic disorder caused by deleterious mutations in the *ADA* gene (OMIM*608958). This is a critical enzyme that aids in the degradation and salvage of purine pathway metabolites and is essential for multiple processes, such as energy transfer and DNA metabolism. Damage to various tissues and lesions in the immune system are caused by toxic accumulation of undegraded metabolites, such as deoxyadenosine. Thymocytes are particularly sensitive, but B cells and NK cells are also frequently affected. Because toxicity to lymphocytes is lineage specific and time dependent, patients may present with T^-B^- or T^-B^+ phenotype. Some may have Omenn syndrome and present with T^+B^- or T^+B^+.[19] Mutations that spare some ADA activity may result in a delayed presentation.[20] Patients with ADA deficiency mostly present with typical SCID features or Omenn syndrome, while patients with "partial deficiencies" can present with CID, autoimmunity, and susceptibility to malignancy.[21] Other associated features include brain anomalies resulting in various levels of developmental delay and bone, liver, lung, and kidney anomalies. Standard diagnostic tools for this disorder demonstrate barely detectable ADA enzyme activity and increased accumulation of purine metabolites in blood or urine. The diagnosis can be confirmed by genetic analysis. Polyethylene glycol-conjugated ADA (PEG-ADA) can be used for "detoxification"; however, salvage of complete or even partial immunity may not be achieved, and efficacy may be short lived.[22] HSCT remains the treatment of choice if a full human leukocyte antigen (HLA)–matched related donor is available. For other patients, gene therapy is gradually becoming standard of care.[23]

COMBINED IMMUNODEFICIENCY (SCID PHENOTYPE)

Combined immunodeficiencies presenting with features typical of SCID are summarized in Table 35.4.

ZAP-70 Deficiency

ZAP-70 deficiency (OMIM #269840) is a rare CID with an autosomal recessive pattern of inheritance. The disease was first described as a novel immunodeficiency with normal numbers of circulating CD4 cells but CD8 lymphopenia. ZAP-70 is a protein tyrosine kinase critical in mediating T-cell receptor signaling. Upon phosphorylation and activation, ZAP-70 phosphorylates a host of downstream molecules culminating in T-cell proliferation and maturation.[24] Patients typically present within the first 2 years of life with repeated severe microbial infections indistinguishable from SCID. However, unlike typical SCID, patients with ZAP-70 deficiency also present with palpable lymph nodes, visible tonsils, and a normal thymus shadow on imaging. Rarely, patients present with autoimmune manifestations, Omenn syndrome, hemophagocytic lymphohistiocytosis (HLH), subcutaneous nodules, or lymphoma.[25,26]

Immune evaluation reveals normal numbers of circulating lymphocytes but marked reduction in CD8 cells.[4,7] Responses to mitogens are depressed; however, the T-cell repertoire of CD4 cells appears normal. The thymus gland is fully developed with normal architecture and normal corticomedullary distinction, including the presence of Hassall corpuscles. However, the corticomedullary ratio appears increased. B-cell number and Ig

TABLE 35.4 Combined Immunodeficiency (Severe Combined Immunodeficiency [SCID] Phenotype)

Disease	Inheritance	Molecular Defect/Presumed Pathogenesis	Features	Treatment
ZAP-70 deficiency	AR	Abnormal T-cell receptor signaling	Normal thymus size; increased corticomedullary ratio	Hematopoietic stem cell transplantation (HSCT)
Major histocompatibility complex (MHC) class II deficiency	AR	Absence of human leukocyte antigen (HLA) class II molecules	Recurrent infections; gastrointestinal manifestations; hepatic abnormalities	Supportive
MHC class I deficiency	AR	Impaired expression of HLA class I molecules	Respiratory infections; skin granulomatous lesions	Supportive
DOCK2 deficiency	AR	Impaired T-cell homing	Viral and bacterial infections	HSCT
CD3γ deficiency	AR	Arrest of thymocyte differentiation at the CD4$^-$CD8$^-$ stage	Thymus size may be normal	HSCT
CARD 11 / BCL10 / MALT1	AR	Abnormal signaling upstream of NF-κB pathway	Opportunistic infections	HSCT
TTC7A deficiency	AR	*TTC7A* gene defects; impaired regulation of transcription, cell cycle, protein degradation and trafficking	Intestinal atresia	HSCT
ORAI-1, STIM-1		Impaired Ca^{2+} fluxes	Autoimmunity; ectodermal dysplasia	HSCT

levels are normal. NK cells may be reduced in some patients (personal communication). Allogeneic HSCT can completely correct the immunodeficiency.

MHC Class II Deficiency (Bare Lymphocyte Syndrome)

MHC class II deficiency is a rare immunodeficiency characterized by the lack of expression of HLA class II molecules (OMIM #209920). The deficiency is caused by mutations in genes encoding transactivating elements critical for regulating HLA class II expression. Such deleterious disease causing mutations have been identified in the class II transactivator (CTIIA), the regulatory factor–associated protein (RFXANK or RFX-B), the fifth member of the regulatory factor X family (RFX5) and the regulatory factor–associated protein (RFXAP).

HLA class II is expressed on thymic epithelial cells, antigen-presenting cells (APCs), and activated lymphocytes. The structure and function of these molecules, which are required for T-cell maturation and antigen presentation, are discussed in detail in Chapters 5 and 6.

Clinical manifestations are highly variable with some patients presenting with features of typical SCID, including oral thrush, *Pneumocystis jiroveci* pneumonia, and failure to thrive. Others may have a milder course. Chronic enterocolitis and sclerosing cholangitis are common and are caused by infections. Meningoencephalitis caused by enteroviruses, herpes simplex, or adenoviruses has also been reported. Autoimmune cytopenias are common.

Historically, the diagnosis relied on demonstrating the lack of MHC class II expression on APCs by flow cytometry. Lymphocyte counts are usually normal, but CD4 lymphocytopenia is frequently observed. *In vitro* responses to mitogens may be normal or depressed, but TREC levels are within normal limits. The definitive diagnosis is made by mutational analysis. Overall, the prognosis for patients with this condition is poor. Patients usually die within the first or second decade of life, as a result of infections or infection-related organ damage. HSCT is offered, but outcomes are usually poor, with a high rate of engraftment failures, severe GvHD, and poor immune reconstitution. Supportive treatment consists of antibiotics and Ig replacement.

MHC Class I Deficiency

This is a rare autosomal recessive immune deficiency characterized by partial expression of HLA class I molecules (OMIM #604571). The decrease in HLA class I expression is caused by mutations in *TAP1* (OMIM *170260) or *TAP2* (OMIM *170261). These endoplasmic adenosine triphosphate (ATP)–binding transporters are essential for processing the association of endogenous peptides with MHC class I molecules. MHC class I molecules that do not bind high affinity peptides do not cross the Golgi apparatus. Recently mutations in the genes encoding tapasin and β_2 microglobulin, involved in antigen and HLA class I molecule processing, have been shown to result in a similar deficiency. Patients present during childhood or later in life with lung infections and/or skin granulomatous lesions. The diagnosis is suspected when HLA class I expression is decreased upon immunophenotyping. Confirmation of the diagnosis is obtained by genetic analysis. Treatment remains supportive in nature.

DOCK2 Deficiency

Dedicator of cytokinesis 2 (DOCK2) deficiency (OMIM #616433) is an autosomal recessive combined immunodeficiency. DOCK2 (OMIM *603122) mediates cytoskeletal reorganization through Rac activation and is required for lymphocyte homing to lymphoid tissues. Indeed, DOCK2-deficient mice exhibit lymphocyte migration defects.

Patients with DOCK2 deficiency experience severe invasive bacterial and viral infections at early childhood. Patients may have T-cell lymphopenia and reduced *in vitro* T-cell proliferation, reduced circulating B cells, defective NK-cell activity, and antibody deficiency.[27] Although experience is very limited, HSCT may cure this condition.

CD3γ Deficiency

CD3γ deficiency (OMIM #615607) is caused by mutations in the CD3γ chain and is transmitted in an autosomal recessive manner. Patients can present with a SCID phenotype or with delayed atypical features, consisting of infections or autoimmunity. This unusual heterogeneous clinical pattern has been observed in patients of Turkish or Spanish descent. Autologous T cells have partial TCR/CD3 expression with only mild lymphocytopenia ($T^+B^+NK^+$). *In vitro* responses of mitogens are depressed, and TRECs can be reduced. Although some of these patients present as typical SCIDs or with severe colitis, other family members with an identical mutation appear well, even into adulthood.

CARD11/BCL10/MALT1 (CBM) Complex Deficiencies

CARD11/BCL10/MALT1 deficiencies are autosomal recessive immunodeficiencies caused by mutations in signaling molecules upstream of the NF-κB pathway. Upon TCR or BCR stimulation, caspase recruitment domain family, member 11 (CARD11) is phosphorylated, enabling its association with BCL10–MALT1 complex, to form the CBM complex. The CBM complex allows for activation of the NEMO–IκB kinase (IKK) complex, which induces phosphorylation of IκB and causes its breakdown. As a consequence, NF-κB is released and translocates into the nucleus, where it acts as a transcription factor, controlling multiple cellular processes, such as cell proliferation, differentiation, and survival.

Patients with CARD11 deficiency (OMIM #615206) may present with a SCID phenotype.[28] They also suffer gingivitis and aphthous ulcers. Patients with MALT1 deficiency (OMIM #615468) present at a later age during childhood with severe enterocolitis and recurrent lung infections.[29] Evaluation of the immune system in CARD11 deficiency shows normal number of circulating T cells, B cells, and NK cells, but regulatory T cells (Tregs) are absent (T^+B^+ SCID). Responses to mitogens are reduced but not absent, and TRECs and the TCR repertoire are normal. B cells appear immature, which explains the panhypogammaglobulinemia found in these patients. Patients with MALT1 deficiency present similarly; however, although Ig levels are normal, antibody production is faulty. Cure can be achieved with HSCT.

Autosomal dominant CARD11 deficiency is caused by a dominant negative monoallelic mutation and results in combined immunodeficiency, severe atropy and autoimmunity.

COMBINED IMMUNODEFICIENCY (SCID PHENOTYPE) AND SYNDROMIC FEATURES

TTC7A Deficiency (MIA Syndrome)

TTC7A deficiency is a rare autosomal recessive disorder in which obstructions occur at various levels of the gut, hence the designation *multiple intestinal atresia* (MIA) syndrome (OMIM #243150).[30] Patients may also have severe lung disease and pulmonary calcifications. Immune deficiency is commonly associated with this syndrome and is frequently severe. Patients may present with typical SCID features,[31] albeit with various levels of autologous T cells. The disease is caused by mutations in the *TTC7A* gene (OMIM *609332), which encodes a protein containing nine tetratricopeptide repeat (TPR) domains. These domains are degenerate 34-amino acid repeat motifs and appear to play a role in multiple cellular processes, such as transcription, cell cycle, protein degradation, and trafficking.

Evaluation of the immune system in TTC7A deficiency reveals T-cell lymphopenia and markedly depressed responses to mitogens. The TCR repertoire is restricted, and TRECs are low. The thymus appears dysplastic with poor corticomedullary demarcation, as well as poorly developed Hassall corpuscles. Nonimmune manifestations are striking, with intraluminal calcifications. In some patients, atresias cannot be detected; rather, severe colitis with characteristic features of mucosal atrophy and massive enterocyte apoptosis is present.[32] The prognosis of this condition is poor. Although HSCT can replenish the hematopoietic system,[30] this does not solve the complex and fatal gastrointestinal (GI) manifestations. Combined gut and bone marrow transplantation could be attempted.

Calcium Channel Defects (ORAI-1, STIM-1 Deficiencies)

This is an autosomal recessive disorder caused by mutations in ORAI-1 (OMIM *610277) and STIM-1 (OMIM *605921).[33] ORAI-1 is a transmembrane protein, which at tetramer configuration constitutes the pore-forming structure of a Ca^{2+} channel. STIM-1 is a calcium sensor of the endoplasmic reticulum (ER) and controls calcium entry after ER store depletion. Both Ca^{2+} influx into the cell as well as release from cytosolic stores are mediated through the TCR and BCR, and are critical for downstream signaling. Patients may present in infancy with recurrent severe infections indistinguishable from SCID.[34] Patients with STIM1 deficiency can also present with autoimmune features, such as arthritis and cytopenia, as well as lymphoproliferative manifestations. Both deficiencies may present with severe eczema and syndromic features, including nail dysplasia and anhydrosis, ectodermal dysplasia, and generalized myopathy. Immune evaluation shows normal to elevated lymphocyte counts and TREC levels, and responses to mitogens are depressed. Ig levels are normal, but specific antibodies are variable. TCR-mediated cytosolic rise in Ca^{2+} is low. HSCT can be effective in reversing the immunodeficiency, but not the syndromic features.

COMBINED IMMUNODEFICIENCIES WITH VARIABLE SEVERITY (NON-SCID)

Combined Immunodeficiency With Immune Dysregulation

Combined immunodeficiencies characterized by autoimmune manifestations are summarized in Table 35.5.

IL-2Rα (CD25) Deficiency

IL-2 receptor α (IL-2Rα or CD25) deficiency is a complex immune dysregulation disorder that has an autosomal recessive pattern of inheritance (OMIM #606367). Biallelic mutations in the CD25 gene result in combined immunodeficiency ($T^+B^+NK^+$). The high affinity IL-2 receptor must be composed of all three components, the γ chain (γc), β chain (CD122), and α chain (CD25), to be functional. Although the β and γ chains are constitutively expressed on T cells, expression of the α chain is limited to early thymocytes, Tregs and activated T cells. The inability to form the high affinity receptor results in faulty T-cell differentiation, loss of central and peripheral tolerance, and failure to respond appropriately to infections. Affected infants present with recurrent infections, lymphadenopathy and hepatosplenomegaly, a variety of autoimmune features, and eczematous skin lesions. Patients may present immediately after birth with diabetes mellitus and may suffer severe chronic lung disease or pulmonary bleeding. Autoimmune lymphocytic infiltrates can be identified in multiple tissues, including lung, liver, gut, and bone. Autoimmune disorders include primary cirrhosis, colitis, thyroiditis, or insulin-dependent diabetes mellitus.

Evaluation of the immune system reveals reduced numbers of circulating T cells, which are unable to respond to mitogens or antigens, but normal B cells. The thymus appears dysplastic with absent Hassall corpuscles and loss of corticomedullary demarcations. Thymocytes do not express CD1 and fail to downregulate

TABLE 35.5 Combined Immunodeficiency With Immune Dysregulation

Disease	Inheritance	Molecular Defect/Presumed Pathogenesis	Features	Treatment
Interleukin (IL)-2Rα / CD25 deficiency	AR	Impaired T-cell differentiation	Recurrent infections; autoimmunity	Hematopoietic stem cell transplantation (HSCT)
FOXP3 deficiency	X-linked	Impaired regulatory T cell (Treg) development	Autoimmunity; impaired barrier function in skin and gut; enteropathy	HSCT
IL-10, IL-10R deficiency	AR	Impaired antiinflammatory cytokine production	Early onset inflammatory bowel disease (IBD); autoimmunity	HSCT
IKK2 deficiency	AR	Disruption of NF-κB pathway	Recurrent fungal and viral infections; failure to thrive	HSCT
DOCK8 deficiency	AR	Disruption in cell adhesion and migration	Recurrent infections; atopy; early onset malignancy; autoimmunity	HSCT
RHOH deficiency	AR	Defective Rho guanosine triphosphatase (GTPase) expression resulting in impaired T-cell receptor (TCR) signaling	Human papilloma virus (HPV) infections	HSCT
MAGT1 / CTPS1 / SH2DIA / BIRC 4	X-linked, AR	Variable depending on associated defect	Epstein-Barr virus (EBV) lymphoproliferation	HSCT
PNP deficiency	AR	Abnormal purine metabolism	Recurrent infections; autoimmunity, neurological abnormalities	HSCT
DiGeorge syndrome	AD	Deletion in chromosome 22q11.2 resulting in impaired TBX1 transcription factor, affecting developmental processes	Omenn syndrome; Truncus arteriosus; autoimmunity; cardiac defects; hypocalcemia; dysmorphism; cleft palate; short stature; neuropsychiatric problems	Supportive
Wiskott-Aldrich syndrome	X-linked recessive	Impaired actin cytoskeleton rearrangement and immunological synapse formation	Eczema; autoimmunity; malignancy	HSCT
Ataxia telangiectasia	AR	Impaired DNA repair and cell cycle regulation	Cerebellar ataxia; telangiectasias; predisposition to malignancy; radiosensitivity	Supportive
Nijmegen breakage syndrome	AR	Impaired DNA repair and cell cycle regulation	Microcephaly; growth retardation; predisposition to malignancy; radiosensitivity	Supportive

levels of the antiapoptotic protein Bcl-2. This results in marked reduction of apoptosis and the expansion of autoreactive clones in multiple tissues. The markedly reduced Tregs in the absence of CD25 explains the uncontrolled expansion of these clones, resulting in autoimmunity. HSCT cures this disorder.

FOXP3 Deficiency

Mutations in the Forkhead box protein P3 (*FOXP3*) gene cause an X-linked recessive disorder, also known as IPEX (immune dysregulation/polyendocrinopathy/enteropathy/X-linked) syndrome (OMIM #304790). The protein encoded by FOXP3, termed *scurfin,* is a transcription factor that plays a critical role in the development and function of $CD4^{+}CD25^{+}$ Tregs (Chapter 18). Affected patients suffer from severe autoimmunity in infancy. Severe enterocolitis is often the presenting feature. Insulin-dependent diabetes mellitus, hepatitis, autoimmune cytopenias, and arthritis are frequent. Patients may also have increased frequency of invasive bacterial infections. Circulating lymphocytes are present in normal numbers, and they respond normally to mitogen stimulation. TREC levels are also normal. However, the number and function of $CD4^{+}CD25^{+}FOXP3^{+}$ Tregs is reduced. Patients have normal Ig and antibody levels, as well as increased autoantibodies against various tissues. Immunosuppressive drugs can transiently control the autoimmune phenomena, but HSCT offers the best option for cure, particularly if performed early before irreversible organ damage occurs.[35]

IL-10 and IL-10Rα and IL-10Rβ Deficiencies

IL-10, produced mainly by T cells and macrophages, is a key antiinflammatory cytokine. Biallelic deleterious mutations in IL-10 (OMIM *124092) or its receptor IL-10R, composed of α (OMIM *146933) and β (OMIM *123889) chains, causes early-onset severe inflammatory bowel disease (OMIM #613148, #612567).[36] Patients suffer perianal abscesses and fistulas that often require surgical intervention. Some patients with IL-10/IL-10R deficiency also develop other autoimmune manifestations, such as rheumatoid arthritis. Treatment with immunosuppressive agents may provide partial and temporary remission. Only HSCT proved effective in completely reversing colitis. However, other autoimmune features, such as arthritis, may persist.

IKK2 Deficiency

IKK2 deficiency (OMIM #615592) is a combined immunodeficiency that is inherited in an autosomal recessive manner and is caused by biallelic mutations in the *IKBKB* gene (OMIM *603258), leading to disruption of the nuclear factor-κB (NF-κB) pathway. The NF-κB pathway regulates immune responses, cell activation and growth, as well as cell survival.[34] The canonical pathway encompasses the NF-κB essential modulator (IKKγ, or NEMO), IKK1 (IKKα), and IKK2 (IKKβ). Upon activation of the complex, IKK2 phosphorylates inhibitors of NF-κB, which are degraded, therefore allowing nuclear translocation of NF-κB and DNA binding. Patients present with oral thrush, pulmonary infections, and invasive bacterial infections.[38] Evaluation of the immune system reveals normal numbers of circulating T cells but reduced NK cells. Responses to mitogens are variable, TREC levels are normal, and the TCR repertoire is intact. The B-cell compartment lacks memory B cells ($T^+B^+NK^-$).[38] HSCT is a plausible choice for treatment; however, results so far are very limited and variable. Like other NK-κB pathway disorders, engraftment may be incomplete using current bone marrow transplantation protocols.

DOCK8 Deficiency

Dedicator of cytokinesis 8 (DOCK8) deficiency (OMIM #243700) is an autosomal recessive combined immunodeficiency combined with severe atopy.[39] DOCK8 interacts with Rho guanosine triphosphatases (GTPases) and is believed to be involved in cytoskeletal rearrangement, affecting cell adhesion and motility. Patients suffer recurrent respiratory infections and, less frequently, invasive bacterial and fungal infections. In addition to severe eczema, patients may have extensive warts and molluscum contagiosum. Food allergies, asthma, and anaphylaxis have also been documented. Immunological assessment shows great variability with variable number of T cells and B cells but increased eosinophil counts. Responses to mitogens appear normal. IgG is elevated, but other serum Ig levels and specific antibodies are variable.[39] Treatment is dictated by the various atopic and infectious manifestations.

RHOH Deficiency

This is an autosomal recessive immunodeficiency caused by mutations in the Ras homologue gene family member H gene (RHOH; OMIM *602037), which is mostly expressed in the hematopoietic system. Upon ligation of TCR, RhoH undergoes phosphorylation on tyrosine residues and leads the recruitment of Zap-70 and Lck to downstream signaling pathways. Patients may present with persistent skin infections, granulomatous lung disease, and Burkitt lymphoma. Circulating lymphocytes, Igs, and antibody production are all normal, and *in vitro* responses to mitogens are variable.[40]

COMBINED IMMUNODEFICIENCY WITH EBV-INDUCED LYMPHOPROLIFERATION

EBV-induced chronic infection and lymphoproliferation is the major clinical presentation of a growing number of newly described inherited immunodeficiencies. Inheritance can be X-linked or autosomal recessive according to the gene affected. Hemizygous mutations in the magnesium transporter 1 (*MAGT1*) gene (OMIM *300715) causes reduced TCR-mediated Mg^{2+} as well as Ca^{2+} influx into T cells (OMIM #300853). This ultimately results in abnormal downstream activity of transcription factors. Patient CD4 T cells are reduced, but the total numbers of CD3 cells, B cells, and NK-cells are normal. EBV related infections or lymphoma may develop at any age.[41] A similar clinical picture (OMIM #615897) is found in patients with biallelic CTP synthase 1 (CTPS1; OMIM *123860) mutations.[42] The nucleotide cytidine-5'-triphosphate (CTP) is needed for the metabolism of RNA and DNA. Circulating lymphocyte counts are variable, Ig levels are normal, and responses to mitogens are normal in most patients. Some patients suffer bacterial infections as well as chronic EBV viremia, and HSCT has been sometimes successful.[42]

EBV-induced lymphoid proliferation and lymphoma (OMIM #613011) can also be caused by biallelic mutations in the IL-2 inducible T-cell kinase (*ITK*) gene.[43] X-linked lymphoproliferative syndrome-1 (XLP, OMIM #308240) and -2 (XLP2, OMIM #300635) are caused by hemizygous mutations in *SH2DIA* (XLP1) and *BIRC4* (XIAP), respectively.[44] Patients can present with overwhelming EBV infection causing liver necrosis and failure, aplastic anemia, or with chronic EBV infections culminating in malignant lymphoma. Some patients present with hypogammaglobulinemia and antibody deficiency, and some others may develop hemophagocytic lymphohistiocytosis.[45] XLP2 is more frequently associated with chronic colitis.[46] Lymphocyte numbers and response to mitogens are typically normal, whereas NK-cell activity may be impaired. HSCT can reverse the susceptibility to EBV.[45]

COMBINED IMMUNODEFICIENCY WITH IMMUNE DYSREGULATION AND SYNDROMIC FEATURES

PNP Deficiency

Purine nucleoside phosphorylase (PNP) deficiency (OMIM # 613179) is an autosomal recessive multisystem syndrome caused by mutations in PNP, a key enzyme in the purine salvage pathway. PNP catalyzes the phosphorylation of guanosine, deoxyguanosine, inosine, and deoxyinosine. This pathway is responsible for balancing the production of dephosphorylated purines and degradation to uric acid and salvage back to the nucleotide level. Lymphotoxicity is caused by accumulation of dGTP in mitochondria inducing apoptosis, and T-cell toxicity may be related to the need for high deoxyguanosine phosphorylation activity in T cells. Indeed, PNP-deficient mice created by gene targeting have no detected PNP activity and demonstrate a sharp reduction in T-cell numbers and function due to loss of double positive $CD4^+CD8^+$ progenitors, as well as progressive loss of peripheral T cells.

Patients with PNP deficiency display similar abnormalities in T-cell maturation and function.[4] The thymus appears partially dysplastic. Circulating T-cell numbers are reduced, and T-cell function is depressed. Igs and specific antibody levels may be variable. B- and NK-cell numbers are variable. Patients with PNP deficiency typically present with a triad of features, including recurrent infections, autoimmune manifestations, and various neurological abnormalities. Infections may be consistent with SCID, but severe bacterial infections have been mostly reported in older patients with delayed presentation. Autoimmune features include cytopenias, systemic lupus erythematosus (SLE), and arthritis, and neurological features consist of ataxia and various degrees of behavioral and mental retardation. Diagnosis is confirmed by demonstrating markedly reduced PNP enzyme activity in red blood cells (RBCs) or T cells, reduced uric acid, an increase of deoxyguanosine and other metabolites in blood and urine, and by genetic analysis of the *PNP* gene. HSCT may reverse the immunodeficiency, but not the neurological manifestations.

DiGeorge Syndrome

DiGeorge syndrome (DGS; OMIM #188400) comprises thymic hypoplasia, parathyroid hypoplasia, and outflow tract malformations of the heart. The syndrome is caused by a 1.8–3.0 Mb hemizygous deletion in chromosome 22q11.2. Haploinsufficiency of the *TBX1* gene (OMIM *602054), which is localized within the deleted region, can be solely responsible for this disorder. Less frequently, DGS can be caused by defects in other chromosomes, notably 10p13. DGS is one form of several phenotypically overlapping disorders, including conotruncal anomaly and velocardiofacial syndrome. Clinically, patients may present with neonatal hypocalcemia and susceptibility to opportunistic infections. Cardiac malfunctions typically include tetralogy of Fallot, truncus arteriosus, and interrupted aortic arch. Facial signs consist of retromicrognathia, small teeth, short filtrum, hypertelorism, and low-set and malformed ears. Pulmonary, renal, skeletal, ophthalmological, and GI defects can also occur.

Molecular diagnosis is made by using chromosomal microarray or array comparative genomic hybridization or by sequencing the *TBX1* gene. Because of the poor development of the thymic gland, patients with DGS usually have T-cell lymphopenia, which frequently improves with age but may never reach normal levels. A small group of patients are born with little to no detectable circulating T cells and are therefore designated as suffering from "complete DGS." Management is dictated by the presenting features of this heterogeneous syndrome. Calcium supplements and 1,25-cholecalciferol may be needed to treat hypocalcemia. Blood products, if needed (normally during heart surgery), should be CMV-free and irradiated to eliminate allogeneic T-cell transfer and to prevent GvHD. Thymic tissue transplantation has been attempted in cases with complete DGS, although the effects are difficult to evaluate, since children with DGS tend to improve spontaneously with age.

Wiskott-Aldrich Syndrome

Wiskott-Aldrich syndrome (WAS; OMIM #301000) is a rare X-linked recessive condition caused by mutations in the *WAS* gene. The WAS encoded protein is predominately expressed in hematopoietic cells, and its main function is to activate actin assembly by binding to the Arp2/3 complex. The WAS protein (WASP) is required for TCR-mediated formation of the immunological synapse, which is a process dependent on actin rearrangement. Patients are typically diagnosed in infancy, and their condition is characterized by thrombocytopenia, eczema, and recurrent infections.[47] Platelet volume is reduced and appears to be rapidly cleared from the circulation. The abnormal coagulation in these patients is associated with bloody stools and life-threatening intracranial bleeding. Patients with this syndrome occasionally have increased susceptibility to bacterial, viral, and opportunistic organisms, including *P. jiroveci* pneumonia. Autoimmunity and malignancy, particularly lymphoma, have been frequently reported in these patients. T-cell numbers tend to decrease over time, resulting in mild-to-moderate lymphopenia, and *in vitro* responses to mitogens are frequently normal. Typically, IgM is low, IgG is normal, whereas IgA and IgE are elevated. Splenectomy should be avoided. HSCT successfully cures these patients. Gene therapy for this condition is currently being studied.[48] Recently, a patient with a similar phenotype to WAS was shown to have a mutation in the WASP-interacting protein (WIP). WIP stabilizes WASP, preventing its degradation.

Ataxia Telangiectasia (Louis-Bar Syndrome)

Ataxia telangiectasia (AT; OMIM #208900) is an autosomal recessive multisystem disorder characterized by cerebellar ataxia, telangiectasias, immunodeficiency, and highly increased predisposition to malignancy. AT is a result of deleterious mutations in the *ATM* gene (OMIM *607585), which encodes a key protein involved in protecting DNA repair. ATM is a serine/threonine protein kinase that activates signaling pathways upon double-strand breaks, which occur naturally during Ig and TCR rearrangement or upon stress posed by ionizing radiation. It is also involved in cell cycle control and may act as a tumor suppressor gene.

Affected individuals present in early childhood with progressive truncal ataxia, which is frequently misdiagnosed in infancy as cerebral palsy. Choreoathetosis and dystonia, ocular motor apraxia, and dysarthria are common. At 3–5 years of age telangiectasias become increasingly prominent in the eyes and ear lobes. Typically, patients become wheelchair dependent during the second decade of life. Most patients die by the age of 30 years, with rare survivors who live to the end of the fourth decade. Patients die of progressive restrictive lung disease, infections, or cancer. Lymphoma, leukemia, and ovarian and stomach cancers are the most common malignancies associated with AT, and a high risk of developing breast cancer is a prominent feature of heterozygotes.

Diagnosis of this condition is supported by demonstrating increased frequency of chromosomal breakage and elevated levels of alpha-fetoprotein, which reflect the defect in DNA repair and liver tissue immaturity, respectively. Immune investigations show decreased circulating T cells and abnormal *in vitro* response to mitogens. The thymus is dysplastic, and TREC levels are markedly reduced. IgA is low in about 50% of patients, and IgG levels vary from extremely low to normal levels. IgM may be increased, a phenotype mimicking hyper-IgM syndrome (HIGM). The diagnosis is confirmed by identifying mutations in the *ATM* gene. Treatment is mostly supportive and limited to intravenous immunoglobulin (IVIG) and antibiotics, when indicated, as well as physiotherapy. Nijmegen breakage syndrome (OMIM #251260), an AT-like syndrome, is caused by biallelic hypomorphic mutations in *MRE11,* which encodes a protein involved in cell cycle control and double-strand DNA repair. In addition to the typical AT feature, Nijmegen breakage syndrome is characterized by microcephaly, bird-like facial features, and growth retardation.

COMBINED IMMUNODEFICIENCY WITH IMMUNOOSSEOUS DYSPLASIA

This group of combined immunodeficiencies characterized by walking and growth disturbances is summarized in Table 35.6.

Cartilage Hair Hypoplasia

McKusick-type metaphyseal chondrodysplasia, also known as cartilage hair hypoplasia (OMIM #250250), is an autosomal recessive syndrome that was first recognized in the Amish and later in the Finnish population. The condition is characterized by short-limb dwarfism, various degrees of T-cell immunodeficiency, fine and sparse hair, Hirschsprung disease, and increased risk of malignancy. This syndrome is caused by mutations in the RNase MRP (*RMRP*) gene (OMIM *157660), a ribonuclease that is present in both the nucleus and mitochondria. Uniquely, this gene product does not undergo translation but functions as a protein in the cleavage of RNA during mitochondrial DNA synthesis, as well as clearing of nuclear pre-rRNA. Immune function is variable, and patients may present with profound T-cell lymphopenia, depressed mitogenic responses, and a dysplastic thymus. Some cases present with Omenn syndrome (T^+B^+ SCID). Other patients have various degrees of T-cell lymphopenia. Nonimmune features consist of spondylometaphyseal dysplasia, small caliber and hypoplastic hair

TABLE 35.6 Combined Immunodeficiency With Immunoosseous Dysplasia

Disease	Inheritance	Molecular Defect/Presumed Pathogenesis	Features	Treatment
Cartilage hair hypoplasia	AR	Defect in *RMRP* gene resulting in impaired ribosomal assembly and cell cycle regulation	Metaphyseal dysplasia; fine and sparse hair	Hematopoietic stem cell transplantation (HSCT)
Schimke immunoosseous dysplasia	AR	Impaired DNA stress response enzyme	Epiphyseal dysplasia; renal dysfunction; autoimmunity; non-Hodgkin lymphoma	Supportive
Roifman immunoskeletal syndrome (SPENCD)	AR	Impaired tartrate-resistant phosphatase (TRAP) enzyme function, accumulation of phosphorylated osteopontin	Spondyloenchondrodysplasia; autoimmunity; neurological abnormalities	Supportive

with lack of central pigment core, and skin hypopigmentation. There is an increased risk of non-Hodgkin lymphoma and basal cell carcinoma with this syndrome. HSCT has proven curative of the immune abnormalities in these patients.

Schimke Immunoosseous Dysplasia

The major features of Schimke immunoosseous dysplasia (SIOD; OMIM #242900), an autosomal recessive syndrome, are spondyloepiphyseal dysplasia, progressive renal failure, and moderate cellular immunodeficiency. The disease is caused by mutations in the SWI/SNF matrix actin-dependent regulator of chromatin subfamily A-like gene (*SMARCAL1*) encoding a DNA stress response enzyme. Patients with SIOD present with growth retardation before and after birth. In addition to progressive renal failure, patients may have cerebral vascular accidents, dental and skin abnormalities, and dysmorphic features. Immune function is variable, with mild to moderate lymphopenia and somewhat reduced *in vitro* responses to mitogens. Consequently, they develop recurrent viral and fungal infections. Treatment is supportive, and some patients may benefit from renal transplantation.

Roifman Syndrome

Roifman Syndrome (OMIM #616651) was first described as a novel association of immunodeficiency, spondyloepiphyseal chondro-osseous dysplasia, retinal dystrophy, growth and developmental delay, and distinctive facial dysmorphism.[49]

This syndrome is caused by distinct mutations in the small nuclear RNA (snRNA) gene RNU4ATAC (OMIM*601428), which is essential for minor intron splicing. About 800 genes have one or more minor introns, and these are dependent on the minor spliceosome for correct splicing. Since many of these genes are involved in basic cellular functions, such as DNA replication and repair or RNA processing, incorrect splicing can alter cell function and survival. Mutations in the Stem II domain of the gene appear obligatory for this syndrome. Patients with Roifman Syndrome frequently suffer repeated microbial infections such as otitis media, cellulitis and pneumonia as well as viral and fungal infections (HSV, PJP).

Atopy and autoimmune manifestations are also common, including eczema, asthma, arthritis, hemolytic anemia, colitis, and autoimmune hepatitis. Antibody deficiency and hypogammaglobulinemia are present in all cases, and circulating B cells are usually reduced, especially memory B cells. T cell lymphopenia, low CD8+ T cells, abnormal T cell repertoire, as well as low in vitro responses to antigens are also recorded.

Infections are controlled with immunoglobulin replacement in these patients.

SPENCDI–Roifman Immunoskeletal Syndrome

This syndrome was first described in 2000 as a novel combination of features encompassing combined immunodeficiency, autoimmunity, and spondylometaphyseal dysplasia (OMIM # 607944).[50] One patient died of noninfectious encephalopathy, and some others had arthritis, SLE, and thrombocytopenia, in addition to lung disease. Renella et al.[45] highlighted the type of spondylometaphyseal dysplasia as spondyloenchondrodysplasia (SPENCD), confirming the initial observation of autoimmunity in this syndrome. The spectrum of clinical features was further expanded to include cerebral calcifications and other neurological abnormalities. SPENCD is a rare skeletal dysplasia characterized by metaphyseal and vertebral lesions consisting of chondroid tissue. Over the years, evidence for both autosomal dominant and recessive inheritance was proposed. Mice lacking tartrate-resistant phosphatase (TRAP) display skeletal abnormalities identical to the patients. Consequently, sequencing of the *ACP5* gene (OMIM *171640), which encodes TRAP, identified deleterious biallelic mutations in humans.[52,53] Immune abnormalities include T-cell lymphopenia with reduced responses to mitogens and antigens, and an inability to produce specific antibodies. Patients frequently have autoantibodies, such as antinuclear factor (ANF) and anti-DNA. Treatment so far has been limited to replacement of Igs, antibiotics, and immunomodulation of the various immune manifestations.

COMBINED IMMUNODEFICIENCY WITH BONE MARROW FAILURE

Immune deficiencies associated with bone marrow failure are presented in Table 35.7.

IKAROS Deficiency

IKAROS deficiency is extremely rare combined immunodeficiency associated with bone marrow failure (OMIM #616873). IKAROS is a hematopoietic-specific transcription factor (OMIM *603023) affecting pluripotent stem cells that control T-cell differentiation, B-cell V(D)J recombination, and NK-cell development.[54] Patients may present with recurrent pulmonary infections or disseminated viral infections in teenage years, or during infancy with bone marrow aplasia and near absent B and NK cells.[55] T cells can be normal in number, but the response to mitogens is absent

TABLE 35.7 Combined Immunodeficiency With Bone Marrow Failure

Disease	Inheritance	Molecular Defect/Presumed Pathogenesis	Features	Treatment
Dyskeratosis congenita	X-linked, AR, AD	Defective telomere processing proteins, genomic instability	Nail and skin defects; noninfectious enteropathy; developmental delay; microcephaly; cerebellar hypoplasia	Hematopoietic stem cell transplantation (HSCT)
IKAROS deficiency	AD	Impaired lymphocyte development and proliferation, V(D)J rearrangement	Recurrent infections; severe aplastic anemia	HSCT
AK2 deficiency (reticular dysgenesis)	AD	Impaired mitochondrial energy metabolism, severe lymphopenia	Severe marrow failure; bilateral sensorineural deafness	HSCT

TABLE 35.8 Progressive T-Cell Combined Immunodeficiency

Disease	Inheritance	Molecular Defect/Presumed Pathogenesis	Features	Treatment
STAT1 dysfunction	AR, AD	Impaired cytokine signaling	Autoimmunity	Hematopoietic stem cell transplantation (HSCT)
RelB deficiency	AR	Disruption of NF-κB noncanonical pathway	Failure to thrive; autoimmunity	HSCT
STK4 deficiency	AR	Abnormal apoptosis	Epstein-Barr virus (EBV) lymphoproliferation	HSCT

($T^+B^-NK^-$), and Igs are low. The thymus appears large and hyperplastic. HSCT is the only effective treatment proposed.[55]

Dyskeratosis Congenita

Dyskeratosis congenita (DKC; OMIM #305000) is a complex and fatal disorder caused by mutations in genes that encode telomere processing proteins. Inheritance can be X-linked in the case of mutations in dyskerin, the most common cause of DKC, autosomal recessive when mutations in telomerase reverse transcriptase (*TERT*), *NOP10, NHP2,* and *C160ORF57* are found, or autosomal dominant in cases with mutations in telomerase RNA component (*TERC*), *TERT,* and *TNF2.*[56]

Common hallmarks of this group of disorders are bone marrow failure, repeated infections, nail dystrophy, and leukoplakia. Pulmonary fibrosis and colitis are frequently observed. Some patients have additional syndromic features and present early in life with significant intrauterine growth retardation, short stature, microcephaly, and cerebellar hypoplasia, a constellation of features also known as *Hoyeraal-Hreidarsson syndrome.*[57] Other patients present in adolescence or later in life with progressive bone marrow failure. Immune abnormalities are highly variable, with a wide range of presentations from extreme T-cell lymphopenia and occasional profound B-cell lymphopenia (T^-B^- SCID) to mild lymphopenia. Similarly, T-cell function, serum Ig levels, and specific antibody formation are variable. Allogeneic HSCT has been attempted in patients with DKC, primarily for bone marrow failure or malignancy, with disappointing results because of drug toxicity and frequent failure of engraftment.

Moesin Deficiency (MSN Deficiency)

This is a newly described combined immunodeficiency which is inherited in an X-linked manner and caused by mutations in the cytoskeletal protein moesin (OMIM *309845). Moesin is a member of the ezrin-radixin-moesin family of proteins that link the cortical actin filaments to the plasma membrane. The active form of moesin is predominantly expressed in lymphocytes where it is believed to sustain cell shape. Moesin is dephosphorylated upon TCR or BCR binding, suggesting that it plays a role in TCR mediated cell polarization.

Immune evaluation in these patients revealed T cell lymphopenia, hypogammaglobulinemia, antibody deficiency and neutropenia.

Patients are typically susceptible to severe varicella zoster virus and recurrent molluscum contagiosum. Infections were well controlled with replacement immunoglobulin therapy.

PROGRESSIVE T-CELL OR COMBINED IMMUNODEFICIENCY

Immune deficiencies characterized by a progressive decline in T- and/or NK-cell numbers and functions are summarized in Table 35.8.

STAT1 Dysfunction

Mutations in the gene for the DNA-binding transcription factor—*STAT1* (OMIM *600555)—have been associated with an array of different clinical phenotypes ranging from complete deficiencies (autosomal recessive; OMIM #613796), which are fatal in early infancy, to chronic mucocutaneous candidiasis[58,59] or mild persistent oral thrush (autosomal dominant; OMIM #614892). One subgroup of patients with monoallelic *STAT1* mutations appear to have a distinct course. They suffer from repeated and severe viral and fungal infections that are ultimately fatal, and most of them die before the age of 20 years. Most have severe colitis as well as multiple autoimmune features. Evaluation of immunity is normal in these patients, featuring a progressive decline and loss of T cells and NK cells, decreasing T-cell functions, and a structurally abnormal thymus gland. Usually, by the end of the first decade of life, these patients have all of the typical clinical and immunological features of combined immunodeficiency. In this chapter, we will only describe the phenotype of STAT1 dysfunction.

The *STAT1* gene encodes two isoforms, α and β. STAT1β acts as a dominant negative inhibitor, whereas STAT1α plays a critical role in signaling responses to a variety of cytokines, such as interferon-α (IFN-α), IFN-β, IFN; IL-2, -3, -6, -9, -10, -11, -12, and -15; growth hormone fibroblast growth factor; epidermal growth factor; and others. Cytoplasmic STAT1 is activated through tyrosine phosphorylation by members of the JAK kinase family, which results in its translocation to the nucleus and subsequent binding to target DNA segments. Typically, these patients harbor *de novo* mutations in the DNA-binding domain of STAT1. Although increased tyrosine phosphorylation of STAT1 or even enhanced DNA binding may be observed, STAT1 function is altered and patients' cells produce reduced amounts of IL–2, IFN-γ, and other cytokines, hence the designation of STAT1 dysfunction.[60] HSCT has been attempted with variable degrees of success.[61,62]

RelB Deficiency

This is a newly described CID ($T^+B^+NK^+$) that is inherited in an autosomal recessive mode.[63] The immunodeficiency is caused by mutations in *RELB* (OMIM *604758), a key component of the NF-κB pathway. Lymphocyte antigen receptors, multiple cytokines, and infective agents all use the NF-κB signaling pathway.[64] The pathway consists of five members, including RelA, RelB, c-Rel, and the precursor proteins NF-κB1 (p105) and NF-κB2 (p100), which are processed to the active forms p50 and p52, respectively. Two NF-κB pathways are recognized, the classical pathway, mediated by RelA, and the alternative (non-canonical) pathway, which uses RelB. Activation of RelB requires inducible processing of p100 and p52. Unlike the classical pathway, which is engaged by many receptors, the alternative pathway is used by TNFRSF12A, LTβR, CD40, and BAFF-R.[37] Patients with RelB deficiency present in the first year of life with failure to thrive, recurrent infections, and autoimmune disorders. *In vitro* responses to mitogens are depressed, the TCR repertoire is restricted, and TRECs are abnormal. Production of specific antibodies in response to vaccination is impaired. The thymus appears moderately dysplastic. Both memory T and B cells are reduced or absent.[63,65,66] HSCT has been successful in two patients who received the procedure.

STK4 (Mst1) Deficiency

This is an autosomal recessive combined immunodeficiency characterized by a progressive loss of T cells and declining function (OMIM #614868). The deficiency is caused by biallelic mutations in serine threonine kinase 4 (*STK4*). STK4 functions as a pro-apoptotic factor by promoting Fas-mediated apoptosis. STK4-null mice demonstrate increased apoptosis leading to T- and B-cell lymphopenia, as well as depletion of the thymus gland. Patients present with oral thrush and mucocutaneous candidiasis, recurrent bacterial and viral lung infections, and EBV-induced B-cell lymphoproliferation. Skin manifestations include disseminated warts and molluscum contagiosum. Evaluation of the immune system reveals progressive decrease in T and B cells and an altered TCR Vβ repertoire. Serum Igs, including IgE, are elevated, and autoantibodies have been observed. Autoimmune cytopenias have also been reported. HSCT is the treatment of choice.

DIAGNOSIS, MANAGEMENT, AND TREATMENT OF T-CELL AND COMBINED IMMUNODEFICIENCY

The current approach to management and treatment of primary immunodeficiencies involves prevention and treatment of infections, and when attainable, restoration of a functional immune system. Early diagnosis is critical to minimizing morbidity and improving quality of life, and with the increasing implementation of NBSs, patients now have a greater chance of being diagnosed shortly after birth. Importantly, measures that minimize exposure of patients to infections, including use of irradiated and filtered blood products, avoidance of live vaccines, and nutritional support, should also be considered as part of therapy.

THERAPEUTIC PRINCIPLES

Management of Severe Combined Immunodeficiency

- Patients should be kept in a protective isolation environment, receive prophylactic antibiotics, immunoglobulin replacement, and enhanced nutrition.
- Patients should only receive cytomegalovirus (CMV)–free, irradiated blood products and irradiated breast milk. Breast feeding should be discontinued if the mother is CMV positive.
- Live attenuated vaccines should be avoided.
- The immune defect in severe combined immunodeficiency (SCID), and in many cases of combined immunodeficiency (CID), can be corrected by administration of allogeneic hematopoietic stem cell transplantation (HSCT).
- Best results are achieved using an human leukocyte antigen (HLA)–identical family donor.
- Patients with autologous T cell should receive myeloablative conditioning to improve engraftment and sustained long-term immune reconstitution.
- Gene therapy should be offered in selected types of SCID, such as adenosine deaminase (ADA) deficiency, whereas this procedure remains experimental in other conditions.

Newborn Screening

Patients with SCID usually die of overwhelming infections or suffer associated catastrophic complications once maternally transferred antibodies have been metabolized, usually within 3–6 months after birth. However, if diagnosed early and before the onset of complications, most of these patients can be treated successfully with HSCT.

TRECs, which are produced during programmed gene rearrangement of progenitor T cells in the thymus, were identified as a useful biomarker of both thymic function and the production of newly formed T cells. Since the thymus is invariably dysfunctional in SCID, the use of TREC detection in blood from Guthrie spots serves as an ideal NBS tool.[67]

NBS utilizing TRECs detects most cases of phenotypic SCID, especially those who have lymphocytopenia of less than 300–500 $CD3^+$ cells/μL. It may also detect other conditions where the patient presents at birth with profound lymphopenia—such as ataxia telangiectasia, Nijmegen breakage syndrome, Di George syndrome, CHARGE, Rac2 and DOCK8 deficiencies, as well as trisomy.[21]

However, cases that are phenotypically similar to SCID, such as ZAP-70 deficiency or other CIDs with a severe phenotype ($CD3^+$ >300–500 cells/μL) may be missed. This highlights the need consider other methodologies, such as next generation sequencing, which have the capacity to detect these cases at birth.

Isolation

Patients with primary immunodeficiency are susceptible to recurrent and opportunistic infections, posing an immediate threat to life. As a preventative measure, isolation and the provision of an aseptic environment are recommended and have been shown to protect against infections for extended periods. The prevention and successful treatment of infections are strong determinants of HSCT outcome[2] and is associated with decreased frequency of GvHD. In particular, infection prevention is critical in patients with SCID receiving matched unrelated HSCT, where the risk of GvHD has been reported to be >70% (the main cause of morbidity and mortality).[68]

Hematopoietic Stem Cell Transplantation

HSCT is the only potentially curative therapeutic approach for patients SCID and various forms of CID.[69] The replacement of defective HSCs with cells from a healthy donor, typically bone marrow, allows T-cell precursors to repopulate the native thymus and permits reconstitution of a functional immune system. As a result of advancements made in preparing the host, donor cells, conditioning regimens, and supportive therapy, HSCT survival rates have markedly improved since the early 1980s (Chapter 82).

Donor Source

Transplantation involving related HLA-identical donors (RID), usually a sibling, carry the most favorable prognosis, with survival rates of 90 to 100%. Complete immune reconstitution occurs rapidly and carries a low risk of GvHD. However, since such donors are available in less than 20% of cases, donor sources including HLA-matched unrelated donors (MUD) or HLA-mismatched related donors (MMRD) can be considered.

The survival rates for MUD HSCT are greater than 70%,[70] with some centers reporting rates of up to 83%[68,71] and low incidences of GvHD.[72] This modality has proven especially effective in cases with CID. Experience with MMRD HSCT in the past has been disappointing,[73,74] but recent reports suggest improved survival rates for these transplants, although follow-up in many of the cases is still too short.[2] Such transplants involving haploidentical donors require the marrow to be depleted of T cells, and immune reconstitution may be prolonged. Rates of graft failure necessitating a second transplant have been high, along with incomplete immune reconstituation.[72] Complications include opportunistic infections, chronic GVHD, autoimmunity and late deaths.

Nevertheless, regardless of donor source, HSCT prior to 3.5 months of age or before the onset of infection is associated with high rates of survival (95% and 90% respectively) when compared to older children (76%).[2]

Conditioning

Pre-transplant conditioning helps to prepare the recipient marrow for optimal engraftment of donor stem cells and T cell reconstitution, and plays an important role in the long-term outcome of HSCT.[75] Whether conditioning is necessary depends in part on the type of immunodeficiency. For example, whereas myeloablation is necessary for HSCT in patients with sufficient immune cells to reject a graft (as in CID), it may not be indicated for patients with SCID receiving HLA-matched stem cells from a sibling donor. However, it is worth noting that donor B cells rarely engraft in the absence of conditioning regimens, and many patients remain on lifelong immunoglobulin replacement with increased risk of infective lung damage and T lymphocyte senescence.

Myeloablative conditioning with busulfan and cyclophosphamide was once commonly used for HSCT in patients with SCID and CID, although a protocol including busulfan and fludarabine is currently proposed due to risk of veno-occlusive disease (VOD).[76] Alternative conditioning agents include treosulfan (1-treitol-1,4-bis-methanesulfonate), which has both myeloablative and immunosuppressive properties, as well as an improved patient safety profile.

Reduced intensity conditioning (RIC) involving incomplete ablation of the bone marrow may also be beneficial in permitting engraftment and immune reconstitution. However, where such regimens may offer improved survival and reduced toxicity, there have also been reports of increased incidences of mixed donor chimerism and graft failure. In these instances, repeat HSCT or donor lymphocyte infusions may be necessary, which may pose further risk to patients.[77]

HSCT without conditioning has been attempted in patients who are too sick to tolerate pre-transplant chemoablation, however, this is associated with a high frequency of complications and mortality.[78] In such circumstances, it is recommended that the patient be stabilized until a full protocol can be implemented.

Prophylaxis and Treatment of Graft versus Host Disease

GvHD is caused by T cell mediated attack of the host tissue by the graft, and is a significant cause of morbidity and mortality to allogeneic HSCT recipients. The principal affected organs are skin, gastrointestinal tract and liver, with the current grading system (grades I-IV) based upon the extent of organ involvement.[79] Risk factors for GvHD include donor-host incompatibility as well as the source of stem cells, and can be significantly reduced by introducing prophylactic immunosuppression and depletion of donor T cells before infusion. The primary response to such first-line treatment is the most important predictor of long term survival in patients with acute GvHD.

Options for pharmacologic prophylaxis include monotherapy with a calcineurin inhibitor (such as cyclosporine or tacrolimus), or in combination with steroids, methotrexate or mycopholate mofetil. The efficacy of steroids in improving the outcome and probability of survival is well documented.[80,81]

T cell depletion is recommended for HSCT involving HLA-mismatched donors. However, while depletion of donor T cells using physical separation techniques, monoclonal antibodies or anti-thymocyte globulin can reduce GvHD, it is also associated with delayed immune reconstitution and an increased incidence of post-transplant lymphoproliferative disorders and life-threatening infections.[79]

Gene Therapy

Gene therapy represents a therapeutic option for patients with primary immunodeficiency who lack a HLA-identical donor[62] (Chapter 85). By using retroviruses to introduce a functional copy of the defective gene into the patient's own HSCs, this technique negates the need for a donor and is not associated with complications normally arising from HLA-incompatibility.

To date, gene therapy has been demonstrated to be extremely effective in treating patients with ADA deficiency, resulting in a variable increase in T and B cell numbers as well as significant improvement of immune function. Importantly, there have been no reports of serious adverse events or malignancy to this indication.[23]

However, early attempts to utilize gene therapy in other forms of primary immunodeficiency were met with mixed results, and highlighted various pitfalls related to the first-generation vectors.[83] A number of serious adverse events, including leukemia, were caused by preferential integration of retroviral vectors in the proximity of transcription initiation sites, leading to increased and deregulated expression of oncogenes. Use of a γ-retrovirus vector in patients with Wiskott-Aldrich syndrome resulted in partial or complete resolution of immunodeficiency, however 7 of the 10 patients receiving gene therapy developed malignancy.[84] Furthermore, in patients with chronic granulomatous disease, long term immune reconstitution was variable and myelodysplastic syndrome was reported in 3 of 10 patients. Between 1999 and 2006, trials for IL2RG deficiency initially reported successful immune reconstitution in 18 of 20 recipients; however clonal proliferation caused by insertional mutagenesis occurred in 5 recipients after 2 to 5 years.[85] Taken together, these results prompted a discontinuation of these trials, and a call to develop novel and much safer vector constructs, including self-inactivating γ-retrovirus and lentiviral vectors. To date, follow-up studies in patients with IL2RG deficiency utilizing a self-inactivating γ-retrovirus vector have so far been encouraging, with reports of excellent immune reconstitution and no evidence of insertional mutagenesis.[83] In an ongoing study,[48] the application of lentiviral vectors in Wiskott-Aldrich syndrome have demonstrated multilineage engraftment of corrected cells, both in the bone marrow and peripheral blood, with no reports of clonal expansion or leukemia.

With the advancement in technology and progress made in vector design, gene therapy is a viable option for challenging forms of primary immunodeficiency, in which current available treatments are insufficient.

Gene Editing

Gene editing technologies allow precise correction of the defective genomic DNA site and are thus a promising alternative to HSCT (Chapter 85). Using engineered endonucleases, such CRISPR/Cas9, zinc finger nucleases, and transcription activator-like effector nucleases, double-stranded breaks are introduced into specific sequences of the genomic DNA and repaired through homology-directed repair. In the presence of a DNA repair template containing a functional version of the defective gene, the double-stranded break is accurately repaired with the corrected sequence incorporated. While gene editing provides greater control over site integration than retrovirus-based gene therapy, the technique is still in the pre-clinical phase, with developments being made to limit errors such as off-target mutagenesis. To date, zinc finger nucleases have been utilized to correct IL2RG mutational hotspots in hematopoietic and embryonic stem cells.[86,87] Transplantation of these cells into immunodeficient mice resulted in multilineage hematopoietic differentiation.[86] Successful attempts to correct SCID patient-specific induced pluripotent stem cells in-vitro have also been reported, although further studies are needed to address the tumorigenic potential of such cells.

CONCLUSIONS

The traditional clinical and laboratory definitions of profound T-cell deficiencies, although still helpful in many cases, are no longer sufficient. This is highlighted by gene defects, such as FOXP3 deficiency, a fatal T-cell deficiency with normal *in vitro* responses to mitogens as well as TREC levels. It is, therefore, important to broaden our definitions to include severe autoimmune/inflammatory conditions as main or prominent clinical features of profound T-cell deficiency, in addition to recurrent opportunistic infections and susceptibility to specific infectious agents. Other clinical features, such as granulomata and lymphoma, have also been reported recently as sole manifestations of T-cell immunodeficiencies. With the increased recognition of genetic aberrations linked to these diseases, we are likely to see a great expansion of novel phenotypes associated with known genotypes, although not necessarily with identical mutations.

ON THE HORIZON

- Whole exome sequencing can be technically performed using DNA derived from a dried blood spot. Implementing this technology to newborn screening will immensely improve the outcome of patients.
- In the future, closer attention will be paid to family members of patients with overt immunodeficiency to better understand the clinical impact of monoallelic/carrier genetic aberrations.
- Whole exome sequencing is widely reported to be successful in only 40% of cases. Whole-genome sequencing and development of new bioinformatics algorithms would likely yield better results.
- Although hematopoietic stem cell transplantation can cure patients with T-cell deficiency, in many cases, it remains a high-risk procedure and may not be optimal therapy for a growing number of newly defined combined immunodeficiencies. New developments in gene therapy and gene editing are likely to expand as alternative treatments.

This expansion in diagnosis of profound T-cell deficiencies raises new opportunities to study mechanisms underlying autoimmunity as well as lymphoid malignancies. However, these cases may pose new challenges, such as choice of HSC therapy. Further prospective trials should help to determine the validity of transplantation compared with conservative immunosuppressive therapy or innovative new biological treatments.

With challenges come opportunities, and we are living in an exciting era of improved tools to enhance the understanding of the molecular changes and mechanisms leading to profound T-cell deficiency and associated autoimmunity and malignancy. This wealth of knowledge should help advance novel treatments.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Sullivan KE, Stiehm ER Stiehm's immune deficiencies, 2014.
2. Pai SY, Logan BR, Griffith LM, et al. Transplantation outcomes for severe combined immunodeficiency, 2000-2009. N Engl J Med [Multicenter Study Research Support, N.I.H., Extramural Research Support, Non-U.S. Gov't]. 2014;371(5):434–46.
3. Shearer WT, Dunn E, Notarangelo LD, et al. Establishing diagnostic criteria for severe combined immunodeficiency disease (SCID), leaky SCID, and Omenn syndrome: the Primary Immune Deficiency Treatment Consortium experience. J Allergy Clin Immunol [Research Support, N.I.H., Extramural Research Support, N.I.H., Intramural Research Support, Non-U.S. Gov't]. 2014;133(4):1092–8.
4. Roifman CM, Somech R, Kavadas F, et al. Defining combined immunodeficiency. J Allergy Clin Immunol [Research Support, Non-U.S. Gov't]. 2012;130(1):177–83.
5. van der Spek J, Groenwold RH, van der Burg M, et al. TREC Based Newborn Screening for Severe Combined Immunodeficiency Disease: A Systematic Review. J Clin Immunol [Meta-Analysis Review]. 2015;35(4):416–30.
6. Arpaia E, Shahar M, Dadi H, et al. Defective T cell receptor signaling and CD8+ thymic selection in humans lacking zap-70 kinase. Cell [Research Support, Non-U.S. Gov't]. 1994;76(5):947–58.
7. Roifman CM, Dadi H, Somech R, et al. Characterization of zeta-associated protein, 70 kd (ZAP70)-deficient human lymphocytes. J Allergy Clin Immunol [Research Support, Non-U.S. Gov't]. 2010;126(6):1226–33.e1.

8. Somech R, Simon AJ, Lev A, et al. Reduced central tolerance in Omenn syndrome leads to immature self-reactive oligoclonal T cells. J Allergy Clin Immunol 2009;124(4):793–800.
9. Villa A, Notarangelo LD, Roifman CM. Omenn syndrome: inflammation in leaky severe combined immunodeficiency. J Allergy Clin Immunol [Review]. 2008;122(6):1082–6.
10. Poliani PL, Facchetti F, Ravanini M, et al. Early defects in human T-cell development severely affect distribution and maturation of thymic stromal cells: possible implications for the pathophysiology of Omenn syndrome. Blood [Research Support, N.I.H., Extramural Research Support, Non-U.S. Gov't]. 2009;114(1):105–8.
11. Nahum A, Reid B, Grunebaum E, et al. Matched unrelated bone marrow transplant for Omenn syndrome. Immunol Res [Research Support, Non-U.S. Gov't]. 2009;44(1-3):25–34.
12. Sharfe N, Shahar M, Roifman CM. An interleukin-2 receptor gamma chain mutation with normal thymus morphology. J Clin Invest [Case Reports Research Support, Non-U.S. Gov't]. 1997;100(12):3036–43.
13. Marcus N, Takada H, Law J, et al. Hematopoietic stem cell transplantation for CD3delta deficiency. J Allergy Clin Immunol [Multicenter Study Research Support, N.I.H., Extramural mResearch Support, Non-U.S. Gov't]. 2011;128(5):1050–7.
14. Pessach IM, Notarangelo LD. Gene therapy for primary immunodeficiencies: looking ahead, toward gene correction. J Allergy Clin Immunol [Research Support, N.I.H., Extramural Research Support, Non-U.S. Gov't Review]. 2011;127(6):1344–50.
15. van der Burg M, van Dongen JJM, van Gent DC. DNA-PKcs deficiency in human: long predicted, finally found. Curr Opin Allergy Clin Immunol 2009;9(6):503–9.
16. Riballo E, Critchlow SE, Teo SH, et al. Identification of a defect in DNA ligase IV in a radiosensitive leukaemia patient. Curr Biol [Comparative Study Research Support, Non-U.S. Gov't]. 1999;9(13):699–702.
17. Lagresle-Peyrou C, Six EM, Picard C, et al. Human adenylate kinase 2 deficiency causes a profound hematopoietic defect associated with sensorineural deafness. Nat Genet [Research Support, N.I.H., Intramural Research Support, Non-U.S. Gov't]. 2009;41(1):106–11.
18. Pannicke U, Honig M, Hess I, et al. Reticular dysgenesis (aleukocytosis) is caused by mutations in the gene encoding mitochondrial adenylate kinase 2. Nat Genet [Research Support, Non-U.S. Gov't]. 2009;41(1):101–5.
19. Roifman CM, Zhang J, Atkinson A, et al. Adenosine deaminase deficiency can present with features of Omenn syndrome. J Allergy Clin Immunol [Case Reports Letter Research Support, Non-U.S. Gov't]. 2008;121(4): 1056–8.
20. Somech R, Lai YH, Grunebaum E, et al. Polyethylene glycol-modified adenosine deaminase improved lung disease but not liver disease in partial adenosine deaminase deficiency. J Allergy Clin Immunol [Case Reports Letter Research Support, Non-U.S. Gov't]. 2009;124(4):848–50.
21. Husain M, Grunebaum E, Naqvi A, et al. Burkitt's lymphoma in a patient with adenosine deaminase deficiency-severe combined immunodeficiency treated with polyethylene glycol-adenosine deaminase. J Pediatr [Case Reports Research Support, Non-U.S. Gov't]. 2007;151(1):93–5.
22. Gaspar HB, Aiuti A, Porta F, et al. How I treat ADA deficiency. Blood [Research Support, N.I.H., Intramural Review]. 2009;114(17):3524–32.
23. Aiuti A, Cattaneo F, Galimberti S, et al. Gene therapy for immunodeficiency due to adenosine deaminase deficiency. N Engl J Med [Clinical Trial, Phase I Clinical Trial, Phase II Multicenter Study Research Support, Non-U.S. Gov't]. 2009;360(5):447–58.
24. Fischer A, Picard C, Chemin K, et al. ZAP70: a master regulator of adaptive immunity. Semin Immunopathol [Research Support, Non-U.S. Gov't Review]. 2010;32(2):107–16.
25. Chan AY, Punwani D, Kadlecek TA, et al. A novel human autoimmune syndrome caused by combined hypomorphic and activating mutations in ZAP-70. J Exp Med [Research Support, N.I.H., Extramural]. 2016;213(2): 155–65.
26. Newell A, Dadi H, Goldberg R, et al. Diffuse large B-cell lymphoma as presenting feature of Zap-70 deficiency. J Allergy Clin Immunol [Case Reports Letter Research Support, Non-U.S. Gov't]. 2011;127(2):517–20.
27. Dobbs K, Dominguez Conde C, Zhang SY, et al. Inherited DOCK2 Deficiency in Patients with Early-Onset Invasive Infections. N Engl J Med [Case Reports Research Support, N.I.H., Extramural Research Support, Non-U.S. Gov't]. 2015;372(25):2409–22.
28. Stepensky P, Keller B, Buchta M, et al. Deficiency of caspase recruitment domain family, member 11 (CARD11), causes profound combined immunodeficiency in human subjects. J Allergy Clin Immunol [Case Reports Research Support, Non-U.S. Gov't]. 2013;131(2):477–85.e1.
29. McKinnon ML, Rozmus J, Fung SY, et al. Combined immunodeficiency associated with homozygous MALT1 mutations. J Allergy Clin Immunol [Case Reports Clinical Trial Letter Research Support, Non-U.S. Gov't]. 2014;133(5):1458–62, 62.e1–7.
30. Chen R, Giliani S, Lanzi G, et al. Whole-exome sequencing identifies tetratricopeptide repeat domain 7A (TTC7A) mutations for combined immunodeficiency with intestinal atresias. J Allergy Clin Immunol [Research Support, N.I.H., Extramural Research Support, Non-U.S. Gov't]. 2013;132(3):656–64.e17.
31. Ngan B, Merico D, Marcus N, et al. Mutations in tetratricopeptide repeat domain 7A (TTC7A) are associated with combined immunodeficiency with dendriform lung ossification but no intestinal atresia. Lympho Sign J 2014;01(01):10–26.
32. Avitzur Y, Guo C, Mastropaolo LA, et al. Mutations in tetratricopeptide repeat domain 7A result in a severe form of very early onset inflammatory bowel disease. Gastroenterology [Research Support, N.I.H., Extramural Research Support, Non-U.S. Gov't]. 2014;146(4):1028–39.
33. Lacruz RS, Feske S. Diseases caused by mutations in ORAI1 and STIM1. Ann N Y Acad Sci [Research Support, N.I.H., Extramural Review]. 2015;1356:45–79.
34. Byun M, Abhyankar A, Lelarge V, et al. Whole-exome sequencing-based discovery of STIM1 deficiency in a child with fatal classic Kaposi sarcoma. J Exp Med [Case Reports Research Support, N.I.H., Extramural Research Support, Non-U.S. Gov't]. 2010;207(11):2307–12.
35. Gambineri E, Perroni L, Passerini L, et al. Clinical and molecular profile of a new series of patients with immune dysregulation, polyendocrinopathy, enteropathy, X-linked syndrome: inconsistent correlation between forkhead box protein 3 expression and disease severity. J Allergy Clin Immunol [Multicenter Study Research Support, Non-U.S. Gov't]. 2008;122(6):1105–12.e1.
36. Glocker EO, Kotlarz D, Boztug K, et al. Inflammatory bowel disease and mutations affecting the interleukin-10 receptor. N Engl J Med [Case Reports Research Support, N.I.H., Intramural Research Support, Non-U.S. Gov't]. 2009;361(21):2033–45.
37. Oeckinghaus A, Hayden MS, Ghosh S. Crosstalk in NF-kappaB signaling pathways. Nat Immunol [Review]. 2011;12(8):695–708.
38. Pannicke U, Baumann B, Fuchs S, et al. Deficiency of innate and acquired immunity caused by an IKBKB mutation. N Engl J Med [Case Reports Research Support, Non-U.S. Gov't]. 2013;369(26):2504–14.
39. Zhang Q, Davis JC, Lamborn IT, et al. Combined immunodeficiency associated with DOCK8 mutations. N Engl J Med [Research Support, N.I.H., Intramural Research Support, Non-U.S. Go 't]. 2009;361(21):2046–55.
40. Crequer A, Troeger A, Patin E, et al. Human RHOH deficiency causes T cell defects and susceptibility to EV-HPV infections. J Clin Invest [Research Support, N.I.H., Extramural Research Support, Non-U.S. Gov't]. 2012;122(9):3239–47.
41. Li FY, Chaigne-Delalande B, Kanellopoulou C, et al. Second messenger role for Mg2+ revealed by human T-cell immunodeficiency. Nature [Research Support, N.I.H., Intramural]. 2011;475(7357):471–6.
42. Martin E, Palmic N, Sanquer S, et al. CTP synthase 1 deficiency in humans reveals its central role in lymphocyte proliferation. Nature [Research Support, Non-U.S. Gov't]. 2014;510(7504):288–92.
43. Huck K, Feyen O, Niehues T, et al. Girls homozygous for an IL-2-inducible T cell kinase mutation that leads to protein deficiency develop fatal EBV-associated lymphoproliferation. J Clin Invest [Case Reports Research Support, Non-U.S. Gov't]. 2009;119(5):1350–8.
44. Pachlopnik Schmid J, Canioni D, Moshous D, et al. Clinical similarities and differences of patients with X-linked lymphoproliferative syndrome

type 1 (XLP-1/SAP deficiency) versus type 2 (XLP-2/XIAP deficiency). Blood [Research Support, Non-U.S. Gov't]. 2011;117(5):1522–9.
45. Booth C, Gilmour KC, Veys P, et al. X-linked lymphoproliferative disease due to SAP/SH2D1A deficiency: a multicenter study on the manifestations, management and outcome of the disease. Blood [Multicenter Study Research Support, Non-U.S. Gov't]. 2011;117(1): 53–62.
46. Worthey EA, Mayer AN, Syverson GD, et al. Making a definitive diagnosis: successful clinical application of whole exome sequencing in a child with intractable inflammatory bowel disease. Genet Med [Case Reports Research Support, Non-U.S. Gov't]. 2011;13(3): 255–62.
47. Aldrich RA, Steinberg AG, Campbell DC. Pedigree demonstrating a sex-linked recessive condition characterized by draining ears, eczematoid dermatitis and bloody diarrhea. Pediatrics 1954;13(2):133–9.
48. Aiuti A, Biasco L, Scaramuzza S, et al. Lentiviral hematopoietic stem cell gene therapy in patients with Wiskott-Aldrich syndrome. Science [Clinical Trial, Phase I Clinical Trial, Phase II Research Support, Non-U.S. Gov't]. 2013;341(6148):1233151.
49. Roifman CM. Immunological aspects of a novel immunodeficiency syndrome that includes antibody deficiency with normal immunoglobulins, spondyloepiphyseal dysplasia, growth and developmental delay, and retinal dystrophy. J Allergy Clin Immunol 1997;2:94–8.
50. Roifman CM, Costa TC. A novel syndrome including combined immunodeficiency, autoimmunity, and spondylometaphyseal dysplasia. J Allergy Clin Immunol 2000;5(1):6–9.
51. Renella R, Superti-Furga A. A new case of spondyloenchondrodysplasia with immune dysregulation confirms the pleiotropic nature of the disorder: comment on "A syndrome of immunodeficiency, autoimmunity, and spondylometaphyseal dysplasia" by M.L. Kulkarni, K. Baskar, and P.M. Kulkarni [2006]. Am J Med Genet A [Comment Letter]. 2007;143A(12):1394–5.
52. Briggs TA, Rice GI, Daly S, et al. Tartrate-resistant acid phosphatase deficiency causes a bone dysplasia with autoimmunity and a type I interferon expression signature. Nat Genet [Research Support, Non-U.S. Gov't]. 2011;43(2):127–31.
53. Lausch E, Janecke A, Bros M, et al. Genetic deficiency of tartrate-resistant acid phosphatase associated with skeletal dysplasia, cerebral calcifications and autoimmunity. Nat Genet [Research Support, Non-U.S. Gov't]. 2011;43(2):132–7.
54. Georgopoulos K, Moore DD, Derfler B. Ikaros, an early lymphoid-specific transcription factor and a putative mediator for T cell commitment. Science [Research Support, Non-U.S. Gov't]. 1992;258(5083):808–12.
55. Goldman FD, Gurel Z, Al-Zubeidi D, et al. Congenital pancytopenia and absence of B lymphocytes in a neonate with a mutation in the Ikaros gene. Pediatr Blood Cancer [Case Reports Research Support, N.I.H., Extramural Research Support, Non-U.S. Gov't]. 2012;58(4):591–7.
56. Jyonouchi S, Forbes L, Ruchelli E, et al. Dyskeratosis congenita: a combined immunodeficiency with broad clinical spectrum—a single-center pediatric experience. Pediatr Allergy Immunol [Review]. 2011;22(3):313–19.
57. Walne AJ, Vulliamy T, Kirwan M, et al. Constitutional mutations in RTEL1 cause severe dyskeratosis congenita. Am J Hum Genet [Research Support, Non-U.S. Gov't]. 2013;92(3):448–53.
58. van de Veerdonk FL, Plantinga TS, Hoischen A, et al. STAT1 mutations in autosomal dominant chronic mucocutaneous candidiasis. N Engl J Med [Research Support, Non-U.S. Gov't]. 2011;365(1):54–61.
59. Liu L, Okada S, Kong XF, et al. Gain-of-function human STAT1 mutations impair IL-17 immunity and underlie chronic mucocutaneous candidiasis. J Exp Med [Research Support, N.I.H., Extramural Research Support, Non-U.S. Gov't]. 2011;208(8):1635–48.
60. Sharfe N, Nahum A, Newell A, et al. Fatal combined immunodeficiency associated with heterozygous mutation in STAT1. J Allergy Clin Immunol [Case Reports Research Support, Non-U.S. Gov't]. 2014;133(3):807–17.
61. Faitelson Y, Bates A, Shroff M, et al. LymphoSign Journal. 2014;1(2):87–95.
62. Grunebaum E, Kim VH, Somers GR, et al. Bone marrow transplantation for monoallelic signal transducer and activator of transcription 1 deficiency. J Allergy Clin Immunol 2016.
63. Merico D, Sharfe N, Hu P, et al. RelB deficiency causes combined immunodeficiency. LymphoSign Journal. 2015;2(3):147–55.
64. Millet P, McCall C, Yoza B. RelB: an outlier in leukocyte biology. J Leukoc Biol [Research Support, N.I.H., Extramural Review]. 2013;94(5):941–51.
65. Haynes A, Kamra P, Roifman CM Heterozygous mutations in RelB can be associated with immune dysregulation and lymphoma. LymphoSign Journal. 2016;3(2):55–60.
66. Sharfe N, Merico D, Karanxha A, et al. The effects of RelB deficiency on lymphocyte development and function. J Autoimmun [Research Support, Non-U.S. Gov't]. 2015;65:90–100.
67. Puck JM. Laboratory technology for population-based screening for severe combined immunodeficiency in neonates: the winner is T-cell receptor excision circles. J Allergy Clin Immunol [Research Support, N.I.H., Extramural Research Support, Non-U.S. Gov't]. 2012;129(3):607–16.
68. Grunebaum E, Mazzolari E, Porta F, et al. Bone marrow transplantation for severe combined immune deficiency. JAMA [Comparative Study Multicenter Study Research Support, Non-U.S. Gov't]. 2006;295(5): 508–18.
69. Gennery AR, Slatter MA, Grandin L, et al. Transplantation of hematopoietic stem cells and long-term survival for primary immunodeficiencies in Europe: entering a new century, do we do better? J Allergy Clin Immunol [Historical Article]. 2010;126(3):602–10.e1-11.
70. Dalal I, Reid B, Doyle J, et al. Matched unrelated bone marrow transplantation for combined immunodeficiency. Bone Marrow Transplant [Multicenter Study Research Support, Non-U.S. Gov't]. 2000;25(6):613–21.
71. Mazzolari E, Forino C, Guerci S, et al. Long-term immune reconstitution and clinical outcome after stem cell transplantation for severe T-cell immunodeficiency. J Allergy Clin Immunol [Research Support, Non-U.S. Gov't]. 2007;120(4):892–9.
72. Roifman CM, Grunebaum E, Dalal I, et al. Matched unrelated bone marrow transplant for severe combined immunodeficiency. Immunol Res 2007;38(1-3):191–200.
73. Antoine C, Muller S, Cant A, et al. Long-term survival and transplantation of haemopoietic stem cells for immunodeficiencies: report of the European experience 1968-99. Lancet [Research Support, Non-U.S. Gov't]. 2003;361(9357):553–60.
74. Caillat-Zucman S, Le Deist F, Haddad E, et al. Impact of HLA matching on outcome of hematopoietic stem cell transplantation in children with inherited diseases: a single-center comparative analysis of genoidentical, haploidentical or unrelated donors. Bone Marrow Transplant [Comparative Study Research Support, Non-U.S. Gov't]. 2004;33(11): 1089–95.
75. Patel NC, Chinen J, Rosenblatt HM, et al. Outcomes of patients with severe combined immunodeficiency treated with hematopoietic stem cell transplantation with and without preconditioning. J Allergy Clin Immunol [Comparative Study]. 2009;124(5):1062–9, e1-4.
76. Ben-Barouch S, Cohen O, Vidal L, et al. Busulfan fludarabine vs busulfan cyclophosphamide as a preparative regimen before allogeneic hematopoietic cell transplantation: systematic review and meta-analysis. Bone Marrow Transplant 2016;51(2):232–40.
77. Chiesa R, Veys P. Reduced-intensity conditioning for allogeneic stem cell transplant in primary immune deficiencies. Expert Rev Clin Immunol [Review]. 2012;8(3):255–66, quiz 67.
78. Dvorak CC, Hassan A, Slatter MA, et al. Comparison of outcomes of hematopoietic stem cell transplantation without chemotherapy conditioning by using matched sibling and unrelated donors for treatment of severe combined immunodeficiency. J Allergy Clin Immunol [Comparative Study Multicenter Study Research Support, N.I.H., Extramural Research Support, Non-U.S. Gov't]. 2014;134(4):935–43.e15.
79. Paczesny S, Choi SW, Ferrara JL. Acute graft-versus-host disease: new treatment strategies. Curr Opin Hematol [Review]. 2009;16(6):427–36.

80. Messina C, Faraci M, de Fazio V, et al. Prevention and treatment of acute GvHD. Bone Marrow Transplant [Review]. 2008;41(Suppl. 2):S65–70.
81. Somech R, Kavadas FD, Atkinson A, et al. High-dose methylprednisolone is effective in the management of acute graft-versus-host disease in severe combined immune deficiency. J Allergy Clin Immunol [Letter]. 2008;122(6):1215–16.
82. Fischer A, Hacein-Bey Abina S, Touzot F, et al. Gene therapy for primary immunodeficiencies. Clin Genet [Review]. 2015;88(6):507–15.
83. Hacein-Bey-Abina S, Pai SY, Gaspar HB, et al. A modified gamma-retrovirus vector for X-linked severe combined immunodeficiency. N Engl J Med [Clinical Trial, Phase I Clinical Trial, Phase II Multicenter Study Research Support, N.I.H., Extramural Research Support, Non-U.S. Gov't]. 2014;371(15):1407–17.
84. Braun CJ, Boztug K, Paruzynski A, et al. Gene therapy for Wiskott-Aldrich syndrome—long-term efficacy and genotoxicity. Sci Transl Med [Clinical Trial, Phase I Clinical Trial, Phase II Research Support, Non-U.S. Gov't]. 2014;6(227):227ra33.
85. Cavazzana-Calvo M, Fischer A, Hacein-Bey-Abina S, et al. Gene therapy for primary immunodeficiencies: Part 1. Curr Opin Immunol [Comparative Study Research Support, Non-U.S. Gov't Review]. 2012;24(5):580–4.
86. Genovese P, Schiroli G, Escobar G, et al. Targeted genome editing in human repopulating haematopoietic stem cells. Nature [Research Support, Non-U.S. Gov't]. 2014;510(7504):235–40.
87. Urnov FD, Miller JC, Lee YL, et al. Highly efficient endogenous human gene correction using designed zinc-finger nucleases. Nature 2005;435(7042):646–51.

MULTIPLE-CHOICE QUESTIONS

1. Each of these comments about T-cell deficiencies (severe combined immunodeficiency/combined immunodeficiency [SCID/CID]) are TRUE, except:
   A. Patients may present in infancy with failure to thrive.
   B. The thymus always appears dysplastic.
   C. Malignancy could be the sole presenting feature.
   D. Clinical manifestations may first arise during late childhood or adolescence.
   E. Patients may present with severe hypocalcaemia.

2. Which of the following is INCORRECT?
   A. All patients who present with a SCID phenotype can be detected at birth by T-cell receptor excision circle (TREC) analysis
   B. *In vitro* responses to mitogens can be normal in some cases.
   C. Evaluation of natural killer (NK) cell numbers and function can be beneficial for diagnosis.
   D. Serum immunoglobulin levels can be normal.
   E. CH50 does not contribute to the diagnosis.

3. For which of the following conditions is hematopoietic stem cell transplantation not recommended?
   A. FOXP3 deficiency
   B. Ataxia telangiectasia
   C. Major histocompatibility complex (MHC) class II deficiency
   D. Adenosine deaminase (ADA) deficiency
   E. DOCK2 deficiency

36

Immunodeficiencies at the Interface of Innate and Adaptive Immunity

Jacinta Bustamante, Shen-Ying Zhang, Bertrand Boisson, Michael Ciancanelli,
Emmanuelle Jouanguy, Stéphanie Dupuis-Boisson, Anne Puel, Capucine Picard,
Jean-Laurent Casanova

In the last 20 years, new primary immunodeficiencies (PIDs) affecting the immunity mediated by interferon (IFN)- γ, IFN-α/β-λ, Toll and interleukin-1 receptor (TIR) domain nuclear factor (NF)-κB, Toll-like receptor (TLR)-3 pathway, and interleukin (IL)-17 have been identified. Some of these genetic defects are "conventional" PIDs, associated with a broad range of infections, but others provide a molecular explanation for severe pediatric infectious diseases previously thought to be idiopathic (Table 36.1). These "nonconventional" PIDs may be associated with severe and/or recurrent infections caused by a single family of microorganisms, a situation strongly contrasting with that for "conventional" PIDs. Standard immunological explorations are generally normal in these patients, whether they are susceptible to one or many infectious agents. Despite the lack of a clear immunological abnormality, infections in these patients are typically severe, and often fatal. This chapter is devoted to the description of these PIDs. They include disorders of the IFN-γ circuit associated with the syndrome of mendelian susceptibility to mycobacterial disease (MSMD).[1,2] Combined disorders of IFN-γ- and IFN-α/β-λ-mediated immunity are associated with mycobacterial and viral diseases. We also describe genetic defects affecting primarily the TLR-3 pathway.[3-9] The predominant infectious phenotype of patients with these defects is herpes simplex virus (HSV)-1 encephalitis (HSE) in childhood. We discovered the first genetic defect associated with severe isolated influenza, and mutation of *IRF7.*[10] We describe four PIDs associated with impaired signaling downstream from or via the TIR canonical pathway as a result of mutations of *IRAK4, MYD88, NEMO,* and *NFKBIA.*[11-15] Mutations of *NEMO* and *NFKBIA* also impair the alternative, Toll/IL-1R homologous region (TIR) domain–containing adapter-inducing IFN-β (TRIF)–dependent pathway.[16] The principal infectious phenotype of patients with any of these defects is the occurrence of pyogenic bacterial infections. Finally, we have described a new group of defects affecting IL-17 pathways and conferring predisposition to chronic mucocutaneous candidiasis (CMC).[17-20] All these disorders were initially thought to be rare, but they have since been diagnosed in about 800 patients around the world (see Table 36.1).

KEY CONCEPTS

- New primary immunodeficiencies (PIDs) should be sought in patients with unexplained infectious diseases.
- Children with severe infectious diseases should be repeatedly investigated for known and unknown immunodeficiency conditions.
- The exploration of idiopathic infections leads to the discovery of new PIDs and to a better understanding of immunity to pathogens.

MENDELIAN SUSCEPTIBILITY TO MYCOBACTERIAL DISEASE: GENETIC DISORDERS OF THE IFN-γ CIRCUIT

Genetic disorders of the IFN-γ circuit are associated with a selective susceptibility to weakly pathogenic mycobacteria, such as environmental mycobacteria (EM) and/or Bacille Calmette-Guérin (BCG) vaccines, and *Salmonella* (MSMD, Online mendelian Inheritance in Man [OMIM]: 209950) (Fig. 36.1).[21] These PIDs are caused by mutations in 10 genes involved in IFN-γ–mediated immunity: *IFNGR1* and *IFNGR2,* encoding the two chains of the receptor for IFN-γ, a pleiotropic cytokine secreted by natural killer (NK) and T cells; *STAT1,* encoding a transcription factor essential to the IFN-γR signaling pathway; *IL12B,* encoding the p40 subunit of IL-12 and IL-23, IFN-γ–induced cytokines secreted by macrophages and dendritic cells (DCs); *IL12RB1,* encoding the β_1 chain of the receptor for IL-12 and IL-23, which is expressed on NK and T cells; *NEMO,* encoding NF-κB essential modulator (NEMO), which is involved in the CD40-dependent induction of IL-12; *IRF8,* encoding an IFN regulatory factor inducible by IFN-γ; *TYK2,* which is involved in IL-12–dependent IFN-γ immunity; and *ISG15,* encoding an IFN-γ-inducing molecule that acts in synergy with IL-12.[1,2] The MSMD-causing mutations in *CYBB* affect the respiratory burst selectively in macrophages.[22] Some mutations of *STAT1* and most mutations of *NEMO* are associated with a much broader range of infectious diseases (see below), and complete interferon regulatory factor-8 (IRF8) deficiency is associated with a lack of circulating monocytes and DCs, and severe clinical disease, mimicking combined immunodeficiency (CID).[1,23,24] The molecular and clinical features of MSMD have been reviewed elsewhere.[1]

Complete IFN-γR1 and IFN-γR2 Deficiencies

Autosomal recessive (AR) complete IFN-γ receptor 1 (IFN-γR1) deficiency (OMIM 107470) is caused by recessive null *IFNGR1* mutations precluding the expression of IFN-γR1 on the cell surface or the recognition of its ligand, IFN-γ, by surface-expressed receptors. Affected patients fail to respond to IFN-γ and have high serum concentrations of IFN-γ after infection.[1] Mutations of the *IFNGR2* gene encoding IFN-γR2 lead to a complete loss of cellular responsiveness to IFN-γ because of a lack of receptor

TABLE 36.1 **New Inherited Disorders**

			INFECTIONS							
Gene	**Form**	**Inheritance**	**Mycobacteria**	***Salmonella***	**Viruses**	**HSE**	**Pyogenic Bacteria**	**Fungi**	**EDA**	**Inflammatory Signs**
IFNRG1	Amorphic	AR	++	+	+	–	–	–	–	N
	Hypomorphic	AR	++	+	–	–	+	–	–	N
	Hypomorphic	AD	++	+	–	–	–	+/–[a]	–	N
IFNGR2	Amorphic	AR	++	–	+	–	–	–	–	N
	Hypomorphic	AR	++	–	–	–	–	–	–	N
IRF8	Amorphic	AR	++	–	–	–	–	–	–	N
	Hypomorphic	AD	++	–	–	–	–	–	–	N
IL12RB1	Amorphic	AR	++	++	–	–	–	+/–[b]	–	N
IL12B	Amorphic	AR	++	++	–	–	–	–	–	N
STAT1	Hypomorphic	AD	++	–	–	–	–	–	–	N
	Hypomorphic	AR	++	–	+	–	–	–	–	N
	Amorphic	AR	++	–	++	+	–	–	–	N
	Hypermorphic	AD	–	–	–	–	–	++	–	N
ISG15	Amorphic	AR	++	–	–	–	–	–	–	N
TYK2	Amorphic	AR	++	–	+	–	–	–	–	N
					+					
CYBB	Hypomorphic	XR	++	–	–	–	–	–	–	N
TLR3	Hypomorphic	AD	–	–	+	++	–	–	–	N
UNC93B1	Amorphic	AR	–	–	+	++	–	–	–	N
TRAF3	Hypomorphic	AD	–	–	+	++	–	–	–	N
TRIF	Amorphic	AR	–	–	–	++	–	–	–	N
	Hypomorphic	AD	–	–	–	++	–	–	–	N
TBK1	Hypomorphic	AD	–	–	–	++	–	–	–	N
IRF3	Hypomorphic	AD	–	–	–	++	–	–	–	N
IRF7	Amorphic	AR	–	–	Flu	–	–	–	–	N
NEMO	Hypomorphic	XR	+	+	+	+	++	+	+/–	Weak
NFKBIA	Hypermorphic	AD	–	+	+	+	++	+	+	Weak
IRAK4	Amorphic	AR	–	–	–	–	++	–	–	Weak
MyD88	Amorphic	AR	–	–	–	–	++	–	–	Weak
HOIL	Amorphic	AR	–	–	+	–	++	–	–	Strong
HOIP	Hypomorphic	AR	–	–	+	–	++	–	–	Strong
IL17RA	Amorphic	AR	–	–	–	–	+	++	–	N
IL17RC	Amorphic	AR	–	–	–	–	–	++	–	N
IL17F	Hypomorphic	AD	–	–	–	–	+	++	–	N
ACT1	Amorphic	AR	–	–	–	–	+	++	–	N
RORC	Amorphic	AR	++	–	–	–	–	+	–	N

HSE, herpes simplex encephalitis; EDA, ectodermal anhidrotic dysplasia; AR, autosomal recessive; N, normal; AD, autosomal dominant; VZV, varicella-zoster virus; HSV-1, herpes simplex virus type 1; XR, X-recessive.

[a] one patient with AD IFN-γR1 deficiency presented one episode of *Histoplasma capsulatum* infection, and another patient presented coccidioidomycosis.

[b] one patient with IL-12Rβ1 deficiency presented one episode of *Paracoccidioides brasiliensis* infection, and some patients display forms of chronic mucocutaneous candidiasis.

expression or the expression on the cell surface of nonfunctional receptors (OMIM 147569). All known patients have suffered from disseminated infections caused by EM and/or BCG, with impaired granuloma formation, all requiring continuous multidrug therapy (Fig. 36.2A).[1,25] Infections usually began early in life, often before the age of 3 years. The immunological features of patients with IFN-γR2 deficiency are essentially indistinguishable from those of patients with AR complete IFN-γR1 deficiency. A few of these patients have presented with nontyphoidal salmonellosis, and one presented with recurrent invasive infections with *Listeria monocytogenes.*[25] Viral infections caused by cytomegalovirus (CMV), human herpes virus 8 (HHV-8), respiratory syncytial virus (RSV), pseudorabies virus 3 (PRV-3), and varicella-zoster virus (VZV) have been described in some patients. One patient had a fatal HHV-8–driven Kaposi sarcoma, a second had a pineal germinoma, and a third had a cutaneous squamous cell carcinoma.[1] Multiple antibiotics against mycobacteria should be administered without interruption. Vaccination with live BCG is contraindicated. The prognosis is very poor in the absence of hematopoietic stem cell transplantation (HSCT).[25] HSCT is best carried out once mycobacterial disease has been controlled.[26] HSCT is the only curative treatment for these patients but has proven to be associated with an unusually high rate of graft rejection, making transplantation particularly difficult.[1]

Autosomal Recessive Partial IFN-γR1 and IFN-γR2 Deficiencies

Fifteen patients from 12 unrelated kindred with AR partial IFN-γR1 deficiency (OMIM 107470) have been identified.[27] Two mutations affecting the extracellular domain of IFN-γR1 have been identified: V63G and I87T. Founder effects accounted for the recurrence of both mutations. Patients with AR partial IFN-γR1 deficiency have a less severe clinical phenotype than patients with AR complete IFN-γR1 deficiency. Disseminated BCG was observed in five patients. One patient, who had not been vaccinated with BCG, had curable symptomatic primary tuberculosis (TB). *Mycobacterium abscessus* caused a disseminated infection in one patient. Other nonmycobacterial infections caused by bacteria (*Salmonella, Shigella sonnei, Haemophilus influenzae, Legionella* spp., *Mycoplasma pneumoniae,* and *Klebsiella* spp.), viruses (VZV, RSV, *Molluscum contagiosum*), and parasites (*Toxoplasma gondii, Cryptosporidium*) have been reported. The

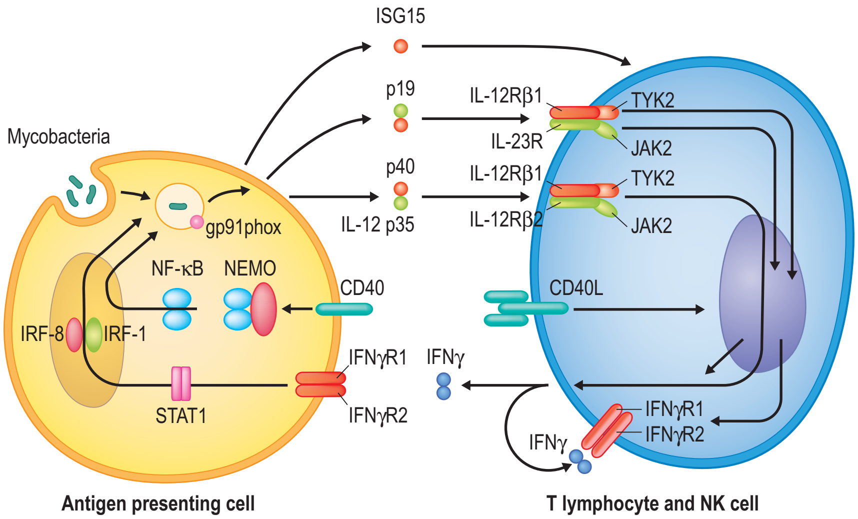


FIG 36.1 Genetic Etiologies of Mendelian Susceptibility to Mycobacterial Disease (MSMD). IL-12 is secreted principally by phagocytes and dendritic cells (DCs) and binds to a receptor consisting of a heterodimer of β_1 and β_2 chains that is expressed specifically on natural killer (NK) and T lymphocytes. IFN-γ is secreted by NK and T lymphocytes and binds to a ubiquitous receptor consisting of two chains, a ligand-binding (IFN-γR1) and a signaling-associated chain (IFN-γR2). Signal transducer and activator of transcription 1 is phosphorylated in response to IFN-γ and is translocated to the nucleus as a homodimer. Nuclear factor (NF)-κB essential modulator (NEMO) is a regulatory protein in the NF-κB pathway. The engagement of CD40 signaling via the NF-κB pathway is important for IL-12 production and for protective immunity against mycobacterial infection. Interferon regulatory factor-8 (IRF8) is an IFN-γ–inducible transcription factor required for the induction of various target genes, including IL-12. TYK2 is necessary for the response to IL-12 in T lymphocytes and NK cells. Gp91^{phox} is a key component of the phagocyte nicotinamide adenine dinucleotide phosphate (NADPH) oxidase.

affected patients had well circumscribed and differentiated tuberculoid granulomas (Fig. 36.2B). The cellular phenotype is characterized by an impaired response to high concentrations of IFN-γ, with receptor expression on the cell surface. Five mutations in six patients with PR IFN-γR2 deficiency (OMIM 147569) have also been reported.[1] Patients have had BCG, *M. abscessus, M. bovis, M. elephantis, M. fortuitum,* and *M. simiae* infections. The patients had high plasma IFN-γ concentrations, and various cell types displayed an impaired response to IFN-γ. All hypomorphic mutations of *IFNGR2* lead to abnormally high levels of N-glycosylation of the corresponding protein and its retention in the endoplasmic reticulum (ER).[28] Thus a diagnosis of PR IFN-γR1 or IFN-γR2 deficiency should be considered in children and adults with mycobacterial infections and a mild clinical and histological phenotype.[1,28] For these patients, antimycobacterial treatment can eventually be stopped, but not within the first year after the infection is brought under control. Recombinant IFN-γ is indicated for the treatment of PR IFN-γR deficiencies. Patients should then be closely monitored. HSCT is not indicated, given the relatively mild infectious phenotype.[1]

Autosomal Dominant Partial IFN-γR1 and IFN-γR2 Deficiencies

More than 70 patients were found to have autosomal dominant (AD) partial IFN-γR1 deficiency (OMIM 107470).[1] These patients have a heterozygous small frameshift deletion in *IFNGR1,* most bearing the same 818del4 heterozygous mutation (or another mutation in the same region), downstream from the segment encoding the transmembrane domain, but upstream from the recycling motif and JAK1 and STAT1 docking sites. This heterozygous mutation defines the first hotspot for small deletions. It is dominant negative and decreases cellular responses to IFN-γ in heterozygous cells and patients. The clinical phenotype is characterized by EM and BCG infections, often affecting the bones.[1,25] A diagnosis of mycobacterial osteomyelitis is highly suggestive of AD IFN-γR1 deficiency.[25] Only 5% of the patients in the largest cohort of patients with heterozygous 818del4 ever published have presented with nontyphoidal salmonellosis, one has presented a single episode of infection with *Histoplasma capsulatum,* eight are asymptomatic, and two died of disseminated mycobacterial infection at the ages of 17 and 27 years, respectively. Antimycobacterial antibiotic treatment could be combined with IFN-γ injection in cases of disseminated infection. Antibiotics can be stopped, but not within the first year after the infection is brought under control. Vaccination with live BCG is contraindicated. The clinical outcome of patients with AD IFN-γR1 deficiency, like that of patients with AR partial IFN-γR1 and IFN-γR2 deficiencies, is much better than that of children with AR complete IFN-γR deficiency because there is some residual IFN-γ signaling. There is, therefore, a strict correlation between *IFNGR1* genotype and cellular and clinical phenotype.[25]

One heterozygous mutation of *IFNGR2* has been reported to contribute to an AD form of partial IFN-γR2 deficiency.[29] This mutation, 186delC, creates a premature codon stop upstream from the segment encoding the transmembrane domain. The affected patient had a mild form of mycobacterial infectious disease, and her father was asymptomatic. Haploinsufficiency of the *IFNGR2* locus was observed in Epstein-Barr virus (EBV) B cells and T lymphocytes.[29]

Complete IL-12Rβ1 and IL-12P40 Deficiencies

More than 200 patients with AR complete IL-12 receptor β1 (IL-12Rβ1) deficiency (OMIN 601604), with no cellular expression or with expression of the mutated form, have been identified.[1,30] The cellular phenotype of patients with IL-12Rβ1 deficiency is a lack of response of natural killer (NK) and T lymphocytes to IL-12 and IL-23, resulting in low levels of IFN-γ production. The clinical phenotype is characterized by EM/BCG infections, with half the patients also presenting nontyphoidal salmonellosis.[30] Most of the children inoculated with live BCG vaccine

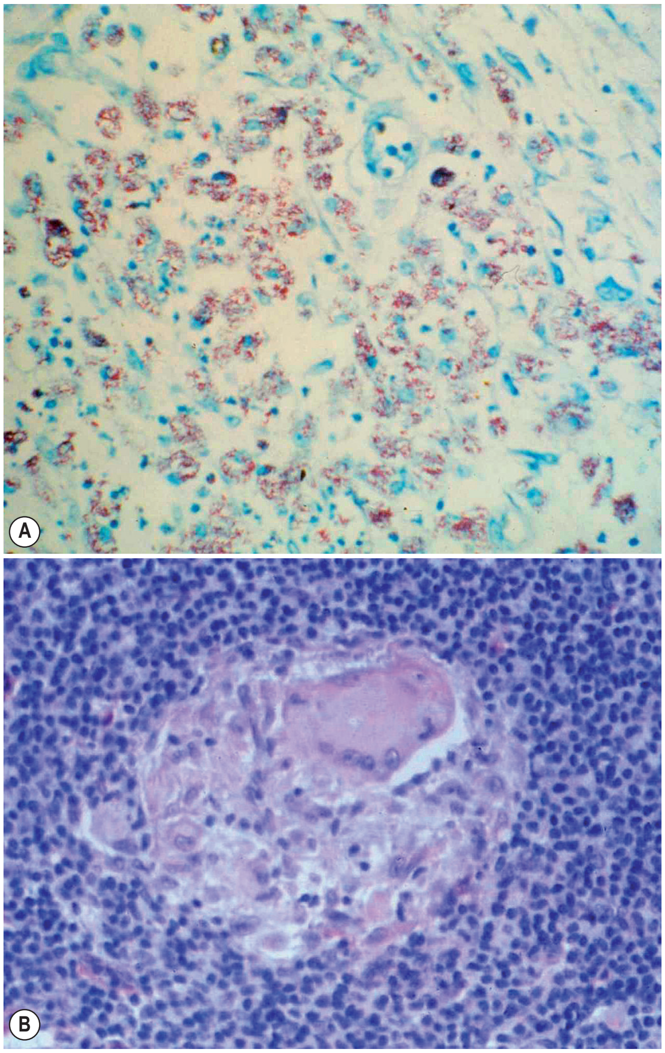


FIG 36.2 **Two Types of Granuloma.** (A) The lepromatous-like type consists of poorly defined, poorly differentiated granulomas, with few, if any giant cells and lymphocytes, but widespread macrophages loaded with acid-fast bacilli. (B) The tuberculoid type consists of well circumscribed and differentiated granulomas, with epithelioid and multinucleated giant cells containing very few acid-fast rods, surrounded by lymphocytes and fibrosis, occasionally with central caseous necrosis.

developed clinical infection, and other mycobacterial infections are caused by EM (alone or associated with other infections) or by *M. tuberculosis.* The low penetrance for the case-definition phenotype of BCG/EM disease led to the discovery of tuberculosis as the sole infectious phenotype in several patients, providing the first cases of mendelian predisposition to tuberculosis. Infections caused by *Paracoccidioides brasiliensis, Coccidioides* spp., *Histoplasma* spp., *Cryptococcus neoformans, Toxoplasma* spp., and *Leishmania* spp. have been reported in a few cases. A substantial fraction of patients with IL-12Rβ1 deficiency present mild forms of CMC,[30-32] an observation probably related to the impairment of the IL-23–IL-17 circuit in these patients. All the survivors were well in 2010, when the oldest of these patients was 51 years old.[1,30]

Fifty patients with AR complete IL-12p40 deficiency (OMIM 161561) have been reported.[1,33] Children with this deficiency produce abnormally low levels of IFN-γ because of a lack of stimulation through the IL-12R. This defect can be partially corrected, in a dose-dependent manner, with exogenous recombinant IFN-γ.[1,33] All children inoculated with live BCG vaccine developed clinical infection, and infections caused by EM or *M. tuberculosis* have also been reported. Other infections, such as salmonellosis, klebsiellosis, and nocardiosis, have been described. Five of the reported patients are asymptomatic. Thirteen of the 50 patients died before the age of 8 years, and one patient died at the age of 34 years.[1,33]

For IL-12p40 and IL-12Rβ1 deficiencies, antibiotic treatment can be combined with IFN-γ injection in cases of disseminated infection.[1] Antibiotics can be stopped, but not within the first year after the infection is brought under control. Infectious *Salmonella* disease should be treated with antibiotics or with IFN-γ and antibiotics in patients with recurrent salmonellosis. Thus a diagnosis of IL-12p40 or IL-12Rβ1 deficiency should be considered in children and adults with mycobacterial and *Salmonella* infections or mycobacterial infections and CMC with a mild clinical and histological phenotype. The vaccination of these patients with live BCG is contraindicated.[1,33] The prognosis is, nevertheless, good, partly because of the incomplete clinical penetrance of primary mycobacterial infection, the favorable response of infections to treatment, and the rarity of recurrent or multiple mycobacterial infections. Indeed, children with primary mycobacterial disease can mount a fully protective immune response against a secondary mycobacterial disease.

THERAPEUTIC PRINCIPLES

Treatment of Mendelian Susceptibility to Mycobacterial Disease (MSMD) Patients

- Vaccination with live Bacille Calmette-Guérin (BCG) is contraindicated.
- Multiple antibiotics against mycobacteria should be administered without interruption in patients with complete interferon (IFN)-γR1 or IFN-γR2 deficiency.
- Antimycobacterial antibiotics may be associated with IFN-γ injections in selected patients with partial IFN-γR1, IFN-γR2, IRF8, CYBB, or signal transducer and activator of transcription 1 (STAT1) deficiency, complete interleukin (IL)-12p40, IL-12Rβ1, or ISG15 deficiency.
- Hematopoietic stem cell transplantation (HSCT) should be considered in selected patients with complete IFN-γR1 or IFN-γR2 deficiency.

Autosomal Dominant Partial STAT1 Deficiency

Nine kindreds with 21 genetically affected cases have been reported to have heterozygous mutations in *STAT1* causing AD partial STAT1 deficiency (OMIM 600555).[1,24] STAT1 is a critical transducer of IFN-mediated signals, either as STAT1 homodimers (γ-activating factor [GAF]) or as STAT1/ STAT2/IRF9 trimers (IFN-stimulated gene factor 3 or ISGF3) (Fig. 36.3). These heterozygous *STAT1* mutations reduce the cellular response to IFN-γ, but not that to IFN-α.[24] Mycobacterial infections caused by BCG, EM, or TB are observed, ranging from local to disseminated disease. Osteomyelitis is frequent in these patients (eight of 15 patients).[1] The other six patients are asymptomatic. All patients were well at their last evaluations and were between four and 38 years old in 2005. Antimycobacterial treatment can be stopped, but not during the first year after the infection is brought under control. Vaccination with live BCG is contraindicated. Observations of affected patients have suggested that STAT1 and GAF are required for human IFN-γ-mediated mycobacterial immunity. In conclusion, patients with AD STAT1 deficiency have clinical and cellular phenotypes (*i.e.,* susceptibility

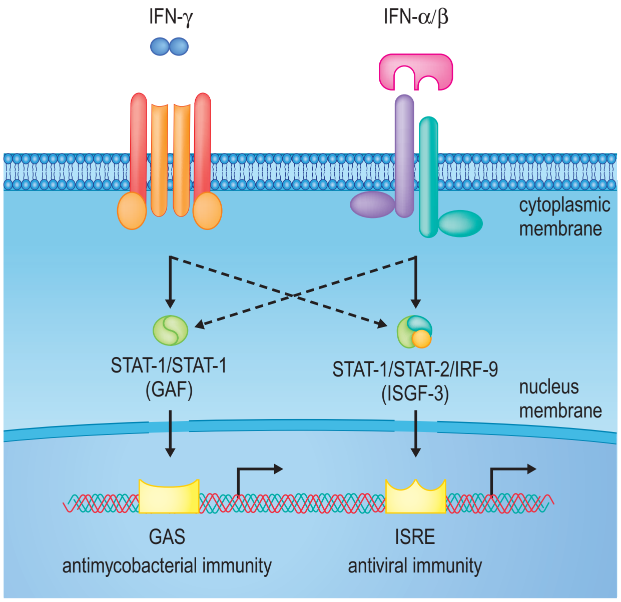


FIG 36.3 STAT-1 Pathway. The binding of homodimeric IFN-γ to its tetrameric receptor leads to activation of the constitutively associated JAK kinases JAK1 and JAK2, which then phosphorylate tyrosine residues in the intracellular part of interferon (IFN)-γR1. Upon IFN-γ stimulation, unphosphorylated signal transducer and activator of transcription 1 (STAT1) molecules are directly recruited to IFN-γR1 docking sites. They are then phosphorylated and released into the cytosol as phosphorylated STAT1 homodimers. These homodimers form γ-activating factor (GAF), which is translocated to the nucleus. GAF binds γ-activating sequences (GAS) present in the promoters of target genes. Following monomeric IFN-α/β stimulation, STAT2 is recruited to the phosphorylated IFN-αR1 chain of the heterodimeric IFN-αR, which is in turn phosphorylated by JAK1 and TYK2. This leads to the phosphorylated STAT2-mediated recruitment of STAT1, which is then phosphorylated. Active phosphorylated STAT1/STAT2 heterodimers are released into the cytosol and translocated to the nucleus with interferon regulatory factor-9 (IRF-9), to form interferon-stimulated gene factor-3 (ISGF-3) heterotrimers. ISGF-3 binds IFN-α/β sequence response elements (ISRE) in the promoters of target genes via the DNA-binding domains of STAT1 and IRF9. In humans, recessive complete STAT1 deficiency results in impaired responses to both IFN-γ and -α/β. It is associated with a specific syndrome, different from mendelian susceptibility to mycobacterial disease (MSMD), of susceptibility to both mycobacteria (impaired IFN-γ-mediated immunity) and viruses (impaired IFN-α/β-mediated immunity).

to mycobacterial disease, osteomyelitis in particular, and impaired GAF activation) similar to those of patients with AD IFN-γR1 deficiency. These patients should be treated in a similar manner.[1]

Complete and Partial IRF8 Deficiency

Two types of IRF8 deficiency have been reported in the last few years (OMIM 601565).[23] Disseminated BCG disease in a child born to consanguineous parents led to the discovery of a lack of circulating monocytes and DCs. The severity of the disease, mimicking CID, rendered HSCT necessary.[1,23] Candidate gene studies showed that the child had AR complete IRF8 deficiency. The IFN-γ circuit was, therefore, profoundly disrupted. Other mechanisms may account for BCG disease and vulnerability to other infectious agents. In two other patients with a history of pure MSMD, the same heterozygous mutation, T80A, was found in *IRF8*, defining an AD IRF8 deficiency.[23] The mutation was severely hypomorphic and dominant negative. Interestingly, the IFN-γ circuit did not appear to be disrupted in tests on whole blood. Nevertheless, the two unrelated patients with AD IRF8 deficiency lacked circulating CD1c CD11c DCs.[23] This population of cells produces large amounts of IL-12 in normal individuals, and its absence therefore probably contributed to the MSMD phenotype.

AR Complete ISG15 Deficiency

Recently, six patients with complete AR ISG15 deficiency have been identified (OMIM 616126).[34,35] Three of the six patients suffered from clinical disease caused by BCG vaccine. The only features common to all patients were calcifications of the basal ganglia in the brain. The cellular phenotype was characterized by an impaired, but not abolished, IFN-γ production upon whole blood activation by BCG plus IL-12, as seen in patients with IL-12p40 or IL-12Rβ1 deficiency.[1,34,35] The addition of recombinant ISG15 rescued the production of IFN-γ by T and NK cells of the patients. A high level of ISG mRNA was detected in all patients with ISG15 deficiency. The absence of intracellular ISG15 in the patients' cells prevents the accumulation of USP18, a potent regulator of IFN-α/β signaling, and amplifies the IFN-α/β induced responses. This novel MSMD-causing gene acts like both an IFN-γ–inducing secreted cytokine and an intracellular negative regulator of IFN-α/β.[34,35] These patients display not only MSMD but also a novel form of type I interferonopathy because intracellular human (but not mouse) ISG15 is paradoxically a negative regulator of IFN immunity.

AR Complete TYK2 Deficiency

AR complete TYK2 deficiency was first described in a Japanese patient as a genetic etiology of hyper-immunoglobulin E (IgE) syndrome in 2006 (OMIM 176941). However, five more families and seven more patients have since been described.[2] The patients harbor homozygous nonsense mutations, deletions, or insertions generating a premature stop codon and resulting in undetectable levels of the TYK2 protein.[2,36] Clinically, all the patients display mycobacterial and/or viral diseases, which appear to be the main clinical features of AR complete TYK2 deficiency. One patient suffered only from BCGosis, two only from tuberculosis, one only from HSV meningitis, three had mycobacterial and viral diseases, and one had meningitis of unknown origin, although all were vaccinated with BCG and had probably encountered the common viruses of childhood.[2] The patients' cells display an impaired, but not abolished response to IL-12/IL-23 and IFN-α/β, almost certainly accounting for their susceptibility to mycobacteria and viruses, respectively. The response to IL-10 is weak, but with no overt clinical manifestations. An impaired response to IL-23 is also observed, preventing the differentiation of abolished CD4 T cell response into Th17 cells *in vitro*, but the patients have normal levels of circulating IL-17$^+$ T cells, probably accounting for their lack of CMC.[2] An intact response to IL-6 has been observed in all patients except for the first patient described, who displayed the symptoms of hyper-IgE syndrome. The impaired response to IL-6 in this patient was found to be independent of WT-TYK2 expression, potentially accounting for some of the unusual clinical features in this patient. Overall, AR complete TYK2 deficiency leads to isolated or combined mycobacterial and viral diseases.[2]

NEMO and *CYBB* Deficiencies

An X-linked recessive (XR) form of MSMD (XR-MSMD) (OMIM 300248) was discovered in three unrelated kindreds. This condition was caused by two neighboring mutations (E315A and R319Q) in *NEMO*.[1,37] These two missense mutations, which affect two neighboring residues in the leucine zipper domain that normally form a salt bridge, were entirely responsible for impairment of the CD40-triggered induction of IL-12 production by monocyte-derived cells following stimulation with CD40L-expressing T cells. Four maternally related male patients from two successive generations presented *M. avium* infection only, and the other two, unrelated patients presented only mycobacterial disease. These patients had a purely mycobacterial infectious phenotype, with no sign of anhidrotic ectodermal dysplasia (EDA).[1,37]

A second form of XR MSMD has been reported (OMIM 300645)[22] in seven patients from two unrelated French families. The *CYBB* or gp91^{phox} mutations identified in these patients were previously unknown and affect the transmembrane domain and an extracellular loop. *CYBB* encodes the major component of the nicotinamide adenine dinucleotide phosphate (NADPH) oxidase complex.[22] Mutations in any of the genes encoding molecules from this complex generally result in chronic granulomatous disease (CGD). However, unlike the cells of patients with CGD, neutrophils, monocytes, and monocyte-derived dendritic cells (MDDCs) from these seven patients with XR MSMD displayed perfectly normal levels of NADPH oxidase activity.[22] These patients produce only small amounts of neutrophil and monocyte gp91^{phox} protein, but normal amounts of cytochrome b_{558}. However, monocyte-derived macrophages and EBV-B cells display much more profound defects of gp91 ^{phox} expression, resulting in impaired cytochrome b_{558} assembly and low levels of NADPH oxidase activity.[22] Thus NADPH oxidase activity in human macrophages is a crucial effector mechanism for protective immunity to tuberculous mycobacteria. None of the seven patients suffered from any other infectious diseases. These patients are all adults and are currently well. Vaccination with live BCG is contraindicated in patients with these deficiencies. None of the patients described, to date, has had prophylactic treatment. Infections can be treated with multiple antibiotics.[1]

INHERITED DISORDERS OF IFN-γ– AND IFN-α/β–MEDIATED IMMUNITY

AR Complete and Partial STAT1 Deficiency

Nine children from six unrelated kindreds with AR complete STAT1 deficiency (OMIM 600555) have also been identified[24] (unpublished data). These patients carried homozygous mutations that completely abolished cellular responses to both IFN-γ and IFN-α/β. Responses to IL-27 and IL-29 (IFN-λ) were abolished, potentially accounting for the susceptibility of these children to bacteria and viruses, respectively. All patients had disseminated BCG disease. One patient died from recurrent encephalitis caused by HSV, a second died from an undocumented illness thought to be of viral origin, and a third died 3 months after HSCT, from EBV-induced lymphoproliferative disorder. The principal causes of death in children with AR complete STAT1 deficiency are viral diseases or disseminated infection with BCG[38](unpublished data). The outcome of viral illnesses in children with AR complete STAT1 deficiency suggests that the STAT1-dependent response to human IFN-α/β is necessary for viral immunity. The patients with AR complete STAT-1 deficiency had multiple, severe, early-onset viral infections. Multiple antibiotics against mycobacteria and possibly antiviral treatment directed against herpes viruses should be administered without interruption, and HSCT should be considered for these patients. Vaccination with live BCG is contraindicated. Partial, as opposed to complete, AR STAT1 deficiency has also been described in three unrelated kindreds.[24] These patients suffered from mild mycobacterial and viral diseases. Their responses to IFN-γ and IFN-α were impaired, but not abolished, in terms of GAF- and ISGF3-mediated immunity. AR complete and partial STAT1 deficiencies constitute an innate immunodeficiency, ranging from lethal to moderate clinical disease, and a diagnosis of STAT1 deficiency should be considered in infants and children with infectious diseases, particularly (but not exclusively) if mycobacterial and viral in nature.[24]

GENETIC DISORDERS OF THE TLR3–IFN-α, IFN-β, AND IFN-λ PATHWAY

This group of genetic disorders leads to impaired TLR3 signaling and susceptibility to HSE in childhood. The affected patients bear mutations of *TLR3, UNC93B1, TRIF, TRAF3, TBK1,* or *IRF3* (Fig. 36.4). The TLR3 signaling pathway is mediated exclusively by the TRIF adapter and leads to activation of the transcription factors IRF3 and NF-κB (see Fig. 36.4).[39] TRIF recruits TNF receptor–associated factor 6 (TRAF6) and activates TAK1 for NF-κB activation. TRIF also recruits a signaling complex involving TBK1 and IKKε via TRAF3 for IRF3 activation.[39] This signaling pathway induces the production of type I and type III IFNs and inflammatory cytokines and is important in antiviral immunity.[39] UNC-93B is a 12-transmembrane domain protein present in the ER that delivers the nucleotide-sensing receptors TLR3, -7, -8, and -9 from the ER to endolysosomes. TRAF-3, TBK1, and IRF3 have functions downstream from multiple tumor necrosis factor (TNF) receptors and the receptors inducing IFN-α, -β, and -λ production, including TLR3.

TLR3 Deficiency

AD and AR forms of TLR3 deficiency caused by heterozygous, compound heterozygous, or homozygous mutations of *TLR3* were first identified in 2007 (OMIM 613002).[4,8] Six patients from six unrelated kindreds have been reported, and each of these patients developed childhood HSE as a result of AR or AD, complete or partial TLR3 deficiency. There is a high degree of allelic heterogeneity, with three AD partial defects resulting from negative dominance or haploinsufficiency, and two AR defects resulting from complete or partial deficiency. Some subsets of blood leukocytes (MDDCs, NK cells, and CD8 T cells) from patients with TLR3 deficiency have an impaired response to stimulation with TLR3 agonist.[8,40] The fibroblasts of the patients produce low levels of the antiviral molecules IFN-β/-λ in response to TLR3 agonist and HSV-1, leading to higher levels of viral replication and virus-induced cell death than for healthy control cells. Only one child from each of the kindreds identified developed HSE. The patients with TLR3 deficiency developed HSE upon primary infection with HSV-1 during childhood. Four (66%) of the six patients with TLR3 deficiency had at least one late relapse of HSE, whereas relapse generally occurs in only about 10% of HSE patients.[8] All patients with HSC who had TLR3 deficiency were exposed to other viruses during many years of follow-up, as shown by the positive serological results obtained for other herpes viruses, with no subsequent acute events.[8] None of the other individuals from the six kindreds with dominant or homozygous hypomorphic *TLR3* mutations developed HSE despite serologically documented HSV-1 infection.

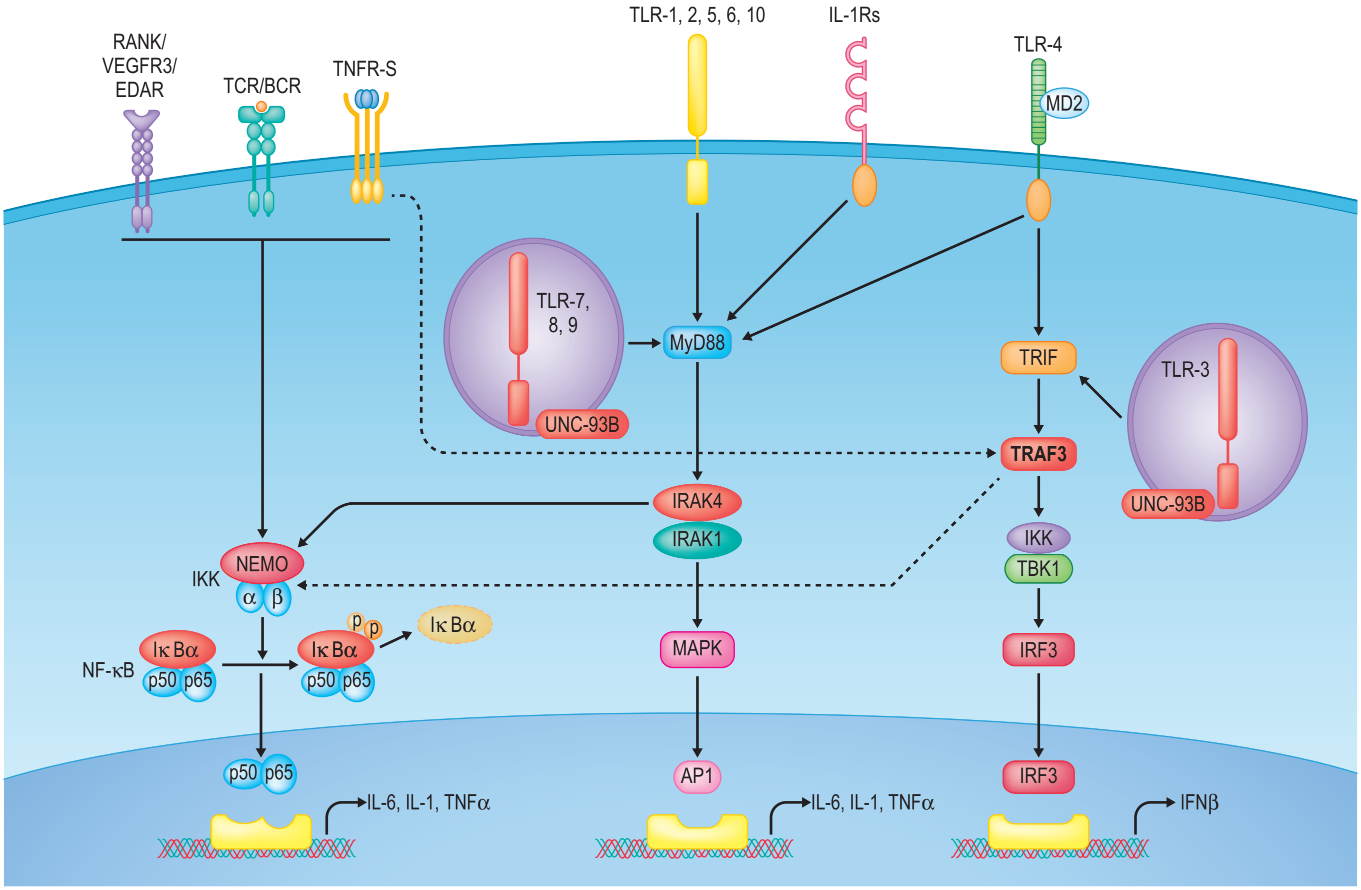


FIG 36.4 **Human Defects Involving the Interferon (IFN)-α/-β, IFN-λ and Nuclear Factor (NF)-κB Pathways.** Immune receptor signaling pathways leading to NF-κB activation can be grouped into four categories on the basis of the surface receptors involved: the developmental receptors RANK, VEGFR3, and EDAR; antigen receptors (T-cell receptor [TCR] and B-cell receptor [BCR]); members of the tumor necrosis factor (TNF) receptor superfamily (TNF-Rs) and members of the TIR superfamily (IL-1Rs/TLRs). The two proteins of the Toll/IL-1R (TIR) signaling pathway (MyD88, IRAK-4) and the two proteins of the NF-κB signaling pathway (NEMO and IκBα) responsible for primary immune deficiencies (PIDs) are shown in red. UNC-93B defects abolish cellular responses to TLR3 and 7–9 agonists. TNF receptor–associated factor 3 (TRAF3) defects impair the cellular responses to TLR3 and other pathways. The three proteins of the TLR3 signaling pathway (TLR3, UNC-93B, and TRAF3) responsible for PIDs are shown in red.

Childhood-onset HSE is, therefore, caused by TLR3 deficiency in an identifiable fraction of patients, particularly those displaying HSE recurrence. Treatment with recombinant IFN-α, in parallel with acyclovir, may help improve the outcome in patients with TLR3 deficiency and HSE. TLR3 deficiency displayed complete penetrance at the cellular level but incomplete penetrance at the clinical level. Multiple factors may affect clinical penetrance, including age at infection with HSV-1, the viral inoculum, and human modifier genes.[8]

UNC-93B Deficiency

AR complete UNC-93B deficiency was identified, in 2006, as the first genetic etiology of isolated HSE (OMIM 610551).[3] Three individuals from two consanguineous kindreds were found to carry homozygous mutations of *UNC93B1*. UNC-93B-deficient fibroblasts and leukocytes did not respond to TLR3, or to TLR7-9 agonists, respectively, in terms of IFN-α, -β, and -λ production. Fibroblasts from patients with UNC-93B deficiency displayed abnormally high levels of viral replication and cell death after infection with HSV-1.[3] Two of the three individuals with UNC-93B deficiency developed HSE. The first patient presented recurrent episodes of HSE at the ages of 11 months, 14 months, and 3½ years. During 14 years of follow-up, this patient has experienced no subsequent acute events. The second patient presented recurrent episodes of HSE at the ages of 5 and 17 years. She is now 21 years old, and her clinical status has not worsened since the second episode of HSE. One sibling of the second patient, who carries the same homozygous mutation in *UNC93B1,* is now 30 years old and has not developed HSE, despite serologically documented HSV-1 infection. The two patients with UNC-93B deficiency and HSE have been exposed to other viruses with no evident clinical manifestations.[3] Treatment with recombinant IFN-α may also be considered, in parallel with acyclovir, in patients with UNC-93B deficiency and HSE.

TRIF Deficiency

The subsequent discovery of AR and AD TRIF deficiencies in children with HSE completely validated the essential role of human TLR3 immunity in anti-HSV-1 defense in the central nervous system (CNS).[7] A patient with AR TRIF deficiency, born to consanguineous Saudi parents, suffered from HSE at the age of 2 years. A homozygous nonsense mutation, R141X, was identified as the cause of her AR TRIF deficiency.[7] Another patient, with AD TRIF deficiency, of mixed European descent (French, Portuguese, and Swiss), suffered from HSE at the age of 21 months. A heterozygous missense mutation, S186L, led to AD TRIF deficiency in this girl.[7]. The two patients with TRIF deficiency have displayed no other unusually severe infectious diseases during follow-up to date. The TLR3 signaling pathway was found to be impaired in cells from patients with both AR and AD TRIF deficiencies, whereas the TRIF-dependent TLR4 pathway was affected only in patients with AR TRIF deficiency. Interestingly, abnormally weak responses to stimulation of the DExD/H-box helicase pathway have been observed in both AR and AD TRIF-deficient fibroblasts. These results demonstrate the importance of TRIF for the TLR3-dependent production of antiviral IFNs in the CNS during primary infection with HSV-1.[7] However, they also suggest that the TRIF-dependent TLR4 and, probably, the DExD/H-box helicase pathways, are largely redundant in host defense in humans.

TRAF3 Deficiency

A French patient with AD TRAF3 deficiency and HSE has recently been identified.[5] The *de novo* germline heterozygous missense mutation in this patient is loss-of-expression, loss-of-function, and dominant-negative. Fibroblasts from the patient displayed impaired responses to TLR3 agonist stimulation, in terms of IFN-β and -λ production. Various TRAF3-dependent pathways were impaired in the patient's cells, including the IFN-α/β-, and IFN-λ–inducing and TNF-R–responsive pathways. However, there was sufficient residual TRAF3-dependent signaling for most defects to remain clinically silent. By contrast, the impaired TLR3 response was symptomatic and caused HSE, implying that the TLR3 pathway is critically dependent on TRAF3 and essential for immunity to HSV-1 in the CNS. The first clinical signs of HSE in the patient with TRAF3 deficiency appeared at the age of 4 years. The patient with TRAF3 deficiency and HSE described here is now 18 years old and has otherwise remained healthy with no prophylaxis.[5] She shows normal resistance to other infectious diseases, including viral diseases, in particular.[5] We cannot yet exclude the possibility that other forms of human TRAF3 deficiency are involved in other human diseases, including viral diseases, but TRAF3 deficiency should be sought in other children with HSE.

TBK1 Deficiency

The identification of AD TBK1 deficiency as a new genetic etiology of isolated childhood HSE was probably the greatest surprise to come out of human studies to date.[6] TBK1 is a kinase at the crossroads of multiple IFN-inducing signaling pathways, including those mediated by TLR3, cytosolic double-stranded RNA (dsRNA) sensors, and dsDNA sensors. Two unrelated children with HSE were found to carry different heterozygous missense mutations in the gene encoding TBK1. Both mutations underlie a phenotype that is dominant, but presumably by different mechanisms: negative dominance (G159A) and haplotype insufficiency (D50A). A defect in poly(I:C)–induced TLR3 responses was detected only in fibroblasts from the patient carrying the dominant-negative *TBK1* mutation.[6] Nevertheless, levels of viral replication and cell death caused by two TLR3-dependent viruses (HSV-1 and VSV) were high in fibroblasts from both patients. However, the peripheral blood mononuclear cells (PBMCs) and fibroblasts from both patients displayed normal IFN responses to the TLR3-independent agonists and viruses tested.[6] AD TBK1 deficiency, therefore, confers narrow, partial cellular phenotypes, probably accounting for the narrow clinical phenotype of these patients, which is limited to HSE. The two patients with TBK1 deficiency suffered from HSE at the ages of 7 years and 11 months, respectively. During subsequent follow-up over a period of years, they have suffered no other unusual infectious diseases, including viral diseases, in the absence of prophylaxis. As in other HSE patients with genetic deficiencies of the TLR3 pathway, treatment with recombinant IFN-α, in parallel with acyclovir, may be considered in patients with TBK1 deficiency and HSE.

IRF3 Deficiency

IRF3 is a transcription factor that controls multiple IFN-inducing pathways, including the TLR3 pathway, which can be triggered by dsRNA, and the pathways triggered by other RNA and DNA sensors. IRF3 is normally activated by TBK1 and/or IKKε kinases. A missense IRF3 mutation has been identified in a 15-year-old patient with HSE.[9] The resulting mutant IRF3 protein could not activate transcription and was dominant through haploinsufficiency. In heterozygous cells, haploinsufficiency for IRF3 impairs the induction of antiviral IFN-α/β in leukocytes following stimulation with various agonists. Moreover, the induction of IFN-β in fibroblasts stimulated with poly(I:C), which mimics viral dsRNA and activates TLR3, was also impaired. Finally, IFN-β induction by HSV1 was impaired in fibroblasts. Overall, consistent with the key role of IRF3 in inducing IFN-α/β, a profound, broad cellular phenotype was observed in patients with AD IRF3 deficiency. Despite the broad impact of AD IRF3 deficiency, in terms of both the pathways and cell types affected, its clinical phenotype in the single patient identified to date is surprisingly narrow, restricted to HSE.[9] Moreover, clinical penetrance is incomplete, as a heterozygous relative of the patient is healthy. Overall, the discovery of IRF3 deficiency in a patient with HSE suggests that IRF3 is a key component in TLR3-dependent, IFN-mediated, CNS-intrinsic immunity to HSV-1. It also adds weight to the emerging notion that impaired TLR3-IFN immunity can underlie HSE of childhood, suggesting that IFN treatment may be beneficial to patients.[9]

IRF7 Deficiency

AR complete IRF7 deficiency was identified, in 2015, as the first genetic etiology of severe isolated influenza (OMIM 616345).[10] A 30-month-old girl born to nonconsanguineous French parents was admitted to intensive care for 3 weeks and required respiratory support for 10 days because of pandemic H1N1 (pH1N1) 2009 influenza A virus (IAV)–induced acute respiratory distress syndrome. This patient carries compound heterozygous mutations, p.F410V and p.Q421X, in *IRF7,* encoding IRF7.[10] Each of her parents has a single heterozygous mutation, and neither has any significant infectious disease history. IRF7 is required for the production and amplification of the antiviral type I (IFN-α/β) and III (IFN-λ) IFNs. The patient's fibroblasts displayed impaired IFN-β and IFN-λ1 production in response to various stimuli, including IAV. Consequently, IAV replication resulted in titers 100 times higher in the patient's fibroblasts than in fibroblasts from a healthy donor. Furthermore, type I IFN production was abolished in the patient's plasmacytoid DCs, a cell type that normally produces constitutively high levels of IRF7 and, therefore, large amounts of type I IFN. Moreover, pulmonary epithelial cells differentiated *in vitro* from patient-specific induced pluripotent stem cells displayed uncontrolled virus replication and impaired IFN-β production. The patient had no other serious viral infections despite being seropositive for a myriad respiratory viruses and is currently healthy at the age of 8 years.[10] Thus the IRF7-dependent IFN amplification pathway is crucial for the mounting of antiviral responses to influenza virus but may be redundant for other viruses.

GENETIC DISORDERS OF NF-κB–MEDIATED IMMUNITY

This group of inherited disorders leads to impaired NF-κB signaling and strong susceptibility to pyogenic bacteria. Affected patients bear mutations of *NEMO, NFKBIA, IRAK4, MYD88, HOIL1,* or *HOIP* (see Fig. 36.4).[16,41,42] Patients with ectodermal dysplasia, anhidrotic (EDA) and immunodeficiency (EDA-ID) syndrome carry either XR hypomorphic mutations of *NEMO,* AD hypermorphic mutations of *NFKBIA*, or AR amorphic mutations of *IKKB.*[43] Diverse mutations have been found in *NEMO,* associated with various cellular and clinical phenotypes (see above and below).[44] Patients with AR amorphic mutations of *IRAK4* or *MyD88* present a more restricted, purely immunological defect, with specific impairment of the TIR-IL receptor–associated kinase (IRAK) signaling pathway.[13] Patients with NEMO, IκBα, IKKB, HOIL1, or HOIP deficiencies are susceptible to multiple infectious agents, including pyogenic bacteria, mycobacteria, and viruses. By contrast, patients with IRAK-4 or MyD88 deficiencies seem to be specifically prone to pyogenic bacterial diseases.

NEMO Deficiency

XR-EDA-ID is a rare PID associated with a developmental disorder (OMIM 300291). Patients with XR-EDA-ID carry hypomorphic mutations of *IKBKG,* which encodes IKK-γ/NEMO, a protein essential for activation of the ubiquitous transcription factor NF-κB. Up to 100 patients with NEMO deficiency have been reported.[16,43] The only known consistent immunological abnormality in patients with NEMO deficiency is an absence of serum antibodies against carbohydrates, which has been found in most of the patients described to date. Some patients have hyper-IgM syndrome, and a few have NK-cell abnormalities. Most patients with NEMO deficiency fail to produce IL-10 in response to activation with TNF-α in whole-blood assays.[16]

The infectious phenotype of the other NEMO-deficient patients is characterized mostly by infections caused by encapsulated pyogenic bacteria, such as *H. influenzae* and *Streptococcus pneumoniae.* Infections caused by weakly pathogenic microorganisms, such as *M. avium* and *M. kansasii,* have also been diagnosed in patients.[16] Other infectious diseases, including salmonellosis, *Pneumocystis* pneumonia, and viral illnesses caused by HSV and CMV, have been reported. Infectious episodes were marked by a poor clinical and biological inflammatory response.[44] One-third of the patients with NEMO deficiency died from invasive infection, demonstrating the severity of this disorder.[44] About 80% of the patients with NEMO deficiency described to date have had EDA-ID, which is characterized by hypohidrosis, widely spaced cone-shaped or peg-shaped teeth, and hypotrichosis[16] (Fig. 36.5). These features result from defective signaling via the ectodysplasin receptor signaling pathway. In some other patients with NEMO deficiency, osteopetrosis (O) and lymphedema (L) have been found associated with the EDA phenotype.[16] Some patients also have dysmorphia with mild frontal bossing. However, about 10% of patients with NEMO deficiency do not display any of the classic features of the EDA phenotype.[16]

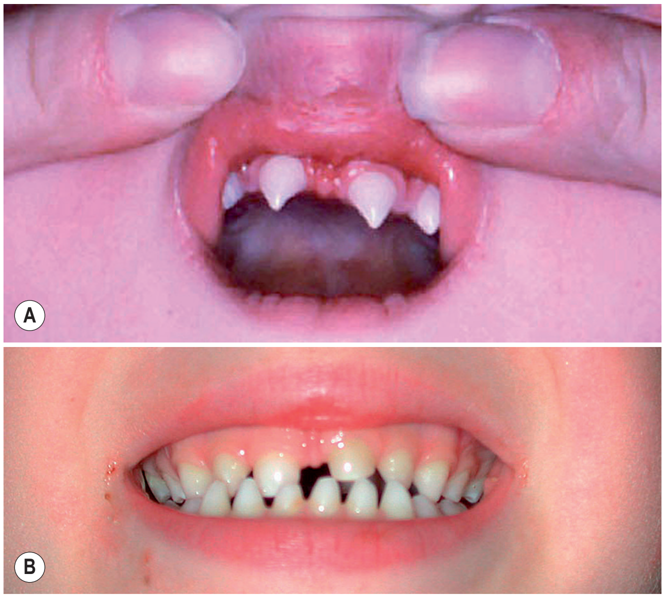


FIG 36.5 Patients With Anhidrotic Ectodermal Dysplasia and Immunodeficiency (EDA-ID). Two patients with EDA-ID, one with widely spaced cone- or peg-shaped teeth (A), the other having conical incisors (B).

Prophylactic trimethoprim–sulfamethoxazole and/or penicillin V treatment should be considered and IgG substitution should be carried out in patients with NEMO deficiency presenting an impairment of B-cell immunity. Patients with functional B-cell immunity should be immunized with *S. pneumoniae* conjugated and nonconjugated vaccines, *H. influenzae* conjugated vaccine, and *N. meningitidis* conjugated and nonconjugated vaccines. Vaccination with live BCG is contraindicated. It is important to initiate empiric parenteral antibiotic treatment against *S. pneumoniae, Staphylococcus aureus, Pseudomonas aeruginosa,* and *H. influenzae* as soon as infection is suspected or the patient develops a moderate fever, without taking inflammatory parameters into account, as patients may die of rapid invasive bacterial infection despite appropriate prophylaxis. HSCT should be considered for patients with severe infectious and cellular phenotypes[16] (Picard unpublished).

IκBα Deficiency

Ten patients with three different hypermorphic mutations of *NFKBIA,* which encodes an inhibitor of NF-κB (OMIM 164008), have been identified since 2003, one of whom displays complex partial mosaicism.[15,16] IκBα deficiency can lead to an impairment of T-cell receptor (TCR) signaling in 70% of patients, and all have impaired TNF receptor (TNFR) and IL1R/TLR responses. Patients with IκBα deficiency have dysgammaglobulinemia and impaired production of specific antibodies. Some also have low proportions of memory CD4 or CD8 T cells or both, excess naïve T cells, or an absence of TCRγ/δ T cells. With the exception of two patients with IκBα deficiency, one with complex mosaicism and one with S36Y mutation, all patients have been found to have the features of EDA.[16] All 10 patients with IκBα deficiency have developed recurrent bacterial infections: pneumonia in five cases, sepsis or meningitis in three cases, and arthritis in one case. They are also prone to opportunistic infections, six of these patients having had *Pneumocystis* pneumonia and CMC. Finally, some patients presented with recurrent diarrhea and/or colitis. Thus a diagnosis of IκBα deficiency should be considered in children with EDA and CID with impaired T-cell immunity. A preventive treatment, including antibiotic prophylaxis with trimethoprim–sulfamethoxazole and/or penicillin V, should be proposed and IgG substitution should be carried out in patients with IκBα deficiency. The recommendations for febrile patients with NEMO deficiency should also be applied to patients with IκBα deficiency. HSCT has been reported in four patients with severe IκBα deficiency causing CID.[45,46]

IL-1 Receptor–Associated Kinase-4 Deficiency

Inherited IRAK-4 deficiency (OMIM 607676) is an AR disorder first described in 2003.[11] In total, 49 patients have since been identified, from 32 kindreds.[13,16] The blood cells of the patients fail to produce the proinflammatory cytokines IL-1β and IL-18 upon stimulation with all known TLR agonists (with the exception of TLR3 agonists). Patients with IRAK-4 deficiency seem to have normal antigen-specific T- and B-cell responses, as shown in routine immunological investigations, with two notable exceptions.[47] First, the glycan-specific IgG and IgM antibody responses to pneumococcal and AB glycans (allohemagglutinins of the ABO system) are impaired in up to one-third of the cases explored.[13] Second, serum IgE and IgG4 concentrations are high in up to two-thirds and one-third, respectively, of patients tested.[13]

THERAPEUTIC PRINCIPLES

Patients with Inherited Disorders of Toll and IL-1 Receptor (TIR)-Mediated Immunity

- Patients should receive conjugated and nonconjugated vaccines against encapsulated bacteria (*Pneumococcus, H. influenzae, Meningococcus*).
- A preventive treatment, including antibiotic prophylaxis with cotrimoxazole plus penicillin V, should be administered throughout the life of the patient.
- Monthly prophylactic administrations of intravenous or subcutaneous immunoglobulins should be considered in selected patients.
- Empiric parenteral antibiotic treatment against *Streptococcus pneumoniae, Staphylococcus aureus,* and *Pseudomonas aeruginosa* should be initiated as soon as an infection is suspected or if the patient develops a moderate fever, without taking inflammatory parameters into account.
- Hemopoietic stem cell transplantation (HSCT) should be considered in selected patients with NEMO and IκBαα deficiency.

IRAK-4–deficient patients suffer from recurrent infections caused by pyogenic bacteria, mostly *S. pneumoniae, S. aureus,* and *P. aeruginosa,* with little or no fever or inflammatory response.[13] The leading pathogen responsible for infections in these patients is *S. pneumoniae,* which was found in half the cases of documented invasive infection (septicemia, meningitis, abscesses, or arthritis), whereas *S. aureus* and *P. aeruginosa* were found in 14% and 19% of such episodes, respectively. Patients with IRAK-4 deficiency also suffer from noninvasive pyogenic bacterial infections, mostly affecting skin and upper respiratory tract, with necrotizing infections particularly common. The principal bacterial strains isolated during noninvasive infections in patients with IRAK-4 deficiency are *S. aureus* in 43% of episodes, *P. aeruginosa* in 22% of episodes, and *S. pneumonia* in 16% of episodes. All reported sudden invasive infections occurred before the age of 14 years. IRAK-4 deficiency is a life-threatening disease, resulting in the deaths of 20 of the 49 known patients, all of whom died before the age of 8 years. Eleven of these patients died from invasive pneumococcal disease. There is an overall trend toward improvement with age, as shown by the seven adult patients doing well with no treatment at ages of 18 to 37 years.

Patients with IRAK-4 deficiency should also be immunized with *S. pneumonia, H. influenzae, N. meningitidis* conjugated vaccine (and also nonconjugated vaccine when available). A preventive treatment, including antibiotic prophylaxis with cotrimoxazole plus penicillin V, should be administered throughout the life of the patient. Given the severity of bacterial infection during childhood and the defective antibody production found in some patients with IRAK-4 deficiency, we also recommend the administration of empiric IgG replacement until the patient is at least 10 years old. This prophylaxis seems to decrease the incidence of invasive bacterial infections.[13] It is important to initiate empiric parenteral antibiotic treatment against *S. pneumoniae, S. aureus,* and *P. aeruginosa* as soon as an infection is suspected or if the patient develops a moderate fever, without taking inflammatory parameters into account because patients may die of rapid invasive bacterial infection despite appropriate prophylaxis. A diagnosis of IRAK-4 deficiency should be considered in children presenting with recurrent pyogenic infection and poor inflammatory responses.

MYD88 Deficiency

MyD88 deficiency is an AR disorder recently described in 22 patients (OMIM 612260).[12,13,16] Patients with MyD88 deficiency display a lack of production of IL-6 by whole blood and no CD62L shedding from granulocytes following activation with most of the TLR and IL-1R agonists tested, with the exception of TLR3, which signals through a MyD88-independent pathway.[12] Thus there seems to be no overt defect of leukocyte development in patients with MyD88 deficiency, and antigen-specific T- and B-cell responses appear to be normal, as shown by routine immunological analyses, in most cases.[16] Serum IgE and IgG4 concentrations were high in up to one-half and one-third, respectively, of the patients tested.[13] Some of the modest, subclinical abnormalities of B-cell responses observed, such as the production of low levels of antibodies against carbohydrates in some patients, may thus reflect impaired TACI responses, rather than impaired TLR and IL-1R responses. MyD88 deficiency, like IRAK-4 deficiency, confers a predisposition to severe bacterial infection, with impairment of the ability to increase plasma C-reactive protein (CRP) concentrations and to mount fever at the beginning of infection. However, pus formation has been observed at various sites of infection.

Patients with MyD88 deficiency present a narrow susceptibility to invasive pyogenic bacterial infections.[16] *S. pneumoniae* was involved in 41% of the documented invasive episodes in patients with MyD88 deficiency, whereas *S. aureus* and *P. aeruginosa* were found in 20% and 16% of such episodes, respectively. Most patients with MyD88 deficiency suffered from their first bacterial infection before the age of 2 years. Nine patients died from invasive bacterial infections, all before the age of 4 years, and most before the age of 1 year.[16] Seven of these patients died from invasive pneumococcal disease. However, MyD88 deficiency seems to improve with age, and none of the patients has presented invasive bacterial infection after adolescence.[13] MyD88 deficiency also confers a predisposition to noninvasive pyogenic bacterial infections, mostly affecting skin and the upper respiratory tract. The principal bacterial strains found during noninvasive infections were *S. aureus* in 53% of patients, *S. pneumoniae* in 20% of patients, and *P. aeruginosa* in 13% of patients. Most patients with MyD88 deficiency have had noninvasive bacterial infections, with half of the patients suffering from their first noninvasive bacterial infection before the age of 2 years. Patients with MyD88 deficiency continue to suffer from skin infections, sinusitis, or pneumonia, even after reaching adolescence. Treatment recommendations for patients with MyD88 deficiency are identical to those for those with IRAK-4 deficiency (see above). A diagnosis of MyD88 deficiency should be considered in children presenting with recurrent pyogenic infection with poor inflammatory responses.

THERAPEUTIC PRINCIPLES

Treatment of Patients with HSE Susceptibility

- Treatment with interferon (IFN)-α, in parallel with acyclovir, may help to improve disease outcome in patients with Toll-like receptor 3 (TLR3) pathway deficiencies, in particular.
- Serological monitoring for herpes simplex virus 1 (HSV-1) infection should be considered in individuals carrying mutations of TLR3 pathway genes but with no anti-HSV-1 antibody detectable in serum.
- In the absence of an effective vaccine against HSV-1, acyclovir may be considered an appropriate prophylactic treatment in individuals carrying mutations of TLR3 pathway genes, even if serologically negative for HSV-1.

HOIL1 Deficiency

The linear ubiquitination chain assembly complex (LUBAC) is assembled by a complex containing HOIL1 (also known as RBCK1), SHARPIN, and HOIP (also known as RNF31). LUBAC adds head-to-tail linear polyubiquitin chains to substrate protein and has been implicated in NF-κB signaling, together with NEMO, RIPK1, and ASC1. AR *HOIL1* (*HOIL1*/*RBCK1,* OMIM# 610924) deficiency, caused by homozygous or compound heterozygous mutations of the *HOIL1* gene, has been identified in three patients. These patients had a paradoxical clinical phenotype combining an autoinflammatory syndrome with pyogenic bacterial diseases, which led to their death.[41] LUBAC deficiency results in an impaired response to TNF and IL-1β in fibroblasts or to CD40L in B cells, but hyperresponsiveness to IL-1β in monocytes. All three patients had recurrent systemic inflammatory symptoms, present from the first few months of life, as a result of monocyte overactivation. They also developed recurrent pyogenic bacterial infections caused by *S. pneumoniae, H. influenzae, Escherichia coli, Staphylococcus* spp., and *Enterococcus,* partly because of their inability to produce antipolysaccharide antibodies. One patient also had chronic CMV infection and another had *Giardia intestinalis* infection.[41] All three patients also had amyotrophy, muscle weakness, and failure to thrive, partly as a result of muscular amylopectinosis complicated by myopathy and cardiomyopathy.[41] More recently, 14 patients with HOIL1 deficiency from 10 unrelated kindreds have been reported, with identification on the basis of neuromuscular and cardiac involvement secondary to amylopectinosis. Patients with HOIL1 deficiency should be immunized (with *S. pneumoniae* conjugated and nonconjugated vaccines, *H. influenzae* conjugated vaccine, and *N. meningitidis* conjugated and nonconjugated vaccines). Prophylactic penicillin V treatment and IgG substitution should also be considered. Heart transplantation should be considered for patients with heart failure. The autoinflammation is more difficult to control. One patient underwent HSCT, which corrected the autoinflammation and the predisposition to infectious disease.

HOIP Deficiency

AR *HOIP* (*HOIP* is also known as *RNF31,* OMIM 612487) deficiency caused by a homozygous mutation has also been identified in one patient. This patient displayed an autoinflammatory syndrome, pyogenic bacterial diseases, amylopectinosis, and lymphangiectasia.[42] At the cellular level, HOIP deficiency causes the same impaired responses to TNF and IL-1β agonists found in patients with AR HOIL-1 deficiency. Systemic inflammation results from the overactivation of monocytes in response to stimulation with IL-1β. Immunological investigations showed T-cell lymphopenia, an impaired proliferative response to anti-CD3, and hypogammaglobulinemia associated with nonprotective antibody responses to *S. pneumoniae* and *H. influenzae,* but with a good antibody response to tetanus toxoid. In addition to recurrent autoinflammation, this patient developed generalized lymphadenopathy after immunization with BCG. She also had recurrent episodes of diarrhea, amyotrophy, muscle weakness, and failure to thrive,[42] partly because of muscle amylopectinosis. Patients with HOIP deficiency should be treated in the same way as those with HOIL-1 deficiency.

GENETIC DISORDERS OF TH17-MEDIATED IMMUNITY

CMC is characterized by persistent or recurrent infections of skin (intertrigo, angular cheilitis), mucous membranes (oral, esophageal, genital), and nails (onychomycosis) caused by *Candida* (*C. albicans* in particular).[48] CMC was first described in the 1960s and was reported to display AD heritance. Isolated CMC usually begins early in infancy and affects otherwise healthy individuals. Invasive candidiasis, dermatophytosis, bacterial infections of the respiratory tract, and cutaneous infections caused by *Staphylococcus* have been reported in some patients. In rare cases, an association with cerebral aneurysm or oral/esophageal squamous cell carcinoma, and thyroid autoimmunity has also been described. These PIDs are caused by mutations of the *IL17RA, IL17RC, IL17F, ACT1,* and *STAT1* genes (Fig. 36.6).[17-20] Some patients with invasive candidiasis carry mutations of the *CARD9* gene.[49] The molecular and clinical features of CMC have been reviewed elsewhere. In particular, this condition is reported in patients with AR autoimmune polyendocrinopathy syndrome type 1 (APS-1 or APECED) (OMIM 240300). APS-1 is caused by mutations of the *AIRE* gene resulting in impaired T-cell tolerance, with 88% of patients developing CMC. High levels of neutralizing autoantibodies against IL-17A, IL-17F, and/or IL-22 have been detected in the serum of patients with APS-1.[50]

AR complete IL-17RA deficiency (OMIM 613953) was first identified in 2011 (see Fig. 36.6). A patient from a consanguineous family from Morocco was found to suffer from recurrent CMC (resistant to local antifungal treatment), cutaneous abscesses and folliculitis caused by *S. aureus,* acute otitis media, and lower respiratory tract infections.[17] No invasive bacterial infectious disease was reported in this patient. A homozygous premature stop codon was found to prevent production of the protein in this patient. The cellular phenotype was characterized by an absence of response to IL-17A and IL-17F homo- and heterodimers and to IL-17E (IL-25), in the patient's fibroblasts.[17] Seventeen other patients from nine unrelated kindreds have since been identified (Puel A).

AD IL-17F deficiency (OMIM 613956) was also first described in 2011 (see Fig. 36.6).[17] Patients from a multiplex kindred from

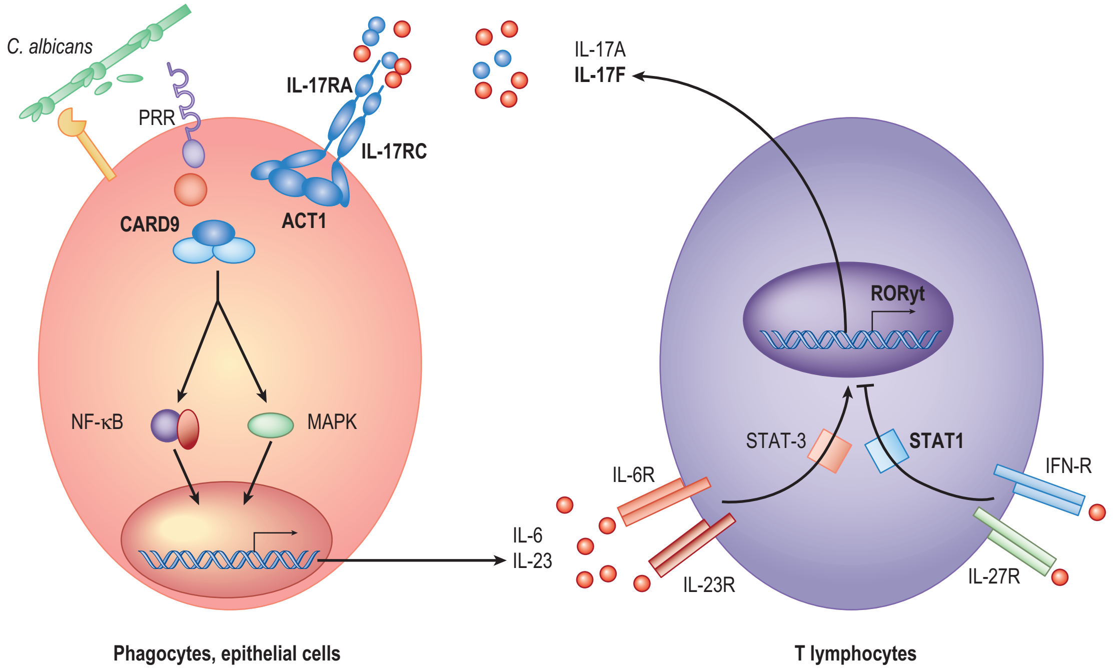


FIG 36.6 Genetic Etiologies Affecting the Interleukin (IL)-17 Pathway. Schematic representation of IL-17 immunity and cooperation between cells recognizing *Candida albicans* (phagocytes and epithelial cells) and cells producing IL-17 cytokines (T and innate [NK] lymphocytes). CMCD-causing mutations (in blue) of IL-17F, IL-17RA, IL17RC, CARD9, and ACT1 impair IL-17 function or response. CMCD-causing gain-of-function mutations (also shown in blue) in *STAT1* impair the development of IL-17–producing T cells.

Argentina displayed isolated CMC (see Fig. 36.6). A monoallelic mutation of *IL17F* was detected and found to have no impact on protein production. However, this mutation greatly decreased the activity of homo- and heterodimers (IL-17F/IL-17F or IL-17A/IL-17F) by affecting binding of the complex to the receptor. Another patient with a monoallelic mutation of *IL-17F* has been reported, but no cellular characterization was carried out.[17]

AR complete IL-17RC deficiency has been reported in three unrelated patients from Turkey and Argentina.[18] All patients suffered from CMC, and none presented invasive or recurrent bacterial infections. The clinical manifestations of infectious diseases in this group of patients resembled those in patients with AD IL-17 deficiency.[18] The three homozygous mutations identified, Q138X, R376X and R378X, confer a loss of IL-17RC protein expression in transfected HEK293T cells, with complete abolition of the response to IL-17A and IL-17 homo- and heterodimers.[18]

Genome-wide approaches led to the discovery of heterozygous *STAT1* missense mutations in patients with CMC disease (D) (OMIM 614162).[20] These mutations, unlike the previously reported mono- or biallelic *STAT1* loss-of-function mutations associated with susceptibility to mycobacterial, intracellular bacterial, and viral infections, were shown to be gain of function (GOF).[24] A recent study from the International STAT GOF group described 274 patients from 167 kindreds from 40 countries on five continents.[20] In total, 76 mutations of the *STAT1* gene have been identified and characterized *in vitro*. CMCD-causing *STAT1* mutations enhance the responses of STAT1 to IFN-α/β, IFN-γ, and IL-27 and of the repressors of IL-17 T-cell development, probably accounting for the low IL-17 T-cell counts in these patients, resulting in CMCD.[20] In AR APS-1 (APECED;)(OMIM 240300), caused by autoimmune regulator (AIRE) deficiency and resulting in impaired T-cell tolerance, 88% of patients develop CMC. High levels of neutralizing autoantibodies against IL-17A, IL-17F, and/or IL-22 have been detected in the serum of patients with APS-1.[50]

AR complete ACT1 deficiency has been reported in two patients from a consanguineous family from Algeria.[19] These two siblings developed CMC. However, one of them presented recurrent episodes of folliculitis decalvans and bilateral blepharitis caused by *S. aureus*. Their cellular phenotypes are characterized by impaired responses to IL-17A and IL-17F in fibroblasts and impaired responses to IL-17E in T lymphocytes.[19]

THERAPEUTIC PRINCIPLES

Treatment of Chronic Mucocutaneous Candidiasis (CMC)

- A preventive antifungal treatment, (principally fluconazole) should be administered in the long term, followed by other antifungal drugs, such as itraconazole or posaconazole.
- Antibiotic prophylaxis should be considered in selected patients with cutaneous staphylococcal disease.
- Monthly prophylactic administration of intravenous immunoglobulin should be considered in selected patients with recurrent pneumonia.
- Granulocyte–colony-stimulating factor (G-CSF), the JAK inhibitor ruxolitinib, or hemopoietic stem cell transplantation (HSCT) may also be considered.

AR complete ROR-γ/ROR-γT RORC deficiency has recently been identified in three unrelated consanguineous families from Israel, Chile, and Saudi Arabia.[51] RORC is a DNA-binding transcription factor that plays an important role in thymopoiesis. The patients suffered from an unusual combination of BCGosis and mild CMC in the absence of other clinical manifestations.[51] The three homozygous mutations identified, S38L, Q329X, and Q441X, affect both IL-17 and IFN-γ immunity. As predicted by studies in mice, the lack of functional ROR-γT protein prevents the development of IL-17-producing T cells, accounting for the chronic candidiasis observed in these patients. Surprisingly, the patients also lacked mucosal-associated invariant T (MAIT) cells and invariant natural killer T cells (iNKT), which normally produce IFN-γ and can inhibit intracellular mycobacterial replication. Moreover, their conventional CD8 α/β and γ/δ T cells, unlike their CD4 α/β T and NK cells, produced only very small amounts of IFN-γ. ROR-γ/ROR-γT thus plays a critical role in MAIT and iNKT development, and in the capacity of CD8 and γ/δ T cells to make IFN-γ. This results in a profound impairment of IFN-γ production by lymphocytes, leading to susceptibility to mycobacterial disease.[51] Antibiotic treatment can be combined with IFN-γ injection in cases of disseminated infection. The vaccination of these patients with live BCG is contraindicated.

Overall, these studies strongly suggest that IL-17 immunity plays a crucial role in defense against CMC in humans. The careful clinical description of a large series of patients would be required to determine the role of these cytokines more precisely.

CONCLUSIONS

ON THE HORIZON

- Unexplained infectious diseases occur in patients, in children in particular, as a result of an inborn error of immunity.
- Patients with severe infectious disease should be repeatedly investigated even if all standard immunological explorations are normal.
- The exploration of these defects, by application of powerful next-generation deep-sequencing technologies, provides tools for molecular diagnosis and genetic counseling and opens up exciting new perspectives benefiting patients.
- The identification of these defects will also guide rational treatment based on improvements in our understanding of these diseases' pathogenesis.

An understanding of the molecular basis of these PIDs affecting the innate immune responses mediated by IFN-γ, IFN-α/β/λ, IL-17, TIRs, or NF-κB has provided detailed insight into the pathogenesis of infections in affected patients, paving the way for genetic counseling and rational treatment design.[52] These PIDs should be sought in patients with unexplained infectious diseases, whether caused by a single or multiple infectious agents, even if all standard immunological explorations are normal.[53] Interestingly, even common infectious diseases, such as TB, invasive pneumococcal disease, and HSE may be favored by mendelian immune disorders. The discovery of many new PIDs opens up exciting new perspectives, not only increasing our understanding of immunity to pathogens but also benefiting patients. It is thought that most children with severe infectious diseases probably suffer from an underlying PID, and they should, therefore, be repeatedly investigated for known and unknown immunodeficiency conditions.

ACKNOWLEDGMENTS

We thank our laboratory colleagues, our collaborators at Necker Hospital for Sick Children and elsewhere, and patients and families for their assistance in making possible the studies contained in this chapter. We thank all the members of the Human Genetics of Infectious Diseases Laboratory for helpful discussions. The Laboratory of Human Genetics of Infectious Diseases is supported by grants from the Rockefeller University Center for Clinical and the Rockefeller University. The Laboratory of Human Genetics of Infectious Diseases is supported by the March of Dimes, the ANR, the INSERM, the FRM, and the PHRC.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Bustamante J, et al. Mendelian susceptibility to mycobacterial disease: genetic, immunological, and clinical features of inborn errors of IFN-gamma immunity. Semin Immunol 2014;26:454–70.
2. Kreins AY, et al. Human TYK2 deficiency: mycobacterial and viral infections without hyper-IgE syndrome. J Exp Med 2015;212:1641–62.
3. Casrouge A, et al. Herpes simplex virus encephalitis in human UNC-93B deficiency. Science 2006;314:308–12.
4. Zhang SY, et al. TLR3 deficiency in patients with herpes simplex encephalitis. Science 2007;317:1522–7.
5. Perez de Diego R, et al. Human TRAF3 adaptor molecule deficiency leads to impaired Toll-like receptor 3 response and susceptibility to herpes simplex encephalitis. Immunity 2010;33:400–11.
6. Herman M, et al. Heterozygous TBK1 mutations impair TLR3 immunity and underlie herpes simplex encephalitis of childhood. J Exp Med 2012;209:1567–82.
7. Sancho-Shimizu V, et al. Herpes simplex encephalitis in children with autosomal recessive and dominant TRIF deficiency. J Clin Invest 2011;121:4889–902.
8. Lim HK, et al. TLR3 deficiency in herpes simplex encephalitis: high allelic heterogeneity and recurrence risk. Neurology 2014;83:1888–97.
9. Andersen LL, et al. Functional IRF3 deficiency in a patient with herpes simplex encephalitis. J Exp Med 2015;212:1371–9.
10. Ciancanelli MJ, et al. Infectious disease. Life-threatening influenza and impaired interferon amplification in human IRF7 deficiency. Science 2015;348:448–53.
11. Picard C, et al. Pyogenic bacterial infections in humans with IRAK-4 deficiency. Science 2003;299:2076–9.
12. von Bernuth H, et al. Pyogenic bacterial infections in humans with MyD88 deficiency. Science 2008;321:691–6.
13. Picard C, et al. Clinical features and outcome of patients with IRAK-4 and MyD88 deficiency. Medicine (Baltimore) 2010;89:403–25.
14. Doffinger R, et al. X-linked anhidrotic ectodermal dysplasia with immunodeficiency is caused by impaired NF-kappaB signaling. Nat Genet 2001;27:277–85.
15. Courtois G, et al. A hypermorphic IkappaBalpha mutation is associated with autosomal dominant anhidrotic ectodermal dysplasia and T cell immunodeficiency. J Clin Invest 2003;112:1108–15.
16. Picard C, JCasanova JL, Puel A. Infectious diseases in patients with IRAK-4, MyD88, NEMO, or IkappaBalpha deficiency. Clin Microbiol Rev 2011;24:490–7.
17. Puel A, et al. Chronic mucocutaneous candidiasis in humans with inborn errors of interleukin-17 immunity. Science 2011;332:65–8.
18. Ling Y, et al. Inherited IL-17RC deficiency in patients with chronic mucocutaneous candidiasis. J Exp Med 2015;212:619–31.
19. Boisson B, et al. An ACT1 mutation selectively abolishes interleukin-17 responses in humans with chronic mucocutaneous candidiasis. Immunity 2013;39:676–86.
20. Toubiana J, et al. Heterozygous STAT1 gain-of-function mutations underlie an unexpectedly broad clinical phenotype: an international survey of 274 patients from 167 kindreds. Blood 2016;127:3154–64.
21. Casanova JL, Abel L. Genetic dissection of immunity to mycobacteria: the human model. Annu Rev Immunol 2002;20:581–620.

22. Bustamante J, et al. Germline CYBB mutations that selectively affect macrophages in kindreds with X-linked predisposition to tuberculous mycobacterial disease. Nat Immunol 2011;12:213–21.
23. Hambleton S, et al. IRF8 Mutations and human dendritic-cell immunodeficiency. N Engl J Med 2011;365:127–38.
24. Boisson-Dupuis S, et al. Inborn errors of human STAT1: allelic heterogeneity governs the diversity of immunological and infectious phenotypes. Curr Opin Immunol 2012;24:364–78.
25. Dorman SE, et al. Clinical features of dominant and recessive interferon gamma receptor 1 deficiencies. Lancet 2004;364:2113–21.
26. Olbrich P, et al. Diagnostic and therapeutic challenges in a child with complete interferon-gamma receptor 1 deficiency. Pediatr Blood Cancer 2015;62:2036–9.
27. Sologuren I, et al. Partial recessive IFN-(R1 deficiency: genetic, immunological and clinical features of 14 patients from 11 kindreds. Hum Mol Genet 2011;20:1509–23.
28. Moncada-Velez M, et al. Partial IFN-gammaR2 deficiency is due to protein misfolding and can be rescued by inhibitors of glycosylation. Blood 2013;122:2390–401.
29. Kong XF, et al. Haploinsufficiency at the human IFNGR2 locus contributes to mycobacterial disease. Hum Mol Genet 2013;22:769–81.
30. de Beaucoudrey L, et al. Revisiting human IL-12Rbeta1 deficiency: a survey of 141 patients from 30 countries. Medicine (Baltimore) 2010;89:381–402.
31. de Beaucoudrey L, et al. Mutations in STAT3 and IL12RB1 impair the development of human IL-17-producing T cells. J Exp Med 2008;205:1543–50.
32. Ouederni M, et al. Clinical features of Candidiasis in patients with inherited interleukin 12 receptor beta1 deficiency. Clin Infect Dis 2014;58:204–13.
33. Prando C, et al. Inherited IL-12p40 deficiency: genetic, immunologic, and clinical features of 49 patients from 30 kindreds. Medicine (Baltimore) 2013;92:109–22.
34. Bogunovic D, et al. Mycobacterial disease and impaired IFN-gamma immunity in humans with inherited ISG15 deficiency. Science 2012;337:1684–8.
35. Zhang X, et al. Human intracellular ISG15 prevents interferon-alpha/beta over-amplification and auto-inflammation. Nature 2015;517:89–93.
36. Minegishi Y, et al. Human tyrosine kinase 2 deficiency reveals its requisite roles in multiple cytokine signals involved in innate and acquired immunity. Immunity 2006;25:745–55.
37. Filipe-Santos O, et al. X-linked susceptibility to mycobacteria is caused by mutations in NEMO impairing CD40-dependent IL-12 production. J Exp Med 2006;203:1745–59.
38. Chapgier A, et al. Human complete Stat-1 deficiency is associated with defective type I and II IFN responses in vitro but immunity to some low virulence viruses in vivo. J Immunol 2006;176:5078–83.
39. Kawai T, Akira S. The role of pattern-recognition receptors in innate immunity: update on Toll-like receptors. Nat Immunol 2010;11:373–84.
40. Zhang SY, Casanova JL. Inborn errors underlying herpes simplex encephalitis: from TLR3 to IRF3. J Exp Med 2015;212:1342–3.
41. Boisson B, et al. Immunodeficiency, autoinflammation and amylopectinosis in humans with inherited HOIL-1 and LUBAC deficiency. Nat Immunol 2012;13:1178–86.
42. Boisson B, et al. Human HOIP and LUBAC deficiency underlies autoinflammation, immunodeficiency, amylopectinosis, and lymphangiectasia. J Exp Med 2015;212:939–51.
43. Paciolla M, et al. Rare mendelian primary immunodeficiency diseases associated with impaired NF-kappaB signaling. Genes Immun 2015;16:239–46.
44. Hanson EP, et al. Hypomorphic nuclear factor-kappaB essential modulator mutation database and reconstitution system identifies phenotypic and immunologic diversity. J Allergy Clin Immunol 2008;122:1169–77, e16.
45. Dupuis-Girod S, et al. Successful allogeneic hemopoietic stem cell transplantation in a child who had anhidrotic ectodermal dysplasia with immunodeficiency. Pediatrics 2006;118:e205–11.
46. Giancane G, et al. Anhidrotic ectodermal dysplasia: a new mutation. J Allergy Clin Immunol 2013;132:1451–3.
47. Ku CL, et al. Selective predisposition to bacterial infections in IRAK-4-deficient children: IRAK-4-dependent TLRs are otherwise redundant in protective immunity. J Exp Med 2007;204:2407–22.
48. Puel A, et al. Inborn errors of mucocutaneous immunity to Candida albicans in humans: a role for IL-17 cytokines? Curr Opin Immunol 2010;22:467–74.
49. Lanternier F, et al. Inherited CARD9 deficiency in otherwise healthy children and adults with *Candida* species-induced meningoencephalitis, colitis, or both. J Allergy Clin Immunol 2015;135:1558–68, e2.
50. Puel A, et al. Autoantibodies against IL-17A, IL-17F, and IL-22 in patients with chronic mucocutaneous candidiasis and autoimmune polyendocrine syndrome type I. J Exp Med 2010;207:291–7.
51. Okada S, et al. Immunodeficiencies. Impairment of immunity to *Candida* and *Mycobacterium* in humans with bi-allelic RORC mutations. Science 2015;349:606–13.
52. Casanova JL. Severe infectious diseases of childhood as monogenic inborn errors of immunity. Proc Natl Acad Sci USA 2015;112:E7128–37.
53. Casanova JL. Human genetic basis of interindividual variability in the course of infection. Proc Natl Acad Sci USA 2015;112:E7118–27.

MULTIPLE-CHOICE QUESTIONS

1. A 2-year-old patient was diagnosed of disseminated osteomyelitis caused by atypical mycobacteria. Immunophenotyping of T, B, and natural killer (NK) cells are normal. Respiratory burst excludes chronic granulomatous disease. What is the correct procedure in the investigation?
   A. Mitogen-induced T-cell proliferation
   B. Test of interferon (IFN)-γ pathway and treatment with multiple antimycobacterial antibiotics.
   C. Observation for 2 weeks while performing an evaluation

2. A 3-year-old boy was diagnosed of herpes simplex encephalitis, but no cutaneous lesion of herpes virus infection was observed. Immunophenotyping of lymphocytes was normal. To investigate the causal immune defect, what would be the best study to undertake?
   A. Observation during 1 week while treating with antibiotics. Acyclovir should be used if no sign of clinical recover.
   B. Test Toll-like receptors (TLR) response to multiple ligands using peripheral blood mononuclear cells.
   C. Test TLR3 responsiveness using dermal fibroblasts.

3. Patients with defect in interleukin-17 (IL-17) immunity are characterized by which of the following infectious phenotype?
   A. Patients present with persistent mucocutaneous candidiasis
   B. Patients present susceptibility to herpes virus
   C. Patients present with recurrent bacterial pneumonia

37

Infections in the Immunocompromised Host

Alexandra F. Freeman, Jennifer M. Cuellar-Rodriguez

As a result of improvements in antimicrobials and immunosuppressant agents, there are increasing numbers of hosts with immune deficiencies, either acquired primarily through genetic defects or acquired secondarily, for example, through treatment for malignancy and autoimmune disease, or after solid or hematopoietic stem cell transplantation (HSCT). Understanding both the genetic defects and the immunological targets of immunosuppressant agents will help further the knowledge of host control of infection. In this chapter, we review the infection spectrum of some of the major classes of primary immunodeficiencies, as well as acquired immunodeficiency, opportunistic infections in patients with human immunodeficiency virus (HIV) and acquired immunodeficiency syndrome (AIDS) are described separately in Chapter 39.

PRIMARY IMMUNODEFICIENCIES

Phagocyte Defects (Chapter 22)

Phagocytic neutrophils and monocytes are key members of the primary immune response. Neutrophils are essential in the initial host defense against microbes. Antimicrobial peptides, cytokines, and chemokines released at the site of microbial entry cause neutrophils to migrate to the site of inflammation (chemotaxis), ingest, and then kill the microbe through oxygen-dependent or oxygen-independent mechanisms. Defects in quantity or quality of neutrophils can predispose to infection, which is primarily with fungi and bacteria (Table 37.1). Defects in monocytes are less frequent and contribute to control of intracellular bacteria, mycobacteria, and fungi.

Chronic Granulomatous Disease

Chronic granulomatous disease (CGD) causes the most common qualitative neutrophil immunodeficiency. Defects in the nicotinamide adenine dinucleotide phosphate (NADPH) oxidase cause an abnormal neutrophil respiratory burst, leading to recurrent bacterial and fungal infections. However, it is interesting that the spectrum of infections is fairly limited, with the most common pathogens being *Staphylococcus aureus, Burkholderia cepacia, Serratia marcescens, Nocardia* spp., and *Aspergillus* spp.[1-3] Other infecting organisms occur less frequently but are rare outside of CGD, and these include *Chromobacterium violaceum, Aspergillus nidulans,* and *Granulibacter bethesdensis.*[3]

Infections in CGD are typically of the lung, lymph nodes, liver, and bone (Fig. 37.1).

Identification of the infecting organism is essential in CGD to guide antimicrobial therapy. CGD is not only characterized by exuberant granulomatous inflammation, both in areas without clear infection, such as in the gastrointestinal (GI) and genitourinary tract, but also in response to infection.[1,2] Therefore although necrotic and purulent centers may form in liver and lymph node infections, the yield with drainage may be poor because of the thicker consistency of granulomatous inflammation. At times, the infected lymph nodes may need resection to cure infections. Likewise, in lung infections, fine-needle aspiration frequently provides a higher yield than bronchoscopy because of the nature of inflammation. In addition, it is important to alert the microbiology laboratory when *G. bethesdensis* is suspected, such as with lymphadenitis, as growth requires special media.

The granulomatous inflammation seen in CGD infections may be intense enough to impede successful treatment of infection solely with antimicrobials and necessitate corticosteroid addition to appropriate antimicrobials. For instance, "mulch pneumonitis" occurs with large inhalations of decaying organic matter, such as mulch.[4] A diffuse pneumonitis associated with *Aspergillus* results and can be quite fulminant, with a high mortality if the inflammatory response is not treated (*e.g.,* with corticosteroids) in addition to antifungals. There is suggestion that the addition of corticosteroids may help with other infections in CGD, such as *Nocardia* pneumonia and *S. aureus* liver abscesses.[5,6]

Trimethoprim–sulfamethoxazole (TMP-SMX) is effective against the majority of bacterial pathogens in CGD and thus is an ideal prophylactic antibiotic and has been shown to significantly decrease bacterial infections.[7] Itraconazole was shown to be effective in preventing some of the fungal infections.[7] The newer triazoles, including voriconazole and posaconazole, have not been studied as prophylactic antimicrobials in this setting but have an extended spectrum and would likely be effective as well. Toxicities of extended use with certain antimicrobials need to be considered; for instance, voriconazole is associated with photosensitivity and increased skin cancers with prolonged use, as well as fluoride toxicity in rare cases.[8,9]

Leukocyte Adhesion Deficiencies

Leukocyte adhesion deficiencies (LADs) result from the inability of neutrophils to migrate to the site of infection.[2] LAD-1 is most frequent, resulting from a defect in β_2 integrin and presenting typically with failure of umbilical cord separation and omphalitis. The spectrum of infection is not as specific as with CGD but is limited typically to bacterial infections. Gingivitis and periodontitis lead to frequent oral bacterial infections, and necrotizing ulcerative skin infections are common, typically with *S. aureus* or gram-negative bacteria (GNB). Although primarily a defect of neutrophils resulting in bacterial infections, viral

TABLE 37.1 Infection Susceptibility of Select Primary Immunodeficiencies

Immunodeficiency	Genetic Defect	Functional Defect	Infection Susceptibility
Chronic granulomatous disease	Gp91^{phox}, p22^{phox}, p40^{phox}, p47^{phox}, p67^{phox}	Defect in nicotinamide adenine dinucleotide phosphate (NADPH) oxidase leading to abnormal superoxide production	Catalase positive bacteria (*e.g., Staphylococcus aureus, Burkholderia, serratia, Nocardia*) and filamentous molds
Leukocyte adhesion disorder-1 (LAD-1)	β_2 integrin	Neutrophil migration	Bacteria, typically *S. aureus,* gram-negative
Severe congenital neutropenia	*HAX1, ELANE*	Neutropenia	Bacteria, typically streptococci, *S. aureus,* gram-negative
X-linked agammaglobulinemia	Bruton tyrosine kinase (BTK)	Absence of B cells, antibody production	Encapsulated bacteria, enterovirus, *Helicobacter,* and related species
CD40 ligand deficiency	CD40 ligand	Impaired B-cell class switching	Encapsulated bacteria, *Pneumocystis jiroveci, Cryptosporidium*
Severe combined immunodeficiency (SCID)	Multiple genes, such as *IL2RG, RAG1/2, ADA, JAK3,* etc.	T-cell lymphopenia, variable B- and natural killer (NK)-cell expression	Bacteria, virus, *P. jiroveci, Candida,* Bacille Calmette-Guérin (BCG)
DiGeorge syndrome	Deletion of chromosome 22q11 in majority	Thymic hypoplasia or aplasia	Viruses, opportunists infrequent
Autosomal dominant hyper-IgE syndrome (Job's syndrome; AD-HIES)	Signal transducer and activator of transcription 3 (STAT3)	Impaired T-helper-17 (Th17) cell differentiation	*S. aureus, Candida,* dimorphic fungi; secondary infection of pneumatoceles with molds, gram-negative bacteria and nontuberculous mycobacteria
Dedicator of cytokinesis 8 (DOCK8) deficiency	DOCK8	T-cell lymphopenia, defect still being delineated	*S. aureus, Candida, Molluscum contagiosum,* human papillomavirus (HPV), herpes viruses
Interleukin-12 (IL-12)/interferon (IFN)-γ axis defects	IFN-γR1, IFN-γR2, IL-12, IL-12R, NEMO, STAT1	Failure of signal transducer and activator of transcription 1 (STAT1) activation, intracellular killing	*Mycobacteria, Salmonella,* dimorphic fungi
Complement C5–9 defects	Specific complements	Impaired membrane attack complex killing	*Neisseria* species

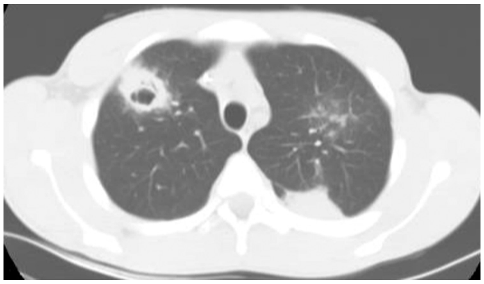


FIG 37.1 Chest computed tomography (CT) scan of a *Nocardia* pneumonia in a 17-year-old man with chronic granulomatous disease.

KEY CONCEPTS

Infection Risks in Primary Immunodeficiencies

- Neutrophil defects present primarily with bacterial and fungal infections.
- Humoral immunodeficiencies present with impaired antibody formation and primarily sinopulmonary infections with encapsulated organisms.
- Severe primarily T-cell defects usually present with opportunists, such as *Pneumocystis.*
- Abnormalities of the interleukin-12 (IL-12)/interferon (IFN)-γ axis commonly present with nontuberculous mycobacteria (NTM) infections.
- Cytokine autoantibody syndromes are being increasingly recognized with infection susceptibility dependent on the involved cytokine.

warts are not infrequent. Fungal infections are unusual. Little is known about antimicrobial prophylaxis for LAD-1, but typically prophylaxis with some coverage for oral flora, *S. aureus,* and GNB is provided, such as with amoxicillin/clavulanate.[7] Other rare neutrophil defects, including Chediak-Higashi and Griscelli syndromes, also are characterized by recurrent pyogenic infections.

Quantitative Phagocyte Defects

Neutropenia is the most common quantitative defect of phagocytes and typically results from cytotoxic agents, as discussed later in this chapter. Autosomal recessive defects in *HAX1* cause Kostmann syndrome, a type of severe congenital neutropenia, which typically presents in infancy with recurrent bacterial infections.[10] Autosomal dominant defects in the neutrophil elastase gene *ELANE* cause severe congenital neutropenia and cyclic neutropenia. Benign ethnic neutropenia is seen primarily in individuals of African descent and is usually asymptomatic. Secondary causes of neutropenia also include antineutrophil antibodies and hypersplenism with sequestration.

Infections associated with neutropenia relate to the etiology as well as the degree and duration of neutropenia. Cyclic neutropenia may be largely asymptomatic, with self-limited fevers and oral ulcers or cervical lymphadenopathy, as opposed to the severe chronic neutropenia of infancy that has more serious infections. Neutropenia is most frequently associated with bacterial infections, often localized to the mouth, cervical lymph nodes, lungs, and perianal area, with both gram-positive bacteria (GPB) and GNB. Fungal infections are more frequent in the prolonged and severe neutropenias associated with hematological malignancies and hematopoietic transplantation. Fever in the setting of neutropenia requires rapid assessment, with particular focus on physical examination of areas with high bacterial colonization, such as the oral cavity and perianal region, as well as sites of indwelling intravenous catheters. Empiric broad-spectrum antibiotics after obtaining appropriate cultures are often prudent.

Monocytopenia has been described in GATA2 deficiency and associated with disseminated Bacille Calmette-Guérin (BCG) with interferon (IFN) regulatory factor 8 (*IRF8*) deficiency.[11,12] The described spectrum of GATA2 deficiency is increasing and includes MonoMAC syndrome, Emberger syndrome with lymphedema and myelodysplasia, familial leukemia, aplastic anemia, and childhood-onset neutropenia.[11] The predominant clinical findings are disseminated mycobacteria; fungal infections, such as disseminated histoplasmosis; warts with cervical and vulvar dysplasia; pulmonary alveolar proteinosis (PAP); myelodysplasia; and leukemia. Onset is typically in later childhood and adulthood, and in addition to monocytopenia, lymphocytopenia may be present. When the patient is symptomatic, treatment is typically with HSCT.

HUMORAL IMMUNODEFICIENCIES (Chapter 34)

Humoral immunodeficiencies are characterized by absent or defective B cells with resultant lack of specific antibody responses, leading predominantly to infections with encapsulated bacteria. There is a spectrum of severity with X-linked agammaglobulinemia (XLA) characterized by a total absence of B cells and presentation in the first years of life and the later presentations of common variable immunodeficiency (CVID).[13]

Humoral immunodeficiencies are characterized by recurrent ear, sinus, and lung infections, with typical bacterial etiologies being *Streptococcus pneumoniae* and *Haemophilus influenzae*. Despite immunoglobulin (Ig) replacement, bronchiectasis may occur, and with it, the spectrum of pulmonary infections may broaden requiring careful microbiological monitoring. Neutropenia may occur with XLA and CD40 ligand deficiency, frequently when replacement Ig is not provided, and may be associated with a broader spectrum of infection, with more severe *S. aureus* and *Pseudomonas* infections. Chronic *Mycoplasma* and *Ureaplasma* spp. infections can occur leading to lung disease and arthritis. *Giardia* can be a source of GI disease. In severe humoral defects, such as XLA, persistent *Helicobacter, Flexispira,* and *Campylobacter* spp. infections are infrequent and are characterized by ulcerative skin lesions.[14,15] These infections require a high index of suspicion to be identified, as special microbiology media with longer lengths of incubation may be required. Eradication of infection, even with a combination of intravenous antibiotics, is often difficult. Severe meningoencephalitis with enteroviral infections may occur and can be quite devastating and difficult to treat. Replacement Ig therapy, given intravenously or subcutaneously, is used to minimize the recurrent encapsulated bacterial infections and diminish the risk of enterovirus infection. The role of prophylactic antibiotics is not well studied and varies among centers.

CD40 ligand deficiency (X-linked hyper-IgM) is the most common of the Ig class switch defects.[13] Sinopulmonary infections similar to XLA typically arise in early childhood. However, as opposed to XLA, *Pneumocystis jiroveci* pneumonia (PJP) occurs and may be the initial infection. Other infections indicative of T-cell defects occur as well, including cytomegalovirus (CMV) infection, toxoplasmosis, and cryptosporidiosis. Cryptosporidial infection can be a chronic and severe problem and may lead to sclerosing cholangitis.

PRIMARY CELLULAR AND COMBINED IMMUNODEFICIENCIES (Chapter 35)

Severe Combined Immunodeficiency

Severe combined immunodeficiency (SCID) results from a severe deficiency in the number or function of T cells. Depending on the genetic defect, B and/or natural killer (NK) cells are affected as well. Presentation is typically very early in life, with significant compromise within the first few months of life. In recent years, there has been more recognition of “leaky” forms of SCID, in which the defects are not complete and presentation can be later and more variable.[16]

SCID is associated with recurrent and persistent viral, bacterial, and fungal infections. PJP frequently occurs in early infancy, and prophylaxis should be provided. Chronic mucocutaneous candidiasis (CMC) is frequent and may require antifungal suppressive therapy. Diarrhea from viral and other etiologies frequently leads to failure to thrive. Severe respiratory viral infections and CMV also occur. In countries where BCG vaccine is given, disseminated infection may result; transplantation may lead to an immune reconstitution syndrome in which sites of BCG infection become inflamed and abscesses may form.

DiGeorge Syndrome

Immune deficiency in DiGeorge syndrome results from varying degrees of thymic hypoplasia or aplasia.[17] Other manifestations include congenital heart disease, characteristic facies and palate anomalies, hypocalcemia, and learning disabilities. With the exception of severe cases in which there is thymic aplasia and severe T-cell lymphopenia, opportunistic infections (OIs) are infrequent. Upper respiratory tract infections predominate, and prophylactic antibiotics may not be necessary. With the more mild defects, live viral vaccination may be considered safe.[18]

Autosomal Dominant Hyper-IgE Syndrome (Job’s Syndrome)

Autosomal dominant hyper-IgE syndrome (Job’s syndrome; AD-HIES), resulting from *STAT3* mutations, has been frequently characterized as a phagocytic defect. However, in recent years, it has been recognized that the immune defects center around the failure of T-helper 17 (Th17) cell differentiation and a decrease in memory T and B cells.[19,20] A lack of Th17 cells resulting in impaired interleukin-17 (IL-17) and IL-22 signaling and diminished antimicrobial peptide upregulation appears to at least explain some of the infection susceptibility.

AD-HIES is characterized by recurrent skin and lung bacterial infections, CMC, eczema, and a variety of connective tissue, skeletal, and vascular abnormalities.[20] *S. aureus* skin abscesses start early in life, and *S. aureus* colonization appears to drive the eczema. *S. aureus, S. pneumoniae,* and *H. influenzae* pneumonias typically start in the first few years of life, and prophylactic antimicrobials, such as TMP-SMX, can be useful in reducing the frequency. Healing of pneumonias appears aberrant, and bronchiectasis and pneumatoceles frequently result. The lung with these parenchymal abnormalities is then predisposed to mold infections (most frequently *Aspergillus* and *Scedosporium* spp.) and gram-negative infections, such as with *Pseudomonas aeruginosa* and nontuberculous mycobacteria (Fig. 37.2). These secondary infections can be quite difficult to eradicate. CMC is the most frequent fungal infection. PJP is much less frequent, typically occurring in infancy, potentially as the first pneumonia.

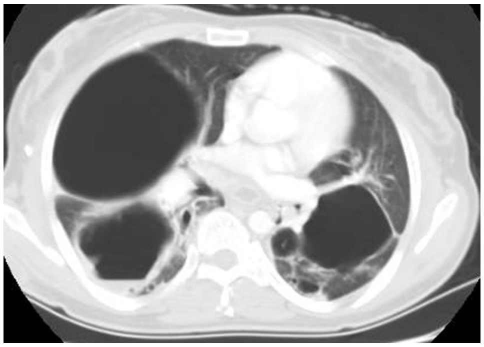


FIG 37.2 Chest computed tomography (CT) scan of a pneumatocele with secondary *Mycobacterium abscessus* and *Pseudomonas aeruginosa* infection in a 25-year-old woman with autosomal dominant hyper-IgE syndrome (Job's syndrome; AD-HIES).

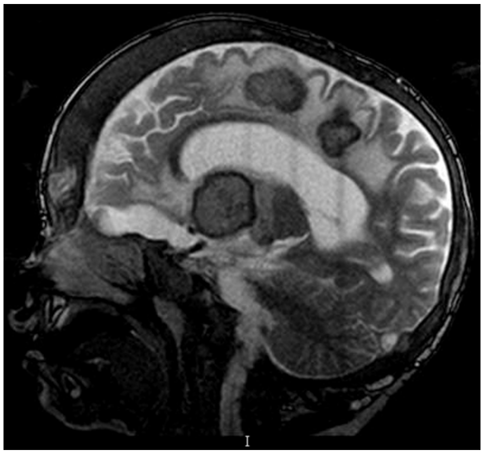


FIG 37.3 Brain magnetic resonance imaging (MRI) scan showing Bacille Calmette-Guérin (BCG) abscesses in a 3-year-old boy with an interferon (IFN)-γ receptor defect.

Disseminated cryptococcal infection has been described causing meningitis or GI infection. *Histoplasma* typically has caused disease of the GI tract, and *Coccidioides* infection has caused meningitis.[21]

DOCK8 Deficiency

Dedicator of cytokinesis 8 (DOCK8) deficiency is a combined defect that is one cause of an autosomal recessive hyper-IgE syndrome (AR-HIES).[22] Compared with AD-HIES, DOCK8 deficiency is characterized by severe cutaneous viral infections in addition to sinopulmonary infections. Recurrent cutaneous infections with herpes simplex virus (HSV) and varicella-zoster virus (VZV) occur, as well as severe, disfiguring warts and *Molluscum contagiosum* infection. DOCK8 deficiency has an overall worse prognosis than AD-HIES, with frequent occurrence of malignancies, typically squamous cell carcinoma and lymphomas.

Defects of the IL-12/IFN-γ Axis

Defects of the IL-12/IFN-γ axis are characterized primarily by mycobacterial infections. Monocyte-derived macrophages ingest intracellular bacteria, such as *Mycobacterium* and *Salmonella,* leading to IL-12 secretion. IL-12 binds to its receptor on T lymphocytes and NK cells leading to IFN-γ secretion, which then binds to its heterodimeric receptor on macrophages, activating microbial killing through signal transducer and activator of transcription 1 (STAT1). Many defects along this pathway have been described, including mutations in IL-12, IL-12 receptor, IFN-γ receptor1 and 2, STAT1, and nuclear factor (NF)-κB essential modulator operon (NEMO).[23,24] Defects along this pathway show varying susceptibility to nontuberculous mycobacteria (NTM), *Salmonella,* and endemic dimorphic fungal infection (Fig. 37.3). For example, individuals with dominant IFN-γ receptor defects retain some IFN-γ signaling and tend to have localized NTM infections, often presenting with osteomyelitis, and infrequently other OIs, such as disseminated histoplasmosis. Treatment of the dominant IFN-γ receptor defects is typically successful with combination antimycobacterial agents with or without exogenous IFN-γ. In contrast, autosomal recessive IFN-γ receptor defects, in which there is typically no residual IFN-γ signaling, have much more severe disease, with earlier onset of disseminated NTM or BCG disease, poor response to therapy with frequent relapse, and a greater susceptibility to infection, including some viral susceptibility. Death in childhood is frequent, and HSCT is typically considered.

Compared with loss of function STAT1 mutations, which are characterized by disseminated NTM infections, gain-of-function STAT1 (GOF-STAT1) mutations have a much broader infection susceptibility pattern and varied presentations.[25] Although initially described as causing CMC, GOF-STAT1 mutations are also associated with disseminated endemic mycoses, such as *Coccidioides* infection, viral infections including those of the herpes family and progressive multifocal leukoencephalopathy (PML), and bacterial infections. GOF-STAT1 mutations can also be associated with NTM infections.

COMPLEMENT DEFICIENCIES (Chapter 21)

Infections with encapsulated organisms are the typical manifestations of complement defects.[26] C3 deficiency is associated with recurrent and severe infections with *S. pneumoniae, H. influenzae,* and *Neisseria meningitidis,* whereas deficiencies of the components of the terminal pathway composing the membrane attack complex, C5-9, result in *N. gonorrhoeae* and *N. meningitidis* infections.

ASPLENIA

The spleen acts as both a source for antibody production and a filter to remove pathogens through phagocytic cells. Spleens can be physically absent, either congenitally or through surgical resection (*e.g.,* with intractable hemolytic anemia or thrombocytopenia, malignancy, trauma), or functionally absent as is seen with increased age in sickle cell disease. Infections associated with asplenia include sepsis with encapsulated bacteria, namely, *S. pneumoniae* and *H. influenzae,* and with *Babesia,* a protozoan that infects erythrocytes and causes fever and hemolysis with asplenia.[26] Similar to babesiosis, malaria infection can be more severe in patients without a spleen. The risk of infection differs with the etiology of asplenia, with splenectomy associated with trauma having the lowest risk as typically small amounts of splenic tissue remain. In addition, infection with encapsulated organisms appears to be more common in children with asplenia than in adults with asplenia.

SECONDARY NON–MEDICATION-ASSOCIATED IMMUNODEFICIENCY

Cytokine Autoantibodies

In recent years, there has been greater recognition of susceptibility to infection associated with anticytokine autoantibodies.[27-30] Autoimmune polyendocrinopathy with candidiasis and ectodermal dysplasia (APOCED) is a disorder caused by mutations in the autoimmune regulator gene (*AIRE*). Many autoimmune manifestations are seen in APOCED, including hypoparathyroidism, diabetes, and adrenal failure as a result of persistent autoreactive T cells caused by failure of thymic deletion. The main infectious complication is severe CMC. Neutralizing autoantibodies to IL-17 and IL-22 have been detected and may contribute to the occurrence of CMC, similar to other defects along the Th17/IL-17/IL-22 pathway, such as AD-HIES.[29,30] Neutralizing autoantibodies to IFN-γ leading most frequently to disseminated NTM infections have been described.[27] This disorder typically presents in adulthood and has an increased prevalence in Asians. Although the primary manifestation of autoantibodies against granulocyte macrophage–colony-stimulating factor (GM-CSF) antibodies is PAP, an increased frequency of OIs, such as with *Nocardia* and *Cryptococcus,* has been described.[30]

THERAPEUTIC PRINCIPLES

Prevention of Infection in Patients With a Defect in Host Defenses

- Use of prophylactic antibiotics in patients at high risk for a specific type of infection
- Immunization to prevent specific bacterial and viral infections:
    - Active immunization, especially in patients who should be capable of mounting an effective response (*e.g.,* before elective splenectomy or before elective initiation of immunosuppressive therapy)
    - Passive immunization through administration of high-titer pooled immunoglobulin to patients exposed to or at high risk of specific viral infections.

INFECTIONS IN PATIENTS RECEIVING IMMUNOSUPPRESSIVE MEDICATIONS

Immunosuppressive drugs have been used for over 60 years[31] and are essential for the management of autoimmune/inflammatory diseases and hematological and oncological malignancies and in transplant recipients. Early immunosuppressive agents lacked specificity and therefore led to serious and numerous adverse effects. Newer immunosuppressive agents have revolutionized the treatment of multiple diseases, but many of these have led to an increase in specific infectious complications. Knowledge of the particular infection or treatment associations is imperative.[32]

Described below are the most common classes of drugs that have been linked to infectious complications. Notably, immunosuppressive agents are frequently used in combination, and therefore establishing clear causality is challenging.

Cytotoxic Agents (*e.g.,* Cyclophosphamide, Methotrexate, Azathioprine)

Initially developed to control neoplastic growth in oncological and hematological malignancies, cytotoxic drugs have also found their niche as an essential component of the conditioning regimen for HSCT and solid organ transplantation (SOT), as immunomodulators in autoimmune and inflammatory disorders, and graft-versus-host disease (GvHD). These drugs inhibit B-cell and T-cell proliferation and hinder cellular and humoral immune responses in a dose-dependent fashion.[33] One of the main side effects of these agents is hematopoietic toxicity; they induce neutropenia to varying degrees. Fever occurs during chemotherapeutic-induced neutropenia in 10–50% of patients with solid tumors and in >80% of those with hematologic malignancies.[34] Common sites of infection include the intestinal tract, lungs, and skin; bacteremia occurs in up to 25% of patients. Probably because of the high incidence of indwelling catheters and widespread use of prophylactic antibiotics, currently coagulase-negative staphylococci are the most common blood isolates, followed by drug-resistant Enterobacteriaceae and nonfermenting GNB (*e.g., P. aeruginosa*).[34] Invasive mold infections (*e.g.,* aspergillosis) typically occur after prolonged neutropenia (>2 weeks). Invasive yeast (usually *Candida* spp.) infections are more commonly seen in patients with severe mucositis and neutropenia. Detailed guidelines for the use of antimicrobial agents in the setting of chemotherapy-induced neutropenic fever have been published.[34] Although bacterial infections are the most common infections, when used as immunomodulators, a higher frequency of OIs, such as PJP, and infections with *Listeria* spp., NTM, *Cryptococcus,* and VZV, has been described.

Glucocorticoids

Depending on the dose and duration, glucocorticoids can induce a vast range of immune defects through decreasing cytokine production (IL-1, IL-6, and tumor necrosis factor-α [TNF-α]), and impairing neutrophil and lymphocyte trafficking and function. Infections are a frequent complication of corticosteroid use.[32] Bacterial infections are the most common infectious complication, but opportunistic fungal, viral, and mycobacterial infections are also seen, particularly with high doses and long duration of systemic therapy. Some associated infections include GPB and GNB, superficial and invasive candidiasis, invasive aspergillosis and nocardiosis, cryptococcosis, PJP, listeriosis, endemic mycoses, tuberculosis (TB), and infections with NTM, and VZV, among others. Outside of the setting of SOT and HSCT, routine use of antimicrobials for patients receiving corticosteroids is not recommended, except for anti-TB treatment (*e.g.,* isoniazid) for those with a positive tuberculin skin test (purified protein derivative [PPD]) or a positive IFN-γ release assay, and TMP-SMX in those receiving high-dose steroid for extended periods.[35] Alternate-day use of corticosteroids has been associated with a lower risk of severe infection; therefore whenever possible, this dosing schedule should be used.[31]

Calcineurin Inhibitors and Mammalian Target of Rapamycin Inhibitors

The calcineurin inhibitors, cyclosporine and tacrolimus, have been cornerstones in improving the outcome of transplant recipients by reducing T-cell function and thus preventing allograft rejection in SOT and GvHD in HSCT. Infectious complications, typically viral, seem to be dose dependent and seen more often when these drugs are used in combination with other immunosuppressants.[32,33] The mammalian target of rapamycin (mTOR) inhibitors sirolimus and everolimus have been associated with a lower incidence of CMV infection, compared with other immunosuppressive agents used in transplantation.[32] Although impaired wound healing has been seen, this has not correlated with an increased incidence of infections.

Mycophenolate Mofetil

Mycophenolate mofetil (MMF) is a prodrug of mycophenolic acid, which inhibits inosine monophosphate dehydrogenase, a key enzyme in purine synthesis. It is a cytotoxic drug with antiproliferative effect on T and B lymphocytes.[33] In addition to its potential myelosuppressive effect, it has been associated in SOT to a higher risk of CMV infection.[32]

Antithymocyte Globulin

Antithymocyte globulin (ATG) is a polyclonal Ig prepared by immunizing horses or rabbits with human thymocytes and harvesting the IgG. These antibodies produce a profound lymphopenia that can last well beyond 1 year. Used for induction or rejection therapy in transplantations, their use has been associated consistently with a greater risk of CMV and posttransplantation lymphoproliferative disease (PTLD).[36] CMV prophylaxis or preemptive therapy is recommended in transplant recipients that receive ATG. Because these are foreign proteins, serum sickness can also develop between 1 and 2 weeks after treatment.

Monoclonal Antibodies and Small Molecules

Monoclonal antibodies (mAbs) and small molecules target specific cellular steps, making it possible to manipulate disease pathways for previously untreatable illnesses. Many of these biological agents interfere with specific aspects of the immune system. Highlighted below are those that have a key infection susceptibility pattern. Others with more subtle increases in infections are also included in Table 37.2.

Tumor Necrosis Factor-α Inhibitors

TNF-α is a cytokine that plays a central role in macrophage and phagosome activation, differentiation of monocytes into macrophages, recruitment of neutrophils and macrophages, and granuloma formation and maintenance. Blockade results in an increased risk of infection, particularly in the first few months after treatment.[37,38] Infections characterized by granulomas have been described most frequently, in particular TB and NTM infections, but also disseminated histoplasmosis and coccidioidomycosis.[39] Other notable infections are listeriosis, legionellosis, salmonellosis, recurrent or relapsing leishmaniasis, overwhelming *Plasmodium falciparum* infection, and invasive mold infections (*i.e.,* aspergillosis and mucormycosis). Risk of certain viral infections also appears to be increased, particularly VZV infection. Several cases of severe hepatitis B virus (HBV) reactivation in patients with positive surface antigen at the start of treatment have been reported, many of whom developed hepatic failure. Patients receiving anti–TNF-α therapy should be screened for TB (latent or active), HBV, and HCV before starting treatment, and treated accordingly.[38] Patients with latent TB should receive isoniazid for 6–9 months. Ideally, patients would have received at least 3 weeks of isoniazid before therapy with biologicals; however, in truly urgent cases, this can be done simultaneously. In patients that are hepatitis B surface antigen (HbsAg) positive, baseline HBV DNA levels should be obtained and antiviral prophylaxis or treatment offered. All HBV-seronegative patients should be vaccinated against hepatitis B and seroconversion documented.

TABLE 37.2 Selected Monoclonal Antibodies and Small Molecules and Their Risk of Infection

Drug Name	Target	Serious Infections
Adalimumab Certolizumab pegol Etanercept Golimumab Infliximab	TNF-α*	TB, NTM infection, endemic mycoses, listeriosis, candidiasis, invasive mold infections, nocardiosis, bacterial infections, reactivation of VZV and HBV infections, severe malaria, and leishmaniasis
Rituximab	CD20	HBV infection reactivation, PML, PCP, enteroviral meningoencephalitis, CMV disease, VZV infection, babesiosis, parvovirus infection, nocardiosis
Anakinra Rilonacept	IL-1	Bacterial infections (cellulitis and pneumonia)
Alemtuzumab	CD52	Herpes viruses reactivation (EBV, VZV, CMV, HHV-6), severe respiratory viral infections (with adenovirus, RSV, parainfluenza virus, influenza virus), PCP, invasive mold infections, histoplasmosis, cryptococcosis, PTLD, bacterial meningitis, toxoplasmosis, PML, parvovirus infection, nocardiosis, TB, NTM infection, acanthamebiasis, and *Balamuthia mandrillaris* infection
Muromonab	CD3	Bacterial infections; listeriosis; nocardiosis; aspergillosis; candidiasis; cryptococcosis; toxoplasmosis; infections with CMV, HSV, VZV, adenovirus, RSV, and enteroviruses; and PCP
Basiliximab Daclizumab	CD25	Bacterial infections; infections with CMV and HSV; EBV-associated PTLD; infections with RSV, adenovirus, influenza, and BK virus; invasive mold infections; candidiasis; nocardiosis; TB; NTM infection; toxoplasmosis
Natalizumab	α-4 Integrin	PML, respiratory viral and bacterial infections, UTI, viral meningitis and encephalitis, CMV infection, IA, PCP, VZV infection, NTM infection, cryptosporidiosis
Abatacept Belatacept	T-cell costimulation blockade	Bacterial infections, EBV-associated PTLD
Alefacept	Inhibits T-cell activation	1 case of *Mycobacterium avium intracellulare* (MAI) bursitis, 1 case of nocardiosis when coadministered with infliximab
Brodalumab	IL 17 receptor A	Mucocutaneous candidiasis
Bortezomib	Proteasome inhibitor	Herpes zoster
Gemtuzumab	CD33	Bacteremia, bacterial pneumonia, and HSV reactivation.

*TNF-α, tumor necrosis factor-α; IL-1, interleukin-1; TB, tuberculosis; NTM, nontuberculous mycobacteria; VZV, varicella-zoster virus; HBV, hepatitis B virus; PML, progressive multifocal leukoencephalopathy; PCP, *Pneumocystis jiroveci* pneumonia; CMV, cytomegalovirus; EBV, Epstein-Barr virus; HHV-6, human herpes virus 6; RSV, respiratory syncytial virus; PTLD, posttransplantation lymphoproliferative disease; HSV, herpes simplex virus; UTI, urinary tract infection; IA, invasive aspergillosis.

Rituximab

Rituximab is a chimeric murine–human mAb that targets CD20 on mature B lymphocytes. It results in rapid and profound depletion of B cells that can last up to several months.[40] In most adults, serum Ig levels remain largely stable, since plasma cells do not express CD20. Initial trials suggested that rituximab had minimal effect on the occurrence of infections; however, more recent meta-analyses have reported a higher risk of infections, especially with repeated administration and in patients with underlying immune defects or concomitant significant immunosuppression.[40] HBV reactivation occurs, and there have been reports of fulminant hepatitis and death after rituximab treatment, particularly when used in combination therapy (*e.g.,* rituximab with cyclophosphamide, doxorubicin, vincristine, and prednisone [R-CHOP]). Also a reverse seroconversion phenomenon has been described, with loss of protective HBV surface antibodies and reactivation.[38] Assessment of HBV status before starting rituximab is recommended. To reduce the risk of reactivation, HBV suppression with antiviral medication should be considered.

Since the initial approval of rituximab, there have been over 280 cases of PML associated with its use. Most cases have been reported in patients with hematological malignancies; however, some have been reported in patients with autoimmune or inflammatory diseases, who take other immunosuppressants as well. Rituximab now carries a black box warning about the risk of PML and death.[41] There have been many reports of PJP following use of rituximab, although most patients also received other immunosuppressive therapies. The need for PJP prophylaxis with its use remains to be determined. The greatest risk of infection appears to be related to viral reactivation, particularly in children; as usual, this is confounded by prior or concomitant immunosuppressive medications. In children with immune-mediated neurological diseases, infections after rituximab use are common; even lethal CMV disease has been reported.[42] Other rare infections that have been described are enteroviral meningoencephalitis, CMV disease, disseminated VZV infection, refractory babesiosis, parvovirus B19 infection, and nocardiosis.[37]

Alemtuzumab

Alemtuzumab is a humanized IgG1 that targets CD52 on B and T lymphocytes, monocytes, macrophages, and NK cells. It lyses these cell populations and results in profound and sustained deficits in cellular and humoral immunity that lasts for several months; CD4 and CD8 cell count reach their nadir approximately 4 weeks after treatment, but median counts remain at <25% of baseline for approximately 9 months.[38] Opportunistic and nonopportunistic infections have been associated with its use, particularly reactivation of herpes virus (*e.g.,* Epstein-Barr virus [EBV], CMV, disseminated VZV) and fungal infections (*e.g.,* PJP, invasive molds, and dimorphic fungi). Moreover, several cases of TB and NTM infection have been reported. The incidence of these OIs varies with the administered dose and whether it is used as a single agent or in combination with other immunosuppressants (*e.g.,* induction therapy [4.5%] versus treatment for rejection [21%] in SOT).[37,38] When used for T-cell depletion in allogeneic HSCT (allo-HSCT) it has been associated with a very high risk of CMV reactivation (50–85%) and disease, severe adenovirus infection, human herpesvirus-6 (HHV-6) encephalitis, respiratory viral infections that frequently progress to pneumonia, and PTLD. Other infections that have occurred are overwhelming bacteremia, bacterial meningitis, toxoplasmosis, PML, disseminated amoebiasis, parvovirus infection, and nocardiosis. All patients receiving alemtuzumab should receive PJP and herpesvirus prophylaxis for a minimum of 2 months after therapy or until CD4 counts are ≥200 cells/μL (Campath package insert). Given the high incidence of CMV reactivation and disease, prophylaxis or preemptive therapy is warranted.

Daclizumab and Basiliximab

Daclizumab and basiliximab are humanized and chimeric mAbs, respectively, that target CD25, the α chain of the IL-2 receptor complex expressed on activated alloantigen-reactive T lymphocytes. They competitively inhibit IL-2 binding and prevent IL-2 mediated activation of lymphocytes and cytokine release. They are mainly used in transplantation to prevent rejection or steroid-refractory GvHD.[37] When used as prophylaxis for rejection in SOT, their use does not seem to be associated with an increased risk in infections. In allo-HSCT, high rates of mortality and of OIs (95%) have been reported in patients receiving daclizumab for steroid-refractory GvHD; however, given the high risk of OIs in this population in general, the actual role of daclizumab is unclear. Infections described in this setting are common bacterial, viral (CMV, BK virus, adenovirus, HSV, RSV, influenza virus), EBV-associated PTLD, invasive fungal infections (mold and yeasts), as well as *Legionella* and *Nocardia* and NTM infections, TB, and toxoplasmosis.[38]

Natalizumab

Natalizumab is a humanized IgG4 that targets the α_4 subunit of $\alpha_4\beta_1$ and $\alpha_4\beta_7$ integrins, found in lymphocytes; it inhibits their binding to cellular adhesion molecules (vascular cell adhesion molecule-1[VCAM-1] and mucosal addressin-cell adhesion molecule-1 [MAdCAM-1]) in the central nervous system (CNS) and the GI tract, thereby attenuating inflammation in these tissues.[38,43] It is associated with a profound decrease in CD4, CD8, and CD19 lymphocytes in cerebrospinal fluid (CSF), and it is used in the treatment of multiple sclerosis and Crohn's disease. Natalizumab was temporarily withdrawn from the market in 2005 after three cases of PML were reported in patients with prolonged exposure. It has now been reintroduced with a black box warning about the risk of PML. Based on retrospective global surveillance data, the overall incidence of PML in natalizumab treated patients is 1.01 cases/1000 patient-years. However, the risk of PML is significantly increased in patients who have detectable anti-JC virus antibodies before the initiation of therapy, had prolonged exposure to natalizumab (>24 months), and had prior use of immunosuppressants (11.1 cases/1000 patients [95% confidence interval (CI) 8.3–14.5]). The incidence in patients with no detectable anti-JC virus antibodies is calculated at 0.09 cases/1000 patients (95% CI 0–0.48).[43] Increases in bacterial and viral infections and OIs have been reported as well, but most patients with these infections were receiving concurrent immunosuppression.

Bortezomib

Bortezomib is a dipeptide boronate proteasome inhibitor that causes G2–M cell cycle arrest and apoptosis, which ultimately impacts T-cell immunity. It is now approved for the treatment of multiple myeloma and mantle cell non-Hodgkin lymphoma. An increased incidence of herpes virus infections, in particular VZV infection, has been reported. The incidence in patients with myeloma treated with bortezomib is around 13–22% and in relapsed/refractory mantle cell lymphoma around 10.3%, with age ≥65 years as the only predictive factor.[44]

INFECTIONS IN SOLID ORGAN TRANSPLANTATION

KEY CONCEPTS

Determinants of the Risk for Infections in Transplant Recipients

- Epidemiological exposures (*e.g.,* residing in a *Histoplasma* endemic area)
- Immunosuppressive drugs
- Antimicrobial prophylaxis
- Type of transplantation (*e.g.,* cord transplant vs matched related donor transplantation)
- Time after transplantation

Despite significant improvement in the survival of SOT recipients, infectious complications remain a major cause of morbidity and mortality. The risk of infection is determined by the interaction of various factors, such as age of the recipient, type of transplantation, invasive procedures, dose and duration of immunosuppressive drugs, epidemiological exposures of both donor and recipient, use of antimicrobial prophylaxis, donor–recipient serological status to certain infections (*e.g.,* CMV infection, EBV infection, toxoplasmosis), and ongoing viral replication (so-called indirect effects).[45] Assessment of a recipient's risk for infection should help tailor specific prophylactic strategies and infectious workup when an infection is suspected.

Age is an important determinant of susceptibility to infections; it impacts the likelihood of prior exposures to microbial pathogens, either by primary infection or vaccination. History of exposure can have either positive or negative effects. Older patients are more likely to have encountered pathogens that can remain latent and reactivate at the time of transplantation (*e.g.,* CMV); younger patients have a higher risk of acquiring primary infections after transplantation, and these infections tend to be more severe than disease secondary to reactivation of a latent infection. In addition, preexisting immunity can have a protective effect against clinical disease (*e.g.,* EBV-associated PTLDs).[45]

The type of allograft affects the specific infectious risk, usually as a result of technical factors associated with the transplantation procedure, but is also inherent to the transplanted organ. For example, urinary tract infections are most common after kidney transplantation, as a result of either catheter placement or ureteric stenting. BK virus is ubiquitous, but BK virus–associated transplant nephropathy is most common after kidney transplantation, and it is an important cause of allograft loss. Infection after liver transplantation frequently results from leaks from biliary and GI anastomoses. Cardiac assist devices in heart transplant recipients are a frequent source of infection. Tracheal anastomotic infections, particularly caused by fungi, are a significant complication of lung transplantation. In single lung transplantations, recurrent infections of the native lung, such as with GNB or fungi, can extend to the transplanted lung. In small intestine transplantation, opportunistic and nonopportunistic viral infections of the GI tract (*e.g.,* with norovirus) are common; these can be severe and even life threatening.[45]

Despite all the differences in individual risk of infection, the general patterns of infection in the absence of antimicrobial intervention are similar among those undergoing SOT. This predictable temporal pattern has enabled the institution of specific prophylactic strategies and in the development of an infection differential diagnosis.[45]

CLINICAL PEARLS

Infections in Solid-Organ Transplant Recipients

- Most infections occur in the first 6 months following transplantation.
- In the first month, most infections are nosocomial or are related to surgical procedures or preexisting infection in the donor or recipient.
- Infections between 1 and 6 months following transplantation are commonly secondary to the use of immunosuppressive drugs, with a predominance of cytomegalovirus infection.
- Infections after 6 months are usually due either to chronic viral infection, acquired earlier, or to chronic allograft dysfunction, necessitating repeated courses of high-dose immunosuppression.

Infections in the First Month After Transplantation

Infections in this period are generally associated with technical complications of the surgery or have a nosocomial etiology. Donor-transmitted infections (*e.g.,* Chagas disease) or reactivation of prior infections (*e.g.,* colonization and infection with multidrug-resistant [MDR] bacteria) can also occur.[45]

Infections 1–6 Months After Transplantation

Before routine antimicrobial prophylaxis, this period was typically characterized by the presence of OIs, such as PJP and CMV infection; the incidence of these infections has been significantly reduced or delayed with TMP-SMX prophylaxis and CMV prophylactic or preemptive therapy. Reactivation of latent infections, such as TB, Chagas disease, endemic mycoses, and cryptococcosis, can occur. Viral infections, such as those with BK virus, adenovirus, RSV, HBV, and EBV, are common (Fig. 37.4). Invasive fungal infections, specifically *Aspergillus* and non-*Aspergillus* mold infections, can be problematic during this period of heightened immune suppression.[45]

Infections 6 Months After Transplantation

This period is less well defined; patients with satisfactory allograft function develop more severe manifestations of community-acquired infections but have a lower risk of OIs. Patients with poor allograft function and, therefore, increased immunosuppression or those with ongoing chronic viral reactivation are at high risk of OIs, such as invasive fungal infections, late-onset CMV infection, nocardiosis, PJP, listeriosis, and EBV-associated PTLD. In addition, lung transplant recipients with chronic graft dysfunction are at high risk of recurrent bacterial pneumonia. Similarly, liver transplant recipients with chronic graft dysfunction frequently develop biliary strictures and recurrent cholangitis.[45]

INFECTIONS IN HEMATOPOIETIC STEM CELL TRANSPLANTATION

Survival after HSCT has significantly improved during the last decade. Factors contributing to the increase in survival are the use of less toxic, nonmyeloablative conditioning regimens, use of peripheral blood stem cells (PBSCs) that leads to faster neutrophil recovery, and use of antimicrobial preventive therapies.[46] Despite these advances, infections are still a frequent cause of morbidity and mortality.

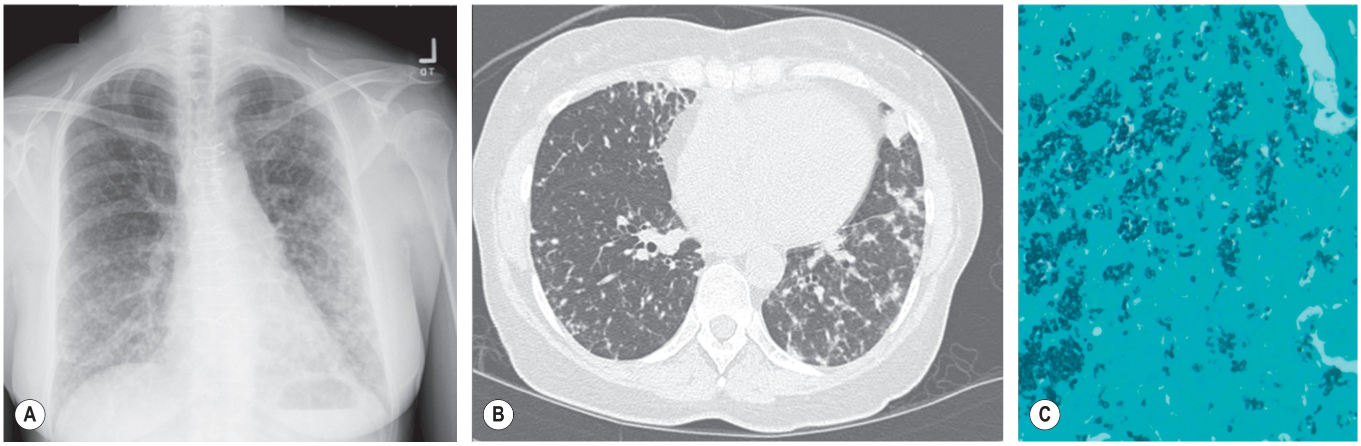


FIG 37.4 Disseminated Histoplasmosis in a Patient from an Endemic Area 3 Months After a Heart Transplantation. (A) Chest X-ray with bilateral reticulonodular infiltrates. (B) Computed tomography (CT) scan of the chest with bilateral patchy and nodular infiltrates. (C) Lung biopsy. Gomori methenamine silver (GMS) stain shows yeast forms compatible with *Histoplasma capsulatum.*

CLINICAL PEARLS

Infections in Bone Marrow Transplant Recipients

- Infections 2–4 weeks posttransplantation are usually caused by profound neutropenia and damage to mucosal surfaces.
- Between the period of engraftment and posttransplant weeks 15–20, OIs predominate and are commonly associated with the development of acute graft-versus-host disease (GvHD) and its treatment.
- Serious infections occurring 4–6 months following transplantation are seen predominantly in patients with chronic GvHD.

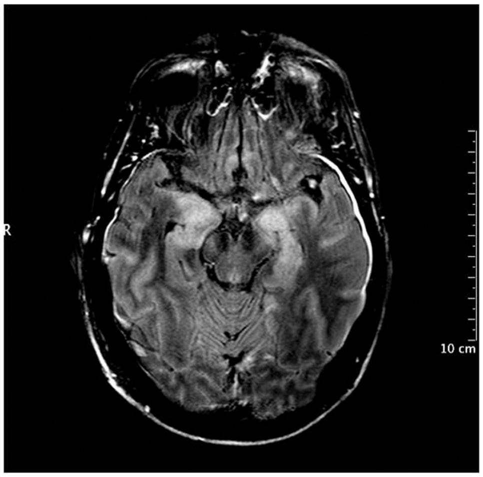


FIG 37.5 Magnetic resonance imaging (MRI) scan of the brain of a patient with human herpesvirus-6 (HHV-6) encephalitis 36 days after a matched unrelated donor hematopoietic stem cell transplantation (HSCT) for diffuse large B-cell lymphoma. Flair axial image shows bilateral medial temporal lobe edema compatible with limbic encephalitis.

The risk for infection is influenced by several factors and varies significantly between type of transplantation (autologous vs allogeneic), stem cell source (bone marrow, PBSC, and cord blood), degree of human leukocyte antigen (HLA) matching, hematopoietic potential of the graft (number of infused CD34$^+$ cells), T-cell depletion, type of conditioning (myeloablative vs reduced intensity vs nonmyeloablative), underlying disease, presence of graft rejection or GvHD, donor–recipient serological status to certain infections (*e.g.,* CMV infection), and use of antimicrobial prophylaxis.[47]

Autologous HSCT (auto-HSCT) refers to the patient serving as his or her own donor to allow administration of myeloablative chemotherapy. Because there is no risk of rejection or GvHD and because after engraftment there is an early and progressive recovery of cell-mediated immunity, the risk for infection is much lower than in allo-HSCT. The major compromises in host defenses occur before engraftment and are usually related to neutropenia and mucosal injury. In contrast, in allo-HSCT, the risk for infection can be divided into three periods: preengraftment, early postengraftment, and late postengraftment periods.

Preengraftment Period

The major impairments in host defenses before engraftment are neutropenia and mucosal injury. This period varies widely between different types of conditioning regimens and stem-cell source. In general, when ablative regimens are used, engraftment occurs any time between 15 and 45 days after transplantation and can be as short as 5–7 days with nonablative regimens.[47] PBSC grafts are associated with faster neutrophil engraftment compared with cord blood units and therefore lower rates of infection.[46] Bacterial infections are most common. However, when neutropenia is prolonged, the risk of invasive fungal infection increases significantly. Without acyclovir prophylaxis, reactivation of severe HSV infection is frequent.

Early Postengraftment Period

The resolution of neutropenia marks the beginning of this period and usually lasts 2–3 months (Fig. 37.5). This period is characterized by a progressive but slow recovery of B- and T-lymphocyte function; thus viral and fungal OIs can occur. The risk of infection is higher in patients who develop acute GvHD and therefore require high dose steroids. Common bacterial infections can still occur and are usually related to indwelling catheters or GvHD of the gut. CMV viremia is common during this period. However, routine use of preemptive or prophylactic therapy has been effective in preventing life-threatening CMV disease. The risk of *Pneumocystis* pneumonia (PCP) has also decreased since the introduction of PCP prophylaxis.[47]

Late Postengraftment Period

This is the period of late immune recovery and ends when the patient regains normal immunity. In the absence of chronic GvHD, this period can last up to 2 years. Viral and bacterial infections of the respiratory tract are common during this period. The development of chronic GvHD delays immune restoration and can extend this period for many years (as long as immunosuppressive drugs are required). Infections with encapsulated bacteria are common in patients with chronic GvHD, as well as opportunistic viral and fungal infections, including VZV and CMV infections, invasive aspergillosis and non-*Aspergillus* mold infections, and PJP, among others.[47]

INFECTIONS OF PARTICULAR IMPORTANCE IN TRANSPLANT RECIPIENTS

Cytomegalovirus Infection

CMV is one of the most important pathogens affecting transplant recipients. It is widely distributed in the general population; seropositivity for CMV varies in different geographical regions and ranges from 30% to 97%.[48] Before widespread use of interventions aimed at reducing the incidence of CMV disease (*i.e.,* pneumonitis, gastroenteritis, hepatitis, etc.), it occurred frequently during the first 3 months after transplantation. In SOT, the highest risk for CMV disease is in patients in whom the donor is seropositive and the recipient is seronegative for CMV (D^+/R^-). CMV has a predilection to invade the allograft, possibly as a result of an aberrant immune response within the graft. The risk of infection varies also with the type of transplantation; recipients of lung, small intestine, and pancreas transplants have a higher risk, whereas the risk in liver and kidney recipients is lower. In SOT, CMV infection has been found consistently to be an independent risk factor for other infectious complications as well as graft rejection.[48] In HSCT, infection is usually the result of reactivation of endogenous virus; hence the highest risk is in seropositive recipients. Without prophylaxis, up to 80% of seropositive recipients will experience CMV infection after HSCT. Seronegative individuals have a 30–40% chance of becoming infected when receiving unfiltered blood products or stem cells from a seropositive donor. Risk factors for CMV disease are acute and chronic GvHD, use of high-dose corticosteroids, use of cord blood, T-cell depletion, and use of mismatched or unrelated donors.[49]

Typically, CMV viremia precedes CMV disease by 1–2 weeks; therefore close monitoring of CMV reactivation by polymerase chain reaction (PCR) assays or pp65 antigenemia allows the detection of early replication, and therefore institution of appropriate antiviral therapy (ganciclovir, valganciclovir, or foscarnet) before the development of end organ disease. This approach, termed *preemptive therapy,* has the advantage of effectively decreasing the incidence of early CMV disease and limiting drug-related toxicities. In addition, a limited amount of viral replication may allow for the development of CMV-specific immune reconstitution. Preemptive therapy is the preferred method for preventing CMV disease in HSCT[49] and is widely used in SOT in patients with intermediate risk for CMV disease. In high-risk solid-organ recipients (D^+/R^-), in particular recipients of lung and small bowel, although preemptive therapy is an accepted alternative, universal prophylaxis is preferred. It has the theoretical advantage of preventing reactivation of other herpes viruses and may be more likely to prevent the indirect effects of CMV.[48,50] Recently, the development of commercially available CMV- specific IFN-γ release assays, which detect CMV-specific T cells in whole blood, may allow for more precise targeting of prophylactic strategies, beyond those currently used.[55]

Other Herpes Viruses

HSV infection is common in both SOT and HSCT recipients during the first month after transplantation. Routine use of antiviral prophylaxis has successfully decreased the incidence of severe HSV infection. However, breakthrough mucocutaneous HSV infection and even disseminated disease can still occur. Reactivation of VZV infection is also common; however, acyclovir prophylaxis has decreased the incidence of disseminated VZV infection significantly. Both severe VZV and HSV infections are typically treated with intravenous acyclovir.[45,46]

Although EBV can cause a wide spectrum of disease, EBV-associated PTLD is the most feared complication. The term *EBV-associated PTLD* is generally used to describe a heterogeneous group of clinical syndromes associated with uncontrolled lymphoproliferation, which can result in true malignancies containing clonal chromosomal abnormalities. Early diagnosis requires a high index of suspicion. EBV viral load monitoring and radiological evaluation can assist in early diagnosis. Reduction of immunosuppression should be the initial strategy in managing the disease. Timing of additional therapies, such as treatment with antivirals and rituximab and chemotherapy, remains controversial.[45,47]

HHV-6 has been associated with disease after transplantation, and although a variety of clinical manifestations have been described, the more convincing association is with encephalitis. The initial presentation may be subtle with memory loss or disorientation. However, it may progress to severe mental status abnormalities and seizures. The diagnosis is established with a positive PCR in the CSF, and absence of other infectious agents. Treatment is with ganciclovir or foscarnet.[47]

Invasive Filamentous Fungal Infections

Invasive pulmonary aspergillosis is the most common invasive filamentous fungal infection. Dissemination, including to the brain, can occur and should be ruled out in all transplant recipients. Risk factors are prolonged neutropenia, high-dose corticosteroids, cord blood transplantation, CMV disease, GvHD, graft rejection, and lung transplantation. Manifestations of filamentous fungi that are almost exclusively seen in lung transplant recipients are tracheobronchitis, characterized by ulceration and cartilage invasion, and bronchial anastomotic infections.[51]

Monitoring serum *Aspergillus* galactomannan in high-risk patients, as well as computed tomography (CT) imaging, can assist in the early diagnosis of invasive aspergillosis. Despite early diagnosis and treatment, the mortality associated with invasive aspergillosis, in particular in HSCT, is still high.[51] Antifungal prophylaxis should be considered in patients in whom prolonged neutropenia is expected, in patients with chronic GvHD receiving high-dose corticosteroids, and in lung transplant recipients. Voriconazole is considered the drug of choice for invasive aspergillosis, but newer-generation azoles, such as posaconazole and isavuconazole, are good alternatives.

Fusariosis, mucormycosis, and dematiaceous and other mold infections are increasingly recognized as significant pathogens in transplant recipients. Clinical disease can be similar to the infection caused by *Aspergillus* spp., with largely sinopulmonary disease. *Fusarium* spp. frequently disseminate hematogenously and can be isolated in blood cultures. Many of these infections are resistant to most antifungal drugs and remain an important cause of transplantation-related mortality.[51]

Invasive Candidiasis

Before routine use of antifungal prophylaxis, invasive candidiasis was the most common invasive fungal infection in transplant recipients. In SOT, most *Candida* infections occur during the first month after transplantation, usually related to the surgical procedure. Recipients of liver, pancreas, and small bowel are at high risk for invasive candidiasis; lung transplant recipients can develop bronchial anastomotic infections.[51] In HSCT, most *Candida* infections occur in the preengraftment period and are associated with mucosal injury, which results from the conditioning regimen and with the widespread use of antibiotics for the treatment of fever and neutropenia. Fluconazole-resistant *Candida* spp. are now increasingly isolated, and echinocandins (caspofungin, micafungin, and anidulafungin) are now considered the drugs of choice in invasive candidiasis.[52] Antifungal prophylaxis reduces the incidence of invasive fungal infections after HSCT, as well as infection-related and overall mortality. In SOT, antifungal prophylaxis is recommended for high-risk liver transplant recipients and in small bowel and lung transplant recipients.

TRANSLATIONAL RESEARCH

ON THE HORIZON

- Cytomegalovirus (CMV), Epstein-Barr virus (EBV), and adenovirus infections remain problematic infections in both hematopoietic-stem-cell and solid-organ transplant recipients, and medications currently available to treat these infections have significant associated toxicities or lack efficacy.
- Adoptive immunotherapy, in which virus specific donor-derived T cells are expanded and infused into the transplant recipient, will likely become more widespread in treating these infections.

Viral reactivation and disease are still a major cause of morbidity and mortality among transplant recipients. CMV, EBV, and adenovirus are particularly problematic once disease develops. Preemptive strategies have substantially decreased the incidence of invasive CMV disease; nevertheless, a significant number of patients still go on to develop invasive disease. Preemptive therapy and treatment are limited by the toxicity associated with the available antivirals and their lack of efficacy against EBV and adenovirus. In addition, development of resistance can complicate therapy. Development of new effective antivirals has been slow. Brincidofovir (CMX001) is an ether lipid ester analogue of cidofovir that has been reported to be 1000-fold more potent than cidofovir against DNA viruses, such as CMV and adenovirus. It is available in an oral formulation, but it lacks the nephrotoxic effects of cidofovir.[53] However, a recent trial in high-risk recipients in HSCT, in which brincidofovir was used to prevent CMV infection, did not achieve its primary endpoint for the prevention of clinically significant CMV infection after transplantation.[53] There are ongoing clinical trials to determine the efficacy and safety of brincidofovir for the treatment of other DNA viruses, such adenovirus and smallpox virus. Letermovir is a new, highly potent anti-CMV agent with a novel mechanism of action targeting the viral terminase subunit pUL56, a component of the terminase complex involved in viral DNA cleavage and packaging. Initial trials have shown that letermovir is effective in reducing the incidence of CMV infection in allo-HSCT and has been used successfully in MDR infections.[54] Phase III clinical trials are currently underway. There are other antivirals in the pipeline, but they are at much earlier stages of development. Because of the scarcity of effective, nontoxic antivirals and because uncontrolled viral infections correlate with a lack of cellular immunity against viruses, several groups are working on the development of adoptive immunotherapy, which is the artificial reconstitution of virus-specific T cells with *in vitro* expanded T cells.

A number of small studies have now shown that virus-specific T cells are effective in controlling CMV, EBV, and adenovirus infections in transplant recipients. The translation of adoptive immunotherapy into the clinic has been limited by technical difficulties associated with the generation of viral-specific cells that are not alloreactive, that can be produced from naïve T cells, and that can be developed rapidly and in a cost-effective way. Cells obtained from the original donor can yield long-lasting immunity, but the process of generating these cells can be cumbersome and often not available in a timely manner. Some groups have explored adoptive transfer of HLA partially matched virus-specific T cells derived from third-party donors, that is, healthy individuals other than the donor of the patient's transplant. Such cells yield high number of virus-specific T-cells, but these cells tend to be short lived.[55] Currently, a few international centers have established cryopreserved banks with virus-specific T-cell lines from normal donors, but these are still mostly for investigational use. Significant advances continue to be seen in the field of cellular therapy, and these could become part of the armamentarium against viruses in the next decade.

CONCLUSIONS

Immunocompromised hosts are surviving longer as antimicrobials improve and as immunosuppressant agents become less toxic. Both monogenetic primary immunodeficiencies and the increasingly specific immunosuppressant agents allow insight into host control of specific microbes. With this knowledge, new types of immunodeficiencies are being recognized, such as the autoantibody cytokine disorders described above. In addition, the knowledge of specific host immunity will help in developing therapies to boost host responses in the increasing challenges of treatment-resistant microbes.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Holland SM. Chronic granulomatous disease. Clin Rev Allergy Immunol 2010;38:3–10.
2. Rosenzweig SD, Holland SM. Phagocyte immunodeficiencies and their infections. J Allergy Clin Immunol 2004;113:620–6.
3. Greenberg DE, Ding L, Zelazny AM, et al. A novel bacterium associated with lymphadenitis in a patient with chronic granulomatous disease. PLoS Pathog 2006;2:e28.
4. Siddiqui S, Anderson VL, Hilligoss DM, et al. Fulminant mulch pneumonitis: an emergency presentation of chronic granulomatous disease. Clin Infect Dis 2007;45:673–81.

5. Freeman AF, Marciano BE, Anderson VL, et al. Corticosteroids in the treatment of severe Nocardia pneumonia in chronic granulomatous disease. Pediatr Infect Dis J 2011;30:806–8.
6. Leiding JW, Freeman AF, Marciano BE, et al. Corticosteroid therapy for liver abscess in chronic granulomatous disease. Clin Infect Dis 2012;54:694–700.
7. Freeman AF, Holland SM. Antimicrobial prophylaxis for primary immunodeficiencies. Curr Opin Allergy Clin Immunol 2009;9:525–30.
8. Rondeau S, Couderc L, Dominique S, et al. High frequency of voriconazole-related phototoxicity in cystic fibrosis patients. Eur Respir J 2012;39:782–4.
9. Tarlock K, Johnson D, Cornell C, et al. Elevated fluoride levels and periostitits in pediatric hematopoietic stem cell transplant recipients receiving long-term voriconazole. Pediatr Blood Cancer 2015;62:918–20.
10. Boztug K, Klein C. Genetic etiologies of severe congenital neutropenia. Curr Opin Pediatr 2011;23:21–6.
11. Hsu AP, McReynolds LJ, Holland SM. GATA2 deficiency. Curr Opin Allergy Clin Immunol 2015;15:104–9.
12. Hambleton S, Salem S, Bustamante J, et al. IRF8 mutations and human dendritic-cell immunodeficiency. N Engl J Med 2011;365:127–38.
13. Conley ME, Dobbs AK, Farmer DM, et al. Primary B cell immunodeficiencies: comparisons and contrasts. Annu Rev Immunol 2009;27:199–227.
14. Cuccherini B, Chua K, Gill V, et al. Bacteremia and skin/bone infections in two patients with X-linked agammaglobulinemia caused by an unusual organism related to Flexispira/Helicobacter species. Clin Immunol 2000;97:121–9.
15. Murray PR, Jain A, Uzel G, et al. Pyoderma gangrenosum-like ulcer in a patient with X-linked agammaglobulinemia: identification of *Helicobacter bilis* by mass spectrometry analysis. Arch Dermatol 2010;146:523–6.
16. Van der Burg M, Gennery AR. Educational paper: the expanding clinical and immunological spectrum of severe combined immunodeficiency. Eur J Pediatr 2011;170:561–71.
17. Kobrynski LI, Sullivan KE. Velocardiofacial syndrome, DiGeorge syndrome: the chromosome 22q11.2 deletion syndromes. Lancet 2007;1443–52.
18. Perez EE, Bokszczanin A, McDonald-McGinn D. Safety of the live viral vaccines in patients with chromosome 22q11.2 deletion syndrome (DiGeorge/velocardiofacial syndrome). Pediatrics 2003;112:e325.
19. Milner JD, Brenchley JM, Laurence A, et al. Impaired T(H) 17 cell differentiation in subjects with autosomal dominant hyper-IgE syndrome. Nature 2008;452:773–6.
20. Freeman AF, Holland SM. Clinical manifestations, etiology, and pathogenesis of the hyper-IgE syndromes. Pediatr Res 2009;65:32R–37R.
21. Odio CD, Milligan KL, McGowan K, et al. Endemic mycosis in patients with STAT3-mutated hyper IgE (Job) syndrome. J Allergy Clin Immunol 2015;136:1411–13.
22. Zhang Q, Davis J, Lamborn IT, et al. Combined immunodeficiency associated with DOCK8 mutations. N Engl J Med 2009;361:2046–55.
23. Al-Muhsen S, Casanova JL. The genetic heterogeneity of mendelian susceptibility to mycobacterial diseases. J Allergy Clin Immunol 2008;122:1043–51.
24. Holland SM. Interferon gamma, IL-12, IL-12R and STAT1 immunodeficiency diseases: disorders of the interface of innate and adaptive immunity. Immunol Res 2007;38:342–6.
25. Toubiana J, Okada S, Hiller J, et al. Heterozygous STAT1 gain-of-function mutations underlie an unexpectedly broad clinical phenotype: an international survey of 274 patients from 167 kindreds. Blood 2016;127:3154–64.
26. Ram S, Lewis LA, Rice PA. Infections of people with complement deficiencies and patients who have undergone splenectomy. Clin Microbiol Rev 2010;23:740–80.
27. Browne SK, Holland SM. Anticytokine autoantibodies in infectious diseases: pathogenesis and mechanisms. Lancet Infect Dis 2010;10:875–85.
28. Kisand K, Boe Wolff AS, Podkrajsek KT. Chronic mucocutaneous candidiasis in APECED or thymoma patients correlate with autoimmunity to Th17 associated cytokines. J Exp Med 2010;207:299–308.
29. Puel A, Doffinger R, Natividad A, et al. Autoantibodies against IL-17A, IL-17F, and Il-22 in patients with chronic mucocutaneous candidiasis and autoimmune polyendocrine syndrome type 1. J Exp Med 2010;207:291–7.
30. Rosen LB, Freeman AF, Yang LM, et al. Anti-GM-CSF autoantibodies in patients with cryptococcal meningitis. J Immunol 2013;190:3959–66.
31. Allison AC. Immunosuppressive drugs: the first 50 years and a glance forward. Immunopharmacology 2000;47:63–83.
32. Mueller NJ. New immunosuppressive strategies and the risk of infection. Transpl Infect Dis 2008;10:379–84.
33. Halloran PF. Immunosuppressive drugs for kidney transplantation. N Engl J Med 2004;351:2715–29.
34. Freifeld AG, Bow EJ, Sepkowitz KA, et al. Clinical practice guideline for the use of antimicrobial agents in neutropenic patients with cancer: 2010 update by the Infectious Diseases Society of America. Clin Infect Dis 2011;52:e56–93.
35. Kang I, Park SH. Infectious complications in SLE after immunosuppressive therapies. Curr Opin Rheumatol 2003;15:528–34.
36. Gaber AO, Knight RJ, Patel S, et al. A review of the evidence for use of thymoglobulin induction in renal transplantation. Transplant Proc 2010;42:1395–400.
37. Salvana EM, Salata RA. Infectious complications associated with monoclonal antibodies and related small molecules. Clin Microbiol Rev 2009;22:274–90.
38. Koo S, Marty FM, Baden LR. Infectious complications associated with immunomodulating biologic agents. Infect Dis Clin North Am 2010;24:285–306.
39. Vergidis P, Avery RK, Wheat LJ, et al. Histoplasmosis complicating tumor necrosis factor–α blocker therapy: a retrospective analysis of 98 cases. Clin Infect Dis 2015;61:409–17.
40. Gea-Banacloche JC. Rituximab-associated infections. Semin Hematol 2010;47:187–98.
41. www.accessdata.fda.gov/drugsatfda_docs/label/2012/103705s5367s5388lbl.pdf.
42. Dale RC, Brilot F, Duffy LV, et al. Utility and safety of rituximab in pediatric autoimmune and inflammatory CNS disease. Neurology 2014;83:142–50.
43. Bloomgren G, Richman S, Hotermans S, et al. Risk of natalizumab-associated progressive multifocal leukoencephalopathy. N Engl J Med 2012;366:1870–80.
44. Piperdi B, Ling Y, Liebes L, et al. Bortezomib: understanding the mechanism of action. Mol Cancer Ther 2011;10:2029–30.
45. Fishman JA, Issa NC. Infection in organ transplantation: risk factors and evolving patterns of infection. Infect Dis Clin North Am 2010;24:273–83.
46. Gooley TA, Chien JW, Pergam SA, et al. Reduced mortality after allogeneic hematopoietic-cell transplantation. N Engl J Med 2010;363:2091–101.
47. Wingard JR, Hsu J, Hiemenz JW. Hematopoietic stem cell transplantation: an overview of infection risks and epidemiology. Infect Dis Clin North Am 2010;24:257–72.
48. Humar A, Snydman D. Cytomegalovirus in solid organ transplant recipients. Am J Transplant 2009;9:S78–86.
49. Ljungman P, Hakki M, Boeckh M. Cytomegalovirus in hematopoietic stem cell transplant recipients. Infect Dis Clin North Am 2010;24:319–37.
50. Kotton CN, Kumar D, Caliendo AM, et al. Updated International Consensus Guidelines on the management of cytomegalovirus in solid-organ transplantation. Transplantation 2013;96:333–60.
51. Nucci M, Anaissie E. Fungal infections in hematopoietic stem cell transplantation and solid-organ transplantation—focus on aspergillosis. Clin Chest Med 2009;30:295–306, vii.
52. Pappas PG, Kauffman CA, Andes DR, et al. Clinical practice guideline for the management of candidiasis: 2016 update by the Infectious Diseases Society of America. Clin Infect Dis 2016;62:e1–50.
53. Marty FM, Winston DJ, Rowley SD, et al. CMX001 to prevent cytomegalovirus disease in hematopoietic-cell transplantation. N Engl J Med 2013;369:1227–36.
54. Chemaly RF, Ullmann AJ, Stoelben S. Letermovir for cytomegalovirus prophylaxis in hematopoietic-cell transplantation. N Engl J Med 2014;370:1781–9.
55. O'Reilly RJ, Prockop S, Hasan AN, et al. Virus-specific T-cell banks for "off the shelf" adoptive therapy of refractory infections. Bone Marrow Transplant 2016;51:1163–72.

MULTIPLE CHOICE QUESTIONS

1. Mucocutaneous candidiasis is thought to occur due to disruption of interleukin-17 (IL-17) signaling. This is observed in all of the following immunodeficiencies EXCEPT:
   - **A.** Loss of function mutations in signal transducer and activator of transcription 3 (STAT3) causing hyper IgE syndrome (HIES)
   - **B.** CD40 ligand deficiency
   - **C.** Gain-of-function (GOF) STAT1 mutations
   - **D.** Autoimmune polyendocrinopathy with candidiasis and ectodermal dysplasia (APOCED)

2. Anticytokine autoantibodies are increasingly being recognized as causing immunodeficiency. The following are described associations EXCEPT:
   - **A.** Anti-interferon (IFN)-γ antibodies and disseminated nontuberculous mycobacteria
   - **B.** Anti–GM-CSF antibodies and Cryptococcus
   - **C.** Anti-IL-17 antibodies and mucocutaneous Candida
   - **D.** Anti-IFNγ antibodies and *Staphylococcus aureus*
   - **E.** Anti–granulocyte macrophage–colony-stimulating factor (GM-CSF) antibodies and *Nocardia*

3. Progressive multifocal leukoencephalopathy has been described in the following settings EXCEPT:
   - **A.** Rituximab
   - **B.** GOF STAT1 mutations
   - **C.** Natalizumab
   - **D.** Chronic granulomatous disease (CGD)
   - **E.** Human immunodeficiency virus (HIV)
   - **F.** Dedicator of cytokinesis 8 (DOCK8) deficiency

4. Photosensitivity is associated with which antifungal?
   - **A.** Fluconazole
   - **B.** Caspofungin
   - **C.** Posaconazole
   - **D.** Voriconazole
   - **E.** Isavuconazole

38

Immune Deficiencies at the Extremes of Age

Jörg J. Goronzy, Claire E. Gustafson, Cornelia M. Weyand

Age is a major factor determining the quality and quantity of an immune response. This age dependency, apparent throughout life, is most obvious at the extremes of age. Both infants and older adults exhibit blunted immune responses to infections and vaccination. However, the underlying mechanisms for immune dysfunction are distinct. Infants are born with limited antigen exposure and an immature immune system, which, however, progressively develops throughout infancy and childhood. The mechanisms underlying the immune dysfunction in older adults are broadly termed *immunosenescence.* The term is reminiscent of *cellular senescence,* which is defined as the loss of the ability to proliferate as a result of an irreversible cell cycle block. However, it is misleading to interpret immune aging only as the accumulation of senescent cells, and similar to aging in general, numerous pathways are involved.[1] The decline in immune competence is not linear. As early as age 40 years, vaccine responses to selected vaccines (*e.g.*, hepatitis B) start to decline. The incidence rate of herpes zoster, a reactivation of latent varicella-zoster virus (VZV), starts to increase at age 50 years, as does morbidity and mortality from influenza infection. A more abrupt transition appears to occur in the eighth decade of life.

In terms of public health, infections are major causes of morbidity in the very young and the very old. Although child mortality rates have dropped by almost 50% between 1990 and 2013, pathogenic infections are still one of the largest causes of infant mortality worldwide, accounting for more than 1.6 million deaths a year. Vaccinations have helped change the infectious landscape in children and young adults, but certain vaccines, such as those against *Streptococcus pneumoniae,* still have limited protective capacity at both extremes of age. Susceptibility to pathogenic infection and ineffectiveness of vaccination is even greater in older adults than in young infants. Moreover, during aging, a functional immune system is important for tissue repair and is of increasing importance in degenerative diseases, and it is vital for cancer surveillance. Immunosenescence is of increasing importance because of the changing population demographics, with increases in the number of individuals older than 65 years in both developed and developing countries.

INFANCY AND THE GENERATION OF AN IMMUNE SYSTEM

Infants first start to develop their immune system *in utero,* maintaining a fine balance between immunological tolerance that helps prevent proinflammatory responses *in utero* and the ability to respond to foreign antigen exposure upon birth. Immunosuppressive regulatory pathways utilized by the fetus for protection against possible infections and maternal–fetal rejection *in utero* are still reflected in the newborn infant immune system, characterized by immune-suppressive, antiinflammatory, and immature cellular responses to foreign antigens. Although newborns and young infants lack the ability to mount effective immune responses against many pathogens, they acquire initial immune protection through the passive transfer of maternal immunoglobulin G (IgG) *in utero* via the placenta (termed "passive immunity")[2] and from secretory IgA and antimicrobial factors present in maternal breast milk. Alternations in passive immunity (*e.g.*, preterm birth) can lead to increased susceptibility to infections and breakdown of immune tolerance in the infant. Passive immunity can be enhanced by vaccination of pregnant mothers to promote the transfer of vaccine-specific IgG *in utero.* However, within the first months of life, as maternal IgG wanes and breastfeeding is discontinued, infants must actively develop their own innate and adaptive immune responses to initiate and maintain immune protection. An overview of the development of innate and adaptive cells, as well as the repertoire of diversity during aging, is outlined in Fig. 38.1.

INNATE IMMUNE DEVELOPMENT

The innate immune system is typically considered the first line of defense against infection. The development of the innate immune system begins *in utero,* with all the classic innate cells types present by the end of the first trimester; monocytes and dendritic cells (DCs) are the earliest cells observed at 4 weeks' gestation (WG) followed by granulocytes and natural killer (NK) cells at 8 WG.[3] These innate immune cells expand in number and mature during gestation; however, at birth the functionality of innate immune cells is still diminished significantly compared with later in life.

DYSFUNCTION OF INNATE IMMUNE CELLS

During early infancy, all cell types of the innate immune system demonstrate some impairment of function, ranging from reduced mobility to skewed cytokine production in response to innate immune stimulation. The most significant functional limitation of neonatal innate immune cells is their collective inability to kill pathogens. Limited chemotaxis in neutrophils is accompanied by a reduced ability to effectively kill pathogens as a result of poor phagocytosis and decreased secretion of neutrophil extracellular traps under inflammatory conditions, which in combination results in decreased killing of infectious pathogens. Moreover, NK cells and plasmacytoid DCs have reduced capacity to prevent

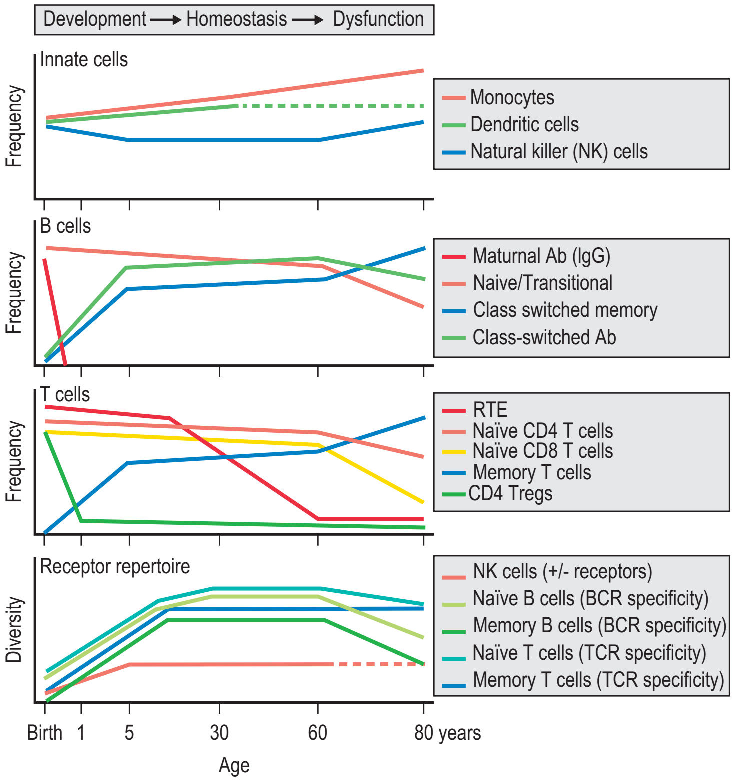


FIG 38.1 Innate and Adaptive Immune Cell Frequencies and Receptor Repertoire Diversity During Aging. The frequencies of different innate and adaptive B and T cells changes over the course of a lifetime (upper three panels). In addition to compartment size, diversity of antigen receptors is essential for immune function. During infancy, diversity of immune populations expands and is then maintained during adulthood (lower panel). Aging is characterized by the general decline in diversity within naïve B- and T-cell populations. Dotted lines indicate that the progression of these populations is either controversial or unknown.

infection through reduced cytotoxic function and lower secretion of interferon- (IFN-α), respectively. In addition, antigen-presenting cells (APCs) are unable to provide effective help to T cells because of reduced expression of costimulatory molecules and skewed cytokine production toward more antiinflammatory, T-helper 2 (Th2)–inducing cytokines, such as interleukin-6 (IL-6). These combined limitations during early infancy lead to increased susceptibility to viral and bacterial infections and also contribute to the reduced functionality of adaptive immune cells.

KEY CONCEPTS

Characteristics of the Developing Innate Immune System

- Reduced antibacterial responses (*i.e.*, phagocytosis, secretion of neutrophil extracellular traps [NETs]) by neutrophils and monocytes
- Impaired neutrophil chemokinesis ("directed movement")
- Diminished capacity of antigen-presenting cells (APCs; *i.e.*, dendritic cells [DCs]) to provide proper costimulatory help to T cells
- Reduced antiviral responses (*i.e.*, interferon (IFN)-α secretion) by plasmacytoid dendritic cells (pDCs)

ADAPTIVE IMMUNE DEVELOPMENT

Adaptive immune cells (T and B cells) begin to develop *in utero* around the start of the second trimester. It is important to note that although lymphocytes initially develop within the fetal liver (≈7 WG) and then in bone marrow (≈12 WG), T-cell maturation occurs in the thymus, and B cells continue maturation in bone marrow *in utero* and after birth. Over the course of gestation, both naïve T- cell and B-cell compartments expand, reaching absolute cell concentrations greater than that of the adults by birth. Thus a defect in cellular generation cannot account for the immune limitations observed in infants. However, the composition of the each lymphocyte population within young infants is distinct from that of adults. At birth, both the newborn infant T- and B-cell compartments consist mainly of naïve and immature cells recently migrated out of the thymus or bone marrow (>90% of the population). Both infant T- and B-cell repertoires have less diversity within their receptors' antigen-binding regions compared with those of adults. T-cell receptors display reduced V-J complexity and fewer amino acid additions. In addition, B cells have decreased B-cell receptor diversity because of less affinity maturation (*i.e.*, somatic hypermutation; Chapter 7) compared with adult populations.

The increased levels of recent thymic emigrants and naïve T cells are maintained both in the periphery and in tissues during early infancy. Effector memory T cells can be found selectively in the lungs and gastrointestinal (GI) tract of young infants, likely because these mucosal tissues are the earliest sites of antigen exposure. Class-switch antibodies produced by B cells are required for mucosal protection (primarily IgA) and systemic protection (primarily IgG) through opsonization, neutralization, or antibody-dependent cellular cytotoxicity (ADCC). However, infants demonstrate significantly lower levels of IgG and IgA compared with adults throughout the first year of life, in both the periphery and mucosae.

INFANCY AND FUNCTIONAL DIFFERENTIATION OF ADAPTIVE IMMUNE CELLS

The acquisition of a mature adaptive immune system requires exposure to antigens for cells to form immunological memory. Not only are infants born with limited antigen exposure and thus little memory, infant responses to antigenic stimulation in the T- and B-cell compartments display altered functions compared with that of adults. The most striking alterations are seen within the CD4 T-cell compartment, characterized by significantly increased frequency of immunosuppressive T regulatory cells (Tregs; 30–40% in infants vs 1–10% in adults).[4] These changes accompany increased Th2 subsets and defective differentiation to effector Th1 and T-follicular helper (Tfh) cells during early infancy. Moreover, decreased expression of Lck in infant CD4 T cells leads to decreased T-cell activation upon antigen exposure.[2] Altered differentiation of CD4 Th subsets and blunted T-cell activation promotes a tolerance-inducing environment good for controlling unnecessary responses to foreign antigen exposure after birth. However, it also leads to inhibition of effective cytotoxic T cells and limits B-cell antibody responses against infectious pathogens and to vaccination (Fig. 38.2).

The reduced antibody titers observed during infancy can be accounted for by the functional inability of infant B cells to undergo class-switching from IgM to IgA or IgG upon antigen stimulation[5] and the poor survival of antibody-secreting plasma cells within the infant bone marrow. Moreover, infant B cells have lower rates of somatic hypermutation compared with those of adults, which is important for affinity maturation and the generation of high-affinity antibodies. Although these defects can be partially attributed to extrinsic T-cell defects, such as poor Tfh generation, intrinsic defects in infant B cells (*e.g.*, lack of costimulatory molecules and receptors to T cell–independent factors) also prevent effective antibody responses. Collectively, these neonatal B-cell defects lead to less effective and shorter-lived antibody responses against infectious pathogens and vaccines during early infancy.

KEY CONCEPTS

Characteristics of the Developing Adaptive Immune System

- High frequencies of naïve and immature T and B cells
- Skewing of CD4 T-cell population towards immunosuppressive regulatory T cells and T-helper 2 (Th2) cells
- Reduced functionality and memory responses of CD8 T cells to viral infections
- Limited ability of B cells to produce high-affinity, class-switched antibodies in response to both T cell–dependent and T cell–independent antigens

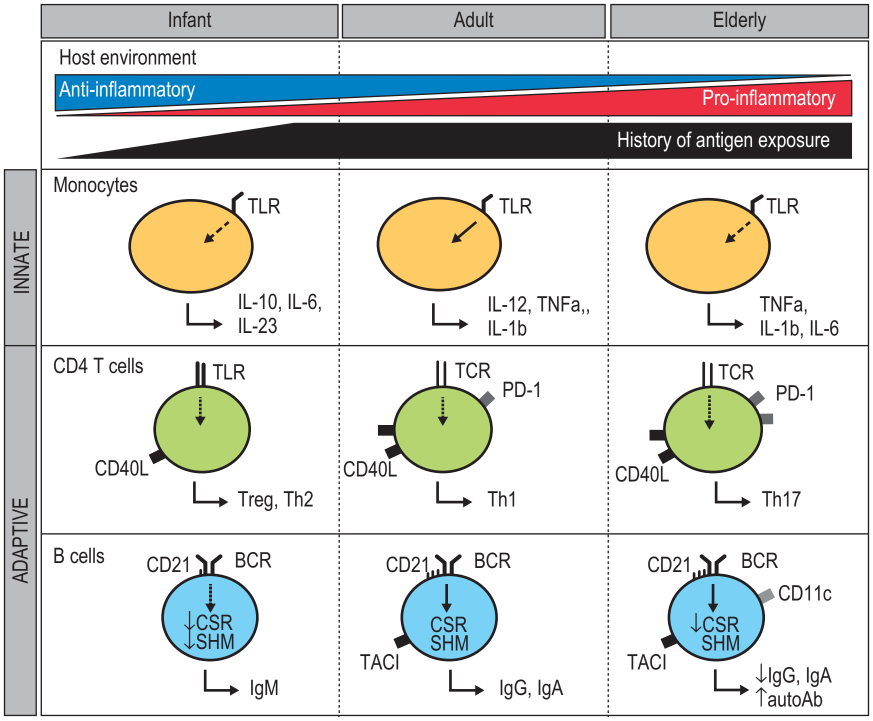


FIG 38.2 **Innate and Adaptive Immune Cell Functionality During Aging.** The infant immune environment is distinctly antiinflammatory characterized by altered Toll-like receptor (TLR) signaling from monocytes, blunted T-cell and B-cell receptor (TCR/BCR) signaling, and skewing toward T regulatory cells (Tregs). Infant B cells are incapable of class-switch recombination (CSR) or affinity maturation via somatic hypermutation (SHM) upon antigen exposure. Altered monocyte TLR signaling and altered function of adaptive immune cells is also observed in older adults, but with a distinct, more proinflammatory outcome.

INFANT IMMUNE DEVELOPMENT AND THE MICROBIOME

One of the first environmental exposures that infants experience is with colonizing bacteria. Within hours after birth, the infant intestinal tract is colonized with nonpathogenic bacteria (termed "microbiome"; Chapter 14). The microbiome can reach concentrations of 10^{10} bacteria per gram of stool by age 7 days, but only mirrors an adult-like composition after age 2 years. Early studies in animal models demonstrated that microbial colonization is essential for the normal development of the immune system during early infancy and throughout life.[6] In particular, IgA, the most abundant antibody within the GI tract, requires the presence of colonizing bacterial for its development and long-term maintenance. Moreover, the specific composition of the microbiome influences the generation of innate lymphoid cells, CD4 T-cells subsets (*e.g.,* Treg, Th17, Tfh), and antibody development. Perturbations in the infant microbiome, caused by such factors as mode of delivery, antibiotic treatment, and diet, may significantly alter developing immune responses.

It has been described that babies delivered via C-section have a microbiome composition distinct from vaginally delivered babies, the latter dominated by skin bacteria (e.g., Staphylococcus) instead of vaginal bacteria (e.g., Lactobacillus, Prevotella). These differences correlated with an increased incidence of allergy and asthma during childhood suggesting a causal link between microbiome and susceptibility to allergies. These findings are preliminary and more evidence is needed to support the notion that the microbiome influences the development of the infant immune system.

CLINICAL CONSEQUENCES FOR CHILDHOOD VACCINATION

Newborn and young infants are highly susceptible to pathogenic infection during the first months to year of life. Indeed, infections are one of the major causes mortality, accounting for more than 24% of global infant mortality. More than 1 million deaths per year are caused by respiratory tract infections (*e.g., Streptococcus pneumoniae, Haemophilus influenzae* type b, respiratory syncytial virus [RSV]) and almost the same amount by intestinal infections (*e.g.,* rotavirus, *Escherichia coli*). The development and distribution of vaccines targeting these pathogens have significantly reduced infant infection rates and subsequent mortality; however, vaccine efficacy can vary, depending on the infant's gestational age at birth, mode of vaccine delivery (oral vs intramuscular), interference by maternal antibodies acquired through passive immunity and via breast milk, and immaturity of infant immune responses, as described above. Additionally, infants require multiple vaccine boosters to elicit and maintain robust protective immunity against infectious pathogens (Chapter 90). Better understanding of the limitations of the human infant immune system that prevent effective immune responses, particularly at mucosal sites where these pathogens initially infect, would facilitate the generation of better vaccines designed to overcome these limitations and induce more rapid and robust immune protection during early infancy.

CLINICAL RELEVANCE

Consequences of Immaturity of the Developing Immune System

- Increased morbidity and mortality from bacterial infections (*e.g.,* pneumococci)
- Increased morbidity and mortality from viral infection (*e.g.,* influenza virus, respiratory syncytial virus [RSV])
- Ineffective primary vaccinations; multiple boosters required for most vaccines to elicit effective protection
- Susceptibility to allergy and asthma when microbiome colonization disrupted during early infancy

OLDER AGE AND IMMUNE CELL GENERATION

The immune system is in constant demand for cellular replenishment to compensate for peripheral losses and cell death. For neutrophils, which have a short half-life, the body needs to produce $\approx 10^{10}$ cells/kg per day. Additionally, for lymphocytes, which are more long-lived but are also more numerous because of their wider tissue distribution, the daily need is in the order of several billion cells. Studies in the 1960s showed that the hematopoietic tissue in bone marrow decreases with age. A similar decline is seen for the frequencies of peripheral CD34$^+$CD45$^+$ hematopoietic stem cells (HSCs). In addition to these numerical differences, the functional potential of HSCs changes with age.[7] Old HSCs exhibit a reduced capacity to regenerate the hematopoietic system. Telomerase expression in peripheral HSCs is not fully protective, and a shortening in telomeric length is seen with age, similar to differentiated mononuclear cells. Lymphoid differentiation potential diminishes with age in favor of myeloid differentiation as a result of altered lineage specification programs driven by DNA damage and epigenetic changes. This shift may contribute to the clinical observation that HSC-derived leukemia preferentially has a lymphoid phenotype in the young and a myeloid one in older adults.

KEY CONCEPTS

Aging Influences Immune Cell Generation and Population Homeostasis

Aging influences immune cell generation and population homeostasis.

- Hematopoietic stem cells (HSCs) are reduced in frequency and biased toward myeloid lineages and against lymphoid lineages.
- Myeloid cell generation is largely intact.
- B-cell generation declines and B-cell repertoire selection is disturbed.
- As a result of thymic involution, naïve T cells are generated from homeostatic proliferation of peripheral T cells.
- Ability to rebuild a T-cell repertoire after lymphocyte-depleting interventions is severely compromised after mid-adulthood.
- Virus-specific effector T-cell population accumulates.

As a consequence, lymphocyte replenishment is more affected by age than that of myeloid lineages. Peripheral neutrophil numbers do not decline and are able to recover after therapy-induced depletion (*e.g.,* chemotherapy). Furthermore, there is no loss in the ability to generate robust neutrophilia in response to infection or other stressors.

In addition to HSC-intrinsic alterations in lineage commitments, defects in bloodborne factors and bone marrow niches contribute to a decline in B-cell generation.[8] As a consequence, recovery after B-cell depletion (*e.g.,* treatment with anti-CD20 antibodies to treat B-cell malignancies or autoimmune diseases) is incomplete and delayed in older adults. Moreover, the frequency of typical memory B cells expressing CD27 declines, whereas other B-cell subpopulations increase. In particular, a functionally distinct B-cell population that responds to Toll-like receptor (TLR) stimulation has been described.[9] These age-associated differences in B-cell selection and expansion may contribute to monoclonal gammopathies (in >5% of individuals older than 70 years; Chapter 80) or increased autoantibodies with increasing age. High-throughput sequencing of Ig genes has the promise to provide insights into the age-associated decline in B-cell responses at the single-cell level. Data so far have indicated that older individuals have prevaccination-expanded B-cell populations that have a high mutation load and are further expanded with vaccination, suggesting that the B-cell response in older individuals relies on the adaptation of a restricted repertoire of memory sequences.

T-cell generation is more affected by age than any other myeloid or lymphoid lineage because of the involution of the thymus. The thymus undergoes dramatic structural changes that begin during childhood and puberty.[10] Thymopoietic niches disappear, and the numbers of thymic epithelial cells and thymocytes decline. In parallel, the thymic perivascular space increases. One major structural change that does not appear to be a compensatory response but, rather, is an active and regulated developmental step in thymic organ transformation is the accumulation and infiltration of adipocytes in the perivascular space. Thymic resection in children undergoing cardiac surgery reproduces many of the T-cell repertoire changes in 20-year-old individuals that are usually seen in 70- to 80-year-olds, confirming that thymic production during the growth period of childhood and adolescence is important. However, throughout adulthood, homeostatic proliferation of naïve T cells accounts for the bulk of T-cell generation.[11] For human CD4 T cells, this process is robust, and frequencies of naïve CD4 T cells only moderately decline with age. In contrast, the naïve CD8 T-cell compartment clearly shrinks. However, repertoire diversity (*i.e.,* the number of different T-cell receptors [TCRs]) remains very high for both naïve CD4 and CD8 T-cell subsets, suggesting that thymic activity is not necessary to prevent holes in the repertoire once a repertoire has formed. Strategies to reactivate thymic T-cell generation will be more important in older patients who cannot restore the naïve T-cell repertoire after a medical intervention, such as chemotherapy or bone marrow transplantation, compared with their healthy counterparts.

T-CELL POPULATION HOMEOSTASIS

The adaptive system responds to antigenic challenges with clonal expansion and differentiation into effector cells followed by clonal downsizing and persistence of long-lived memory T cells. Infections therefore leave a permanent imprint on the immune system, a mechanism on which vaccinations capitalize. However, the pathogen-induced clonal expansion also represents a challenge to homeostatic mechanisms that are supposed to maintain a balance among naïve, memory, and effector cells.[12] This is particularly evident in persisting infections where the offending pathogen cannot be cleared. Herpes virus infections are highly prevalent in an apparently healthy population without causing active disease, as they establish latency. Classic examples are varicella-zoster virus (VZV), Epstein-Barr virus (EBV), and cytomegalovirus (CMV). The effects of these herpes viruses with immune aging differ greatly; VZV tends to relapse with age, with reactivation presenting as shingles. A decrease in the frequency of VZV-specific CD4 memory T cells has been postulated to explain this lack in viral control mechanisms. In contrast, EBV and CMV infections only relapse in severely immunocompromised individuals, but not during normal immune aging. The immune system commits extraordinary resources to controlling CMV, and CMV-specific CD8 T cells can make up a large fraction of the entire T-cell repertoire. Whether this memory inflation has broader implications for immune health remains a matter of controversy. Expansion of the CMV-specific T cells may compromise the size of naïve and central memory T cell repertoires; however, many of the CMV-specific CD8 T cells have the phenotype of end-differentiated effector T cells that lack the expression of CCR7, CD28, and CD27 and therefore do not compete for the same space as naïve cells.

$CD45RA^+CD28^-$ end-differentiated T cells express negative regulatory receptors of the killer lectin-like receptor (KLR), killer immunoglobulin-like receptor (KIR), and immunoglobulin-like transcript (ILT) families that appear to constrain their otherwise unopposed expansion. In spite of these inhibitory receptors, these cells are competent effector T cells capable of producing inflammatory cytokines and therefore possibly contributing to inflammation in older adults. End-differentiated $CD45RA^+$ effector T cells should be distinguished from exhausted CD8 T cells.[13] T-cell exhaustion is seen with chronic stimulation by highly replicating viruses or tumor cells and characterized by the expression of inhibitory receptors programmed death 1 (PD-1), T-cell immunoglobulin and mucin-domain containing-3 (TIM-3), and lymphocyte activation gene 3 (LAG3).[14] T-cell exhaustion is not a general feature *per se* of T-cell aging.

INFLAMMATION, AGING, AND THE AGING HOST ENVIRONMENT

The aging host environment is characterized by the continuous presence of inflammatory mediators independent of acute or chronic disease (Fig. 38.3).[15] Even for the healthy older adult, IL-6 and tumor necrosis factor (TNF) serum levels are twofold to fourfold higher than in young adults. Low-level systemic inflammation plays an important role in the progression of several age-related diseases, including Alzheimer disease, atherosclerosis, and cancer. Moreover, inflammatory markers are associated with several conditions that are characteristic of older adults. IL-6 serum concentrations have been correlated with loss of mobility and advent of disability; increased mortality of older individuals has been shown among those who have higher levels of TNF-α. Increased IL-6 and cross-reactive protein (CRP) serum levels and increased white blood cell (WBC) counts, presumably resulting from increased production of neutrophils, predispose to, and are associated with, frailty. A causative relationship may also exist between the increased production of IL-6 or TNF-α and the age-associated loss in muscle mass, eventually presenting as sarcopenia. Long-lived individuals, such as centenarians, tend to have lower levels of proinflammatory cytokines and increased levels of antiinflammatory mediators, such as cortisone and IL-10, supporting the concept that low-level inflammation is detrimental to healthy aging.

Production of inflammatory cytokines is driven by several mechanisms. Failure of the adaptive immune system leads to a less effective control of chronic viral infections as well as incomplete response to exogenous challenges, resulting in increased and prolonged innate immune activation. Defective epithelial barrier function, as well as decline in the mucosa-associated

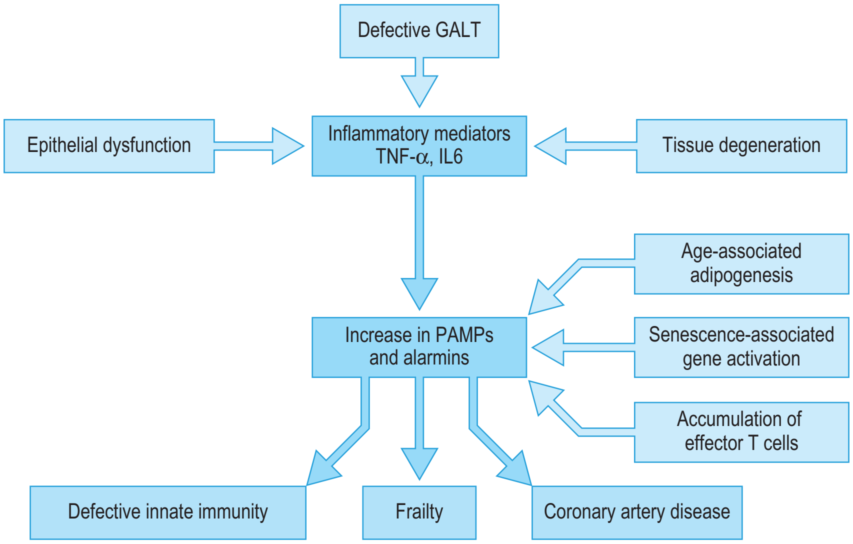

Defective GALT
Epithelial dysfunction
Inflammatory mediators TNF-α, IL6
Tissue degeneration
Age-associated adipogenesis
Increase in PAMPs and alarmins
Senescence-associated gene activation
Accumulation of effector T cells
Defective innate immunity
Frailty
Coronary artery disease

FIG 38.3 **Inflammation in Older Adults.** The schematic diagram depicts possible mechanisms that account for the increased production of inflammatory mediators with age. These mediators contribute to many of the age-associated diseases.

lymphoreticular tissue (MALT), results in increased leakage and increased systemic levels of lipopolysaccharide (LPS) and innate immune activation. Failure in maintaining T-cell population homeostasis and accumulation of effector T cells also favors an inflammatory response.

KEY CONCEPTS

Causes for Increased Constitutive Production of Inflammatory Mediators With Age (Inflammation in Older Adults)

- Activation of the innate immune system as a result of defective epithelial barrier function
- Activation of the innate immune system as a result of defective adaptive immunity
- Accumulation of adipocytes producing inflammatory mediators
- Accumulation and activation of T-effector cell populations
- DNA damage-induced transcription of inflammatory cytokine genes in senescent cells

The immune system, however, is not the only source of inflammatory cytokines. Adipocytes, in part replacing muscle cells in older adults, produce various inflammatory mediators. Cellular senescence has also been associated with the production of cytokines. Persistent DNA damage response signaling not only induces an irreversible cell cycle block that is characteristic of cellular senescence but also initiates a transcriptional program to secrete numerous growth factors, proteases, and inflammatory cytokines, termed the *senescence-associated secretory phenotype* (SASP).[16]

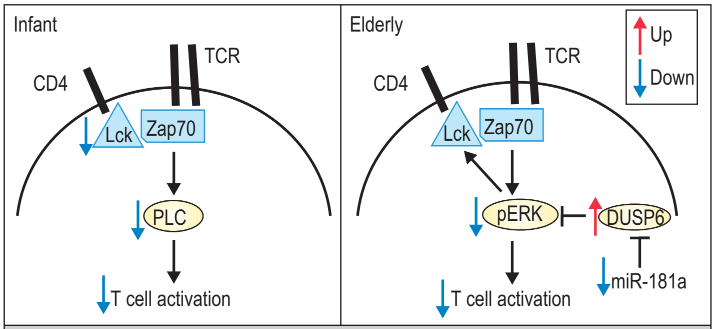
CELLULAR DEFECTS AND SENESCENCE

As described so far, immune aging occurs at the system level with organizational restructuring. Equally important are changes at the single-cell level, which are partly cell-intrinsic and partly caused by the host environment. The increased cytokine concentrations in older adults not only activate but also attenuate signaling pathways. Low responsiveness to cytokine stimuli is frequently seen in those cells that have increased baseline activation of a signaling pathway (*e.g.,* cells that constitutively have increased signal transducer and activator of transcription 3 (STAT3) or STAT1 phosphorylation respond less to IL-6/granulocyte macrophage–colony-stimulating factor (GM-CSF) or type I/II IFNs, cytokines that activate these STATs). Attenuation of signaling pathways by induction of negative feedback loops explains, in part, the reduced responsiveness and functionality of cells of the innate immune system.

Although neutrophil and monocyte/macrophage numbers remain normal, many of their functions decline with age.[17] Decreased chemotaxis in neutrophils delays tissue infiltration; reduced phagocytosis and respiratory burst compromise the ability to control bacterial infections; and TLR-induced monocyte/macrophage activation is dampened in older adults. Declines in responsiveness, for example, to TLR stimulation, are partly reversible *in vitro,* suggesting that they are not intrinsic. Adaptive immune cells are also directly affected by the proinflammatory environment in the aging host.

Equally important are cell-intrinsic changes that appear to be a consequence of the replicative history and failures in cellular processes, such as DNA repair and autophagy. The program most obviously influenced by age is cellular senescence. In all hematopoietic cell lineages, including stem cells, telomeric lengths decline with age. This is of particular importance for T cells because much of their response depends on their ability to proliferate and clonally expand.[18] Telomeric erosion results not only from cumulative replicative history and DNA damage but also from decline in the ability to express telomerase and repair telomeric ends.

Lymphocytes in older adults are more differentiated than those in the young. Although most obvious for CD8 T cells, increasing differentiation can also be noted for B cells and CD4 T cells (Fig. 38.4). Differentiation is generally driven by antigen recognition but could also occur in the absence of exogenous antigen under the influence of cytokines. Proliferation alone may be sufficient to drive initial steps of differentiation. A classic



	Infant	Elderly
	Internal changes	
Telomere length	+++	+
DNA damage	+	+++
Epigenetic modifications	+	+++
	Functional outcomes	
Proliferation capacity	+++	++*
T cell activation	+	+
Effector function	+	+*

+ = low ++ = normal +++ = high *controversial

FIG 38.4 **Intrinsic Differences in CD4 T Cells at the Extremes of Age.** CD4 T cells from infants and older individuals demonstrate similar dysfunctions, including poor activation and reduced effector functions, although as a result of distinct mechanisms. Signaling cartoons illustrate the best documented defects. Moreover, infant and older adult T cells have very different internal changes to their DNA, including telomere length, the amount of DNA damage, and epigenetic modifications.

KEY CONCEPTS

Cellular Dysfunction With Age

- Exposure to aging host environment (*e.g.,* inflammatory cytokines) activates negative regulatory signaling loops.
- Telomeric erosion impairs proliferative competence and restraints clonal expansion.
- End-differentiation reduces functional plasticity.
- Activation of specific gene programs modifies cell function:
  - Gene programs associated with differentiation (*e.g.,* microRNA [miRNA])
  - Gene programs associated with T-cell exhaustion (*e.g.,* expression of programmed death-1 [PD-1])
  - Loss of CD28 on T cells
  - Gain in natural killer (NK) cell–associated regulatory receptors on T cells (*e.g.,* killer immunoglobulin-like receptor [(KIR], killer lectin-like receptor [KLR], immunoglobulin-like transcript [ILT])
  - Senescence-associated gene activation (*e.g.,* inflammatory mediators)

example is the acquisition of memory-like and effector phenotypes with lymphopenia-induced homeostatic proliferation. So-called virtual memory cells, which presumably have never seen an exogenous antigen, have been identified in mouse models; clonal expansions are also seen within the human naïve T-cell compartment. Some of the changes in gene expression that are seen in naïve T cells with age may represent partial differentiation, such as declines in the microRNA miR181a and changes in the expression of phosphatases and other signaling molecules.[18] Increased expression of cytoplasmic phosphatases impairs the TCR-induced activation of naïve T cells. Changes in cell surface molecules that are seen with terminal differentiation, such as the gain in CD57 and the loss of CD27 and CD28 expression, are the most striking. Of functional importance, predominantly for CD8 T cells, is the gain in expression of cell surface receptors that are usually only found in NK cells. Most of these receptors have inhibitory function, but some of them also stimulate. Since expression of these receptors on individual cells is stochastic, the consequences can range from immunosuppression to autoreactivity.

CLINICAL CONSEQUENCES OF IMMUNE AGING—IMMUNODEFICIENCY, AUTOIMMUNITY, AND ACCELERATED DEGENERATIVE DISEASES

The most profound and most noted consequence of human senescence is the increased susceptibility to infections. Upper respiratory bacterial and urinary tract infections are frequent in the older population and less contained by the innate immune system and preexisting adaptive immunity. Not surprisingly, the immune system of an older adult is not able to induce a protective response to new antigens to which the individual has not been exposed to in the past. Clinically important examples are the severe acute respiratory syndrome (SARS) epidemic and West Nile fever virus infection. First-time vaccinations with live viruses, for example, yellow fever virus, are associated with increased morbidity and even mortality in older adults. Despite annual vaccination, influenza infections continue to be associated with high morbidity and mortality. Pneumonia caused by RSV, usually infecting young children, is not uncommon. Immune competence to chronic infections is also compromised with age. The best example here is the reactivation of VZV manifesting as shingles.

CLINICAL RELEVANCE

Consequences of Immune Aging

- Increased morbidity and mortality from bacterial infections (*e.g.*, pneumococci)
- Increased morbidity and mortality from viral infections (*e.g.*, influenza, West Nile fever)
- Reactivation of latent virus (*e.g.*, varicella zoster virus)
- Ineffective primary and booster vaccinations
- Acceleration of degenerative diseases as a result of the production of inflammatory mediators (*e.g.*, atherosclerotic disease, Alzheimer disease, osteoarthritis)
- Increased incidence of autoimmune disease (*e.g.*, polymyalgia rheumatica, giant cell arteritis, rheumatoid arthritis)

Vaccinations against pneumococcal antigens, influenza strains, and VZV are recommended in older adults; however, the efficacy of vaccination is reduced.

Immune aging also predisposes for autoimmune manifestations and a breakdown in self-tolerance.[19] Autoantibodies are a common finding in healthy older adults; many of these autoantibodies are specific for common autoantigens, such as IgG Fc or nuclear components. The risk for several autoimmune diseases, most notably polymyalgia rheumatica and giant cell arteritis, increases with age. Although polymyalgia rheumatica predominantly presents as an activation of innate immunity, giant cell arteritis is clearly a disease of the adaptive immune system with T cell–dependent granulomatous inflammation in the vascular wall of midsized and large arteries.

The low-grade inflammation in the aging host has direct clinical consequences in promoting frailty and sarcopenia and accelerating degenerative diseases, including coronary artery disease, osteopenia, and Alzheimer disease. Accelerated immune aging may be one of the reasons that autoimmune diseases, such as rheumatoid arthritis (RA), are associated with a shorter life span and increased risk for cardiovascular morbidity. Inflammation as a manifestation of accelerated aging has been also implicated in the increased morbidity and mortality of patients with HIV infection in spite of highly active antiretroviral therapy (HAART).[20]

STRATEGIES AND INTERVENTIONS ON THE HORIZON

Vaccinations hold the promise for reducing increased susceptibility to infections in young children and older adults, but improving vaccine responses has proven to be a challenge. Current strategies for targeting more elusive pediatric pathogens (*e.g.,* RSV) included maternal vaccination (*i.e.,* the mother is vaccinated during pregnancy and vaccine-specific IgG is transferred *in utero* to the infant) and the development of infant-specific vaccine adjuvants utilizing newer understanding of the developing infant immune system. Also, further investigations into the role of microbiome and immune development (or dysfunction) may open new strategies to influence immune system development to reduce immune deficiencies and prevent inappropriate hyperactivity.

In older adults, a proper T-cell response to infectious organisms or to vaccines depends on the availability of a T-cell repertoire that includes antigen-receptor specificities to respond to vaccination. Repertoire contraction resulting from thymic involution or memory inflation appears to be less severe than originally thought, and currently explored interventions to restore thymic activity may only be meaningful for selected populations, such as bone marrow recipients. Deficiency in the antigen-presenting system and in costimulatory signals may be overcome by identifying new adjuvants. Increasing the vaccine dose is another promising approach. Live vaccines or self-replicating constructs that also accomplish higher antigen loads may not have a sufficient safety profile in older adults. Direct targeting of signaling defects in older T cells resulting from increased expression of cytoplasmic phosphatases or inhibitory cell surface receptors will be feasible, as shown for checkpoint inhibitors in oncology; however, such approaches need to demonstrate a much better safety profile.

Interventions to influence inflammation in older adults at present needs to be nonspecific, given the multitude of underlying mechanisms. Immunomodulatory therapy is obviously standard practice in patients with autoimmune disease who exhibit accelerated aging and increased all-cause mortality. Calorie restriction to slow immune aging to an extent that it is effective is generally not well accepted. Statins and aspirin are being routinely used to prevent cardiovascular disease; their effect may be mostly antiinflammatory. Future interventions will require the development of mild and low-toxicity medications to reduce low-grade inflammation while not impairing the ability to prevent innate immune activation in response to harmful antigens.

ON THE HORIZON

- Improved vaccination strategies tailored to the infant or aged immune system (novel adjuvants, novel vaccine delivery systems)
- New vaccines for pregnant females to confer passive immunity
- Manipulation of the microbiome composition to influence immune system development
- Thymic rejuvenation (*e.g.* with KGF, IL-7 and other mediators)
- Prevention of chronic infection that accelerate immune aging (*e.g.* immunization for CMV)
- Pharmacological approaches to improve T and B cell activation, clonal expansion and differentiation
- Treatment of inflamm-aging

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Lopez-Otin C, Blasco MA, Partridge L, et al. The hallmarks of aging. Cell 2013;153(6):1194–217.
2. Basha S, Surendran N, Pichichero M. Immune responses in neonates. Expert Rev Clin Immunol 2014;10(9):1171–84.
3. Gollwitzer ES, Marsland BJ. Impact of Early-Life Exposures on Immune Maturation and Susceptibility to Disease. Trends Immunol 2015;36(11):684–96.
4. Simon AK, Hollander GA, McMichael A. Evolution of the immune system in humans from infancy to old age. Proc Biol Sci 2015;282(1821):20143085.
5. Siegrist CA, Aspinall R. B-cell responses to vaccination at the extremes of age. Nat Rev Immunol 2009;9(3):185–94.
6. Round JL, Mazmanian SK. The gut microbiota shapes intestinal immune responses during health and disease. Nat Rev Immunol 2009;9(5):313–23.
7. Geiger H, de Haan G, Florian MC. The ageing haematopoietic stem cell compartment. Nat Rev Immunol 2013;13(5):376–89.

8. Kogut I, Scholz JL, Cancro MP, et al. B cell maintenance and function in aging. Semin Immunol 2012;24(5):342–9.
9. Naradikian MS, Hao Y, Cancro MP. Age-associated B cells: key mediators of both protective and autoreactive humoral responses. Immunol Rev 2016;269(1):118–29.
10. Palmer DB. The effect of age on thymic function. Front Immunol 2013;4:316.
11. Goronzy JJ, Fang F, Cavanagh MM, et al. Naive T cell maintenance and function in human aging. J Immunol 2015;194(9):4073–80.
12. Nikolich-Zugich J, Li G, Uhrlaub JL, et al. Age-related changes in CD8 T cell homeostasis and immunity to infection. Semin Immunol 2012;24(5):356–64.
13. Akbar AN, Henson SM. Are senescence and exhaustion intertwined or unrelated processes that compromise immunity? Nat Rev Immunol 2011;11(4):289–95.
14. Wherry EJ, Kurachi M. Molecular and cellular insights into T cell exhaustion. Nat Rev Immunol 2015;15(8):486–99.
15. Kanapuru B, Ershler WB. Inflammation, coagulation, and the pathway to frailty. Am J Med 2009;122(7):605–13.
16. Tchkonia T, Zhu Y, van Deursen J, et al. Cellular senescence and the senescent secretory phenotype: therapeutic opportunities. J Clin Invest 2013;123(3):966–72.
17. Shaw AC, Goldstein DR, Montgomery RR. Age-dependent dysregulation of innate immunity. Nat Rev Immunol 2013;13(12):875–87.
18. Goronzy JJ, Weyand CM. Understanding immunosenescence to improve responses to vaccines. Nat Immunol 2013;14(5):428–36.
19. Goronzy JJ, Weyand CM. Immune aging and autoimmunity. Cell Mol Life Sci 2012;69(10):1615–23.
20. Deeks SG. HIV infection, inflammation, immunosenescence, and aging. Annu Rev Med 2011;62:141–55.

MULTIPLE-CHOICE QUESTIONS

1. The developing immune system is characterized by:
   A. Increased frequencies of regulatory T cells.
   B. Innate immune system activation.
   C. Frequent presence of autoantibodies caused by tolerance defects.
   D. Hematopoietic stem cells preferentially differentiating into myeloid and not lymphoid lineages.

2. An inflammatory environment in older adults is:
   A. Less pronounced in frail older adults.
   B. Caused by thymic involution.
   C. Characterized by the production of cytokines from various cell types.
   D. Results from failure of naïve T cells to differentiate.

3. The following consideration to vaccinations in older individuals is correct:
   A. Annual influenza vaccinations are not always protective and are therefore not recommended in older individuals.
   B. Vaccinations with live viruses (*e.g.*, yellow fever vaccine) are potentially harmful.
   C. Only booster, but not primary, vaccine responses are impaired in older individuals.
   D. Vaccinations even with inactivated or component vaccine carry a higher risk in older adults.

4. T cells are derived from hematopoietic stem cells that differentiate in the thymus. Which of the following statements is correct?
   A. Thymic involution leads to a rapid loss of naïve T cells.
   B. Upon thymic involution, peripheral homeostatic proliferation of existing T cells to is effective to at least partially compensate for T cell loss.
   C. The thymus starts to involute after the age of 50 years.
   D. Generation of T cells by homeostatic proliferation can increase the T-cell receptor repertoire.

39

Human Immunodeficiency Virus Infection and Acquired Immunodeficiency Syndrome

Susan L. Gillespie, Javier Chinen, Mary E. Paul, William T. Shearer

In recent years, there has been a stabilization and possible reversal of the overall growth of the global acquired immunodeficiency syndrome (AIDS) epidemic. The number of new infections has steadily declined since the late 1990s, and there are fewer AIDS-related deaths as a result of the scale-up of antiretroviral therapy (ART) since 2004. People living with human immunodeficiency virus (HIV) are now living longer and healthier lives.

According to the Joint United Nations Programme on HIV/AIDS (UNAIDS), 36.9 million people worldwide were estimated to be living with HIV or AIDS at the end of 2014 (Fig. 39.1).[1] New HIV infections are declining in most countries. The annual number of incident HIV infections peaked at >3 million in the late 1990s and has steadily declined thereafter. In 2014, 2 million people were newly infected, with approximately 1% of these infections (220 000 cases) occurring in children younger than 15 years of age.[1] Globally, however, adolescents, particularly adolescent girls, are at particular risk for acquiring HIV infection. In 2012, there were 300 000 new infections reported among adolescents aged 15–19 years, which accounted for about 13% of all new infections. In sub-Saharan Africa, 70% of new infections in adolescents occurred among girls.[2] There has been a slow but continued reduction in the number of people dying from AIDS-related conditions worldwide. AIDS mortality peaked at 2.1 million in 2004, and by the end of 2014, it was down to 1.2 million deaths resulting from AIDS-related illnesses.[1] The decline reflects increased access to ART and improvements in the care and support of infected individuals. Mortality among children younger than 15 years of age has also declined as a result of the expansion of services to prevent mother-to-infant transmission of HIV and an increase in access to care and treatment for children. Although access to care and treatment services for HIV-infected children in resource-limited settings is expanding, it is estimated that only 31% of children in need of life-saving ART received it, and an estimated 150 000 children had died of AIDS-related illnesses in 2014. Of concern, HIV/AIDS is the most common cause of death among adolescents in sub-Saharan Africa and the second leading cause of death among adolescents globally.

US Perspective

The Centers for Disease Control and Prevention (CDC) estimate that almost 1.2 million adults and adolescents were living with HIV in the United States at the end of 2014, including 156 300 (12.8%) whose infection was undiagnosed. Despite ongoing prevention efforts designed to reduce the number of new cases of HIV infection, the number of newly diagnosed infections in the United States has remained unchanged from 2006–14, approximately 40 000–50 000 per year. Within the United States, there is great variability in both the geographical and demographic distribution of the disease, with certain segments of the American population being disproportionately affected, specifically men who have sex with men (MSM) and ethnic and racial minorities, including African Americans and Hispanic Americans (Fig. 39.2).[3] In the United States and other developed countries, in contrast to more resource-limited parts of the world, the number of children newly infected with HIV has decreased dramatically as a consequence of successful interventions against perinatal mother-to-child transmission. In 2014, 174 children were diagnosed with HIV infection, 73% of whom had acquired the infection through perinatal transmission.[4] At the same time, new pediatric AIDS cases and AIDS deaths also have plummeted, in large part as a result of powerful combinations of antiretroviral (ARV) drugs.

KEY CONCEPTS

Trends in Human Immunodeficiency Virus (HIV) Infection

- Global rates of HIV infection have stabilized, but there is great heterogeneity in disease incidence among countries and regions.
- Although most countries had a reduction in disease incidence between 2001 and 2009, seven countries had increases in disease incidence by more than 25% over that period.
- The age demographic most affected in the developing world, that is, those aged 25–44 years, includes men and women who are economically productive and women of childbearing potential.
- Worldwide, most infections are acquired through heterosexual contact.
- In the United States, African Americans and men who have sex with men (MSM) are disproportionately infected.
- The numbers of infections transmitted from mother to child are declining as access to prophylactic medications to prevent infection improves.

HIV PATHOGENESIS

HIV Lifecycle

HIV is a lentivirus that targets CD4 T cells by specifically binding the viral Env protein to two cell surface proteins, the CD4 receptor, and either the CCR5 or the CXCR4 chemokine receptor (Fig. 39.3). Other cell targets for HIV are monocytes and dendritic cells (DCs), although they present less expression

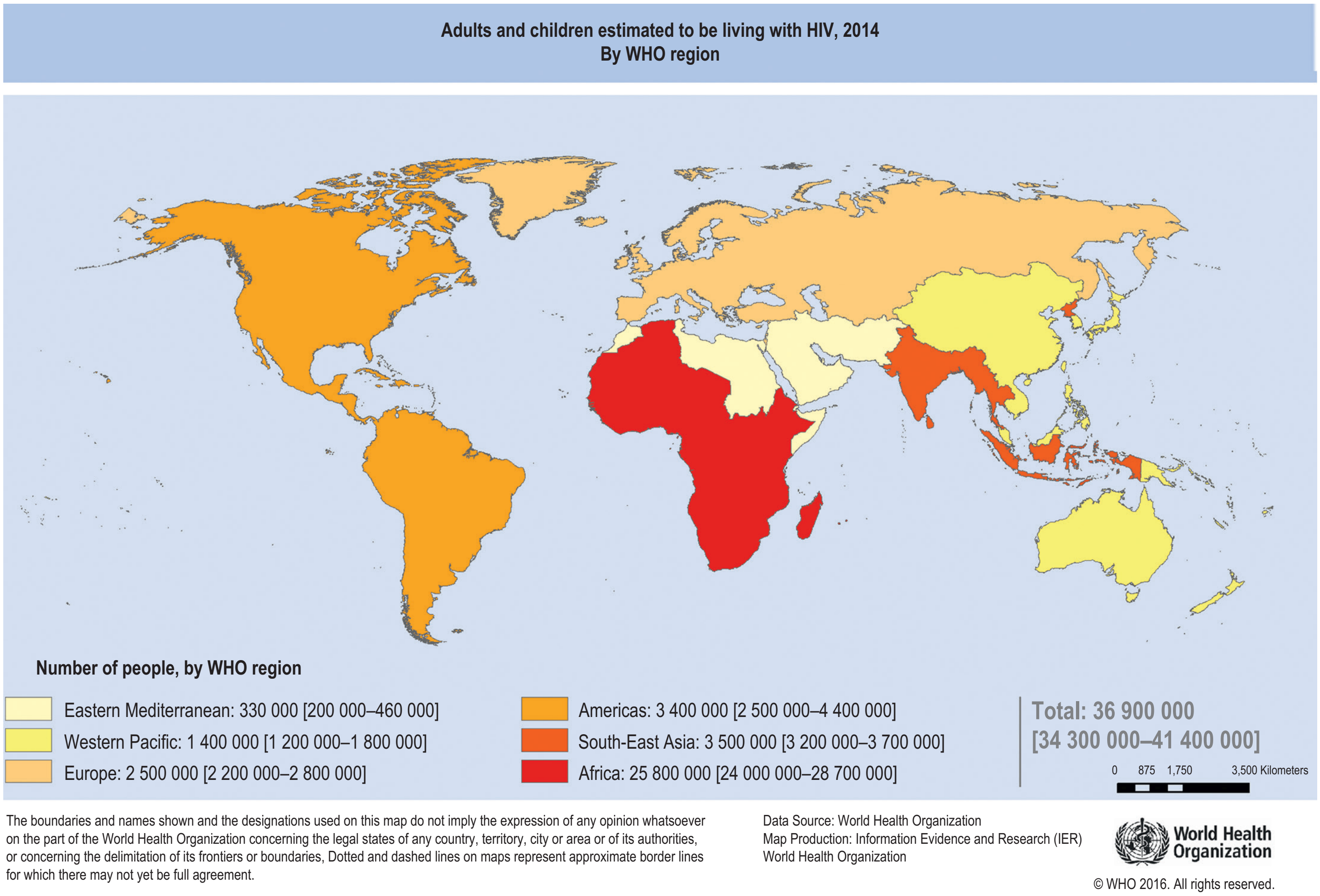

Adults and children estimated to be living with HIV, 2014
By WHO region
Number of people, by WHO region
Eastern Mediterranean: 330 000 [200 000–460 000]
Western Pacific: 1 400 000 [1 200 000–1 800 000]
Europe: 2 500 000 [2 200 000–2 800 000]
Americas: 3 400 000 [2 500 000–4 400 000]
South-East Asia: 3 500 000 [3 200 000–3 700 000]
Africa: 25 800 000 [24 000 000–28 700 000]
Total: 36 900 000
[34 300 000–41 400 000]
0 875 1,750 3,500 Kilometers
The boundaries and names shown and the designations used on this map do not imply the expression of any opinion whatsoever on the part of the World Health Organization concerning the legal states of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries, Dotted and dashed lines on maps represent approximate border lines for which there may not yet be full agreement.
Data Source: World Health Organization
Map Production: Information Evidence and Research (IER)
World Health Organization
World Health Organization
© WHO 2016. All rights reserved.

FIG 39.1 Adults and children estimated to be living with human immunodeficiency virus (HIV), 2014: by World Health Organization (WHO) region, 2014.

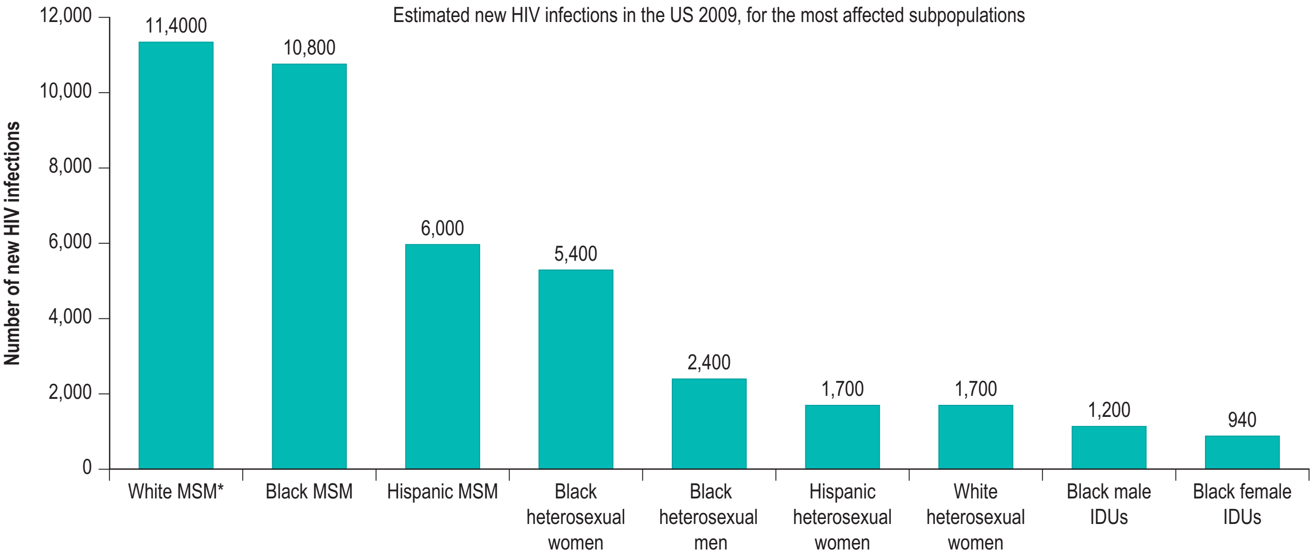
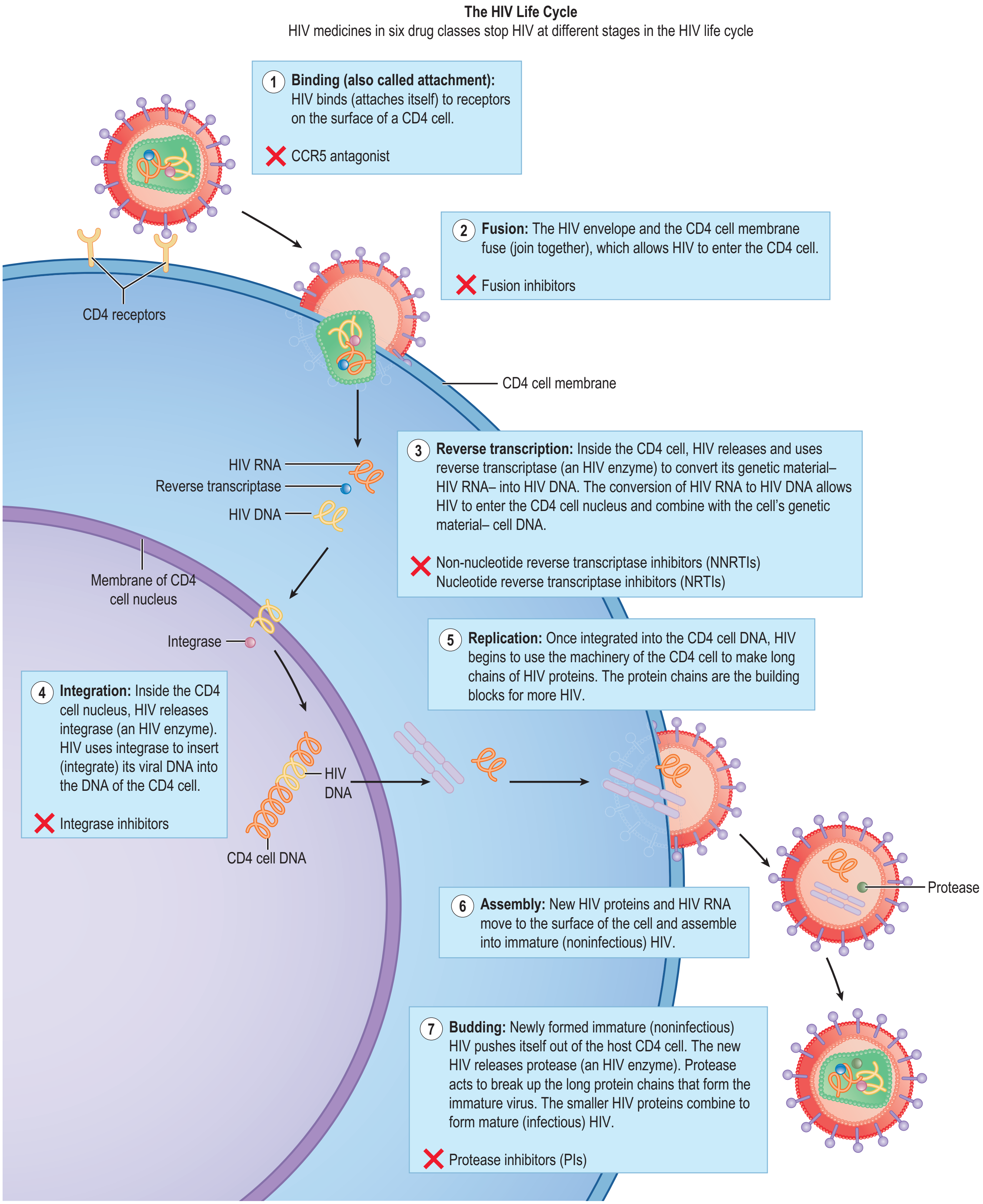


FIG 39.2 US trends in human immunodeficiency virus (HIV) infection by race and mode of transmission. (From Centers for Disease Control and Prevention. HIV surveillance—United States, 1981–2008. MMWR Morb Mortal Wkly Rep 2011; 60: 689–93. Atlanta, GA: US Department of Health and Human Services, Centers for Disease Control and Prevention; 2011. (Also available at http://www.cdc/gov/hiv/topics/surveillance/resources/reports.gov.)



The HIV Life Cycle

HIV medicines in six drug classes stop HIV at different stages in the HIV life cycle

① **Binding (also called attachment):** HIV binds (attaches itself) to receptors on the surface of a CD4 cell.

✗ CCR5 antagonist

② **Fusion:** The HIV envelope and the CD4 cell membrane fuse (join together), which allows HIV to enter the CD4 cell.

✗ Fusion inhibitors

③ **Reverse transcription:** Inside the CD4 cell, HIV releases and uses reverse transcriptase (an HIV enzyme) to convert its genetic material– HIV RNA– into HIV DNA. The conversion of HIV RNA to HIV DNA allows HIV to enter the CD4 cell nucleus and combine with the cell's genetic material– cell DNA.

✗ Non-nucleotide reverse transcriptase inhibitors (NNRTIs)
Nucleotide reverse transcriptase inhibitors (NRTIs)

④ **Integration:** Inside the CD4 cell nucleus, HIV releases integrase (an HIV enzyme). HIV uses integrase to insert (integrate) its viral DNA into the DNA of the CD4 cell.

✗ Integrase inhibitors

⑤ **Replication:** Once integrated into the CD4 cell DNA, HIV begins to use the machinery of the CD4 cell to make long chains of HIV proteins. The protein chains are the building blocks for more HIV.

⑥ **Assembly:** New HIV proteins and HIV RNA move to the surface of the cell and assemble into immature (noninfectious) HIV.

⑦ **Budding:** Newly formed immature (noninfectious) HIV pushes itself out of the host CD4 cell. The new HIV releases protease (an HIV enzyme). Protease acts to break up the long protein chains that form the immature virus. The smaller HIV proteins combine to form mature (infectious) HIV.

✗ Protease inhibitors (PIs)

FIG 39.3 The human immunodeficiency virus (HIV) lifecycle and stage susceptibility to antiretroviral therapy. (From https://www.aidsinfo.nih.gov/education-materials/fact-sheets/19/73/the-hiv-life-cycle.)

of these receptors compared with CD4 T cells. After the virus is attached to the cell, the viral membrane fuses with the cell membrane and viral RNA enters the cytoplasm. It is transcribed by the viral reverse transcriptase into double-stranded DNA, which is translocated to the cell nucleus, and the virus integrase enzyme mediates its integration into the cell genome. The viral proteins Tat and Nef, in addition to T-cell activation factors, induce active transcription and expression of viral RNA and proteins to produce new virus particles that exit the host cell to infect other cells. CD4-independent viral entry has been demonstrated in B cells, astrocytes, and kidney epithelial cells; however, efficient viral replication is not likely to occur in these cells (see Treatment).

HIV Entry Through Mucosal Surfaces

Using dilution techniques that can precisely determine unique virion genome sequences, it was found that about 80% of HIV infections via heterosexual transmission start with a single virion.[5] This is remarkable, considering the high rate of mutation and diversity of viral genome sequences found in infected individuals. In contrast, HIV transmission occurring in MSM and in intravenous drug users is caused by several viral variants, indicating that the multilayered epithelia of the vaginal mucosa and the cervical mucus are likely deterrents for lentiviral infection.[6] Regardless of the route of infection, the time to develop viremia and immunological changes is similar to that of all forms of transmission. Viral replication can be detected in the germinal center of lymph nodes within a week, and viremia is found within 21 days of infection. The most active replication occurs in the gut-associated lymphoid tissue (GALT), with massive depletion of CD4 T cells. Activation and destruction of CD4 T cells are associated with an increase of serum cytokines, which might explain the flu-like syndrome that patients might experience during acute HIV infection.[7] Simultaneously, HIV reservoirs become established with viral infection and integration of viral DNA in resting cells.

T-cell Depletion

Progressive T-cell depletion in HIV infection is induced, in part, by a state of chronic immune activation, which also contributes to noninfectious complications, such as cardiovascular disease.[7] In individuals with HIV infection, immune activation is caused by increased microbial translocation in the gut, direct Toll-like receptor (TLR) stimulation by HIV, and coinfection with other pathogens. This activated state results in increased levels of inflammatory cytokines, such as interleukin-6 (IL-6). Also detected is an increase of lipopolysaccharide (LPS) in blood, which, in turn, can also activate the coagulation cascade and promote thrombosis. Low or normal levels of T regulatory cells (Tregs) are observed in long-term nonprogressors, suggesting their role in anti-HIV immunity, especially in modulating CD8 T-cell immunity. It has also been suggested that Tregs may play a beneficial role by reducing chronic immune activation.[8]

HIV Latency and HIV Reservoirs

The most difficult problem in the efforts toward achieving a cure for HIV infection is the presence of *HIV reservoirs*, defined as resting or nonreplicating cells infected with HIV.[9] Anti HIV drugs are able to efficiently suppress HIV replication; however, when these drugs are stopped, activation of the resting cells results in the production of new virus. Most viral DNA is detected in memory CD4 T cells within lymphoid tissues and peripheral blood. They might produce very low levels of viral RNA and proteins and therefore might not be detected by immunosurveillance. HIV reservoirs are established during the acute phase of the infection, hence the recommendation for early intensive treatment to reduce the population of latently infected CD4 T cells (nonreplicating monocytes, astrocytes, and glial cells are other latently infected cells). Molecular mechanisms that inhibit HIV replication from viral DNA and maintain latent infection include histone deacetylases and methyl transferases with methylation of viral genes, especially of the long-terminal promoter repeat. These epigenetic modifications can be reversed by using T-cell activation signals.

ANTI-HIV IMMUNITY

Immunity against HIV depends mostly on specific cytotoxic CD8 T cells, which recognize and destroy infected cells.[7] These antiviral cells are most efficient when certain combinations of human leukocyte antigen (HLA) and virus strain occur in the host, such as the presence of a cell bearing the HLA-B27 allele and infected with clade B viral strain. However, HIV infection almost always results in global T-cell destruction and exhaustion. Antibody responses are initially directed against the gp41 portion of the Env protein. This response and subsequent anti-HIV antibodies become ineffective because of the rapid production of escape mutants.

HIV Vaccines: Basic Concepts

Several efforts to develop an effective anti-HIV vaccine have attempted to address the genetic diversity of HIV that develops in the infected individual, by identifying epitopes that remain relatively conserved and by inducing neutralizing antibodies as well as cell-mediated immunity.[10] Broadly neutralizing antibodies, which target conserved Env epitopes from different strains and can prevent HIV from establishing infection, might develop in some individuals with chronic infection, offering hope for disease prevention. Such antibodies are found to develop after approximately 2 years of infection, when many generations of escape virions, inducing new antibody epitopes, allow the immune response to refine and select for best affinity. This concept suggests that several booster immunizations might reproduce the process in shorter time. Current approaches involve viral vectors that drive the expression of HIV proteins in host tissue; an example is a canary pox virus carrying the HIV *env* gene. No current vaccines have been able to induce high titers of broadly neutralizing antibodies against HIV Env. Of note, an unexpected, although modest, increased rate of HIV infection was observed in individuals immunized with an adenovirus-based HIV vaccine compared with nonimmunized individuals and was highest in persons who had antiadenovirus antibodies before the trial. The mechanisms explaining this increased infection rate are not clear. Results from the RV144 clinical trial conducted in Thailand showing a 31% reduction of HIV transmission demonstrated that vaccine protection may be a reality. New vector approaches, improving T-cell responses and achieving high titers of broadly neutralizing vaccines, are being investigated to increase vaccine efficacy (see HIV preventive vaccines).

ROUTES OF INFECTION

HIV is transmitted through three routes of infection: sexual, parenteral, and perinatal. In general, higher viral loads, lower

CD4 T-cell count, and larger viral inoculum size are all associated with a greater risk of transmission.

Sexual transmission is the most common mode of infection, responsible for 70–80% of worldwide infections. Besides the factors mentioned above, recipients in penetrative intercourse are more likely to become infected. Anal intercourse carries the greatest risk, followed by vaginal intercourse, and oral intercourse is the least likely to spread the virus. The presence of other sexually transmitted diseases, especially ulcerative lesions, such as those seen in herpes simplex or syphilis infection, increases the risk of transmission.

Parenteral transmission is the second most common route of HIV infection, accounting for 8–15% of all HIV infections. Examples include contaminated needles used by people who are addicted to intravenous drug use (IDU), accidental needle sticks occurring in health care workers, improperly sterilized hospital equipment, and contaminated blood products. Use of contaminated needles by people addicted to IDU is the largest risk factor for parenteral transmission. In 2013, 7% of the estimated 47 500 new HIV infections in the United States were attributed to IDU.[11] Although infection control efforts have greatly reduced the risk, HIV transmission via transfusion of contaminated blood products remains significant, particularly in resource-limited settings where there is a dependence on replacement of blood by family member donors and paid blood donors and where the infrastructure for routine blood screening is suboptimal.

Perinatal transmission accounts for the majority of pediatric HIV cases and for 5–10% of HIV infections in patients of all ages. The virus is capable of infecting the child *in utero*, during labor, and after delivery through breastfeeding. Without preventive therapy, the risk of a child contracting HIV from the mother during gestation or during labor is about 15–40%. The majority of transmission events occur during the passage of the fetus through the birth canal, by exposure of the baby to infected maternal blood, amniotic fluid, and/or cervical and vaginal secretions.[12] Although breastfeeding increases the risk of transmission of the virus from mother to child by another 15–29%, breastfeeding is still recommended to prevent morbidity and mortality related to diarrheal diseases in settings where access to replacement feeding is limited.

IMMUNOPATHOGENESIS

The mechanisms of immunodeficiency induced by HIV are not limited to depletion of its main target cell, the CD4 T cell; HIV infection ultimately leads directly or indirectly to the impairment of every arm of the immune system.[6] The time of progression from infection to the development of AIDS is not the same for all infected individuals, and several explanations have been proposed, including genetic resistance genes and the presence of low-virulence mutant viral particles. The study of long-term nonprogressor individuals has helped to define T-cell and antibody features associated with anti-HIV immunity in the search for the optimal vaccine design, focusing on minimizing the escape variants and inducing persistent specific viral neutralization or inhibition of replication.[10] HIV researchers have also explored the role of chronic immune activation processes that are associated with progression to AIDS, which also helps explain the diverse clinical impact of this viral infection in different individuals.

Mucosal Dendritic Cells: Myeloid Versus Plasmacytoid

Following a mucosal inoculation of HIV, CD4 T cells are generally infected by CCR5-tropic virus, also called *R5* or *M-tropic virus.*[13] After approximately 1 week, the virus becomes detectable in draining regional lymph nodes. Myeloid DCs are not infected with HIV; rather, HIV gp120 binds to DC-specific intercellular adhesion molecule–grabbing nonintegrin (DC-SIGN) in the cell membrane. The resulting complex is internalized as a phagosome and then presented on the cell surface. As there is no fusion of the HIV Env with the DC membrane, infection does not occur. Plasmacytoid DCs, in contrast, express CD4 and the coreceptors CXCR4 and CCR5 and thus become infected by the virus. Infection leads to expression of CCR7, which acts as a homing signal for the lymph nodes; it is in lymph nodes that the virus infects CD4 T cells, and significant viral replication occurs, resulting in detectable viremia and dissemination through lymphoid tissues. Of interest, the target cells for HIV infection, CD4 T cells and monocyte/macrophage cells, are also reservoirs for the virus. HIV-infected cells carry a stable provirus integrated in the cell genome, which is transcriptionally silent until the cell is activated. These cells can reestablish HIV viremia even after prolonged antiviral treatment, since viral reservoirs are unlikely to be soon eradicated.

Gastrointestinal System: Early Target

Within days of infection, 20% of CD4 T cells found in the GALT are infected. Of these, up to 80% are killed, by lytic infection and Fas-mediated apoptosis of both infected and noninfected cells, mainly $CD4^+CCR5^+$ memory T cells. By the time of peak viremia, 60% of the mucosal memory CD4 T cells are infected. Because their destruction is less drastic, circulating CD4 T-cell counts do not reflect the magnitude of CD4 T-cell death taking place in the GALT. Following acute HIV infection, a massive activation of the immune system occurs. Inadequate memory T-cell responses can lead to infection and inflammation of the entire bowel. The presence of immune activators over the massive surface area of the bowel has been hypothesized to result in the profound immune activation seen in HIV. Mucosal Th17 cells are a preferential target for HIV, further weakening mucosal immunity and favoring microbial infection. The viremia decreases spontaneously from as high as 10 million copies of HIV RNA per milliliter during the acute illness phase to a stable level many orders of magnitude lower called the *viral set point.* It is known that higher set points of viral load and lower T-cell counts are loosely predictive of shorter periods of clinical latency. Eventually, a prolonged period of homeostasis between the virus and the immune system collapses, and AIDS ensues.

Chronic Immune Activation and Progression to AIDS

After the acute stage, HIV infection induces a high degree of proliferation and turnover of CD4 and CD8 T cells. The massive immune activation can be explained by the circulation of many soluble factors, including TLR ligands. HIV infection appears to activate plasmacytoid DCs by stimulation of TLR7 to secrete interferon-α (IFN-) and proinflammatory cytokines, using both TLR8 and DC-SIGN to infect these cells.[14] In turn, stimulated DCs activate T cells, which are also being directly activated through TLR ligands. Bacterial products from the mucosa (*e.g.*, LPS) and HIV protein products, such as Nef, Env, Vif, and so on, are mediators of this immune activation. Other microbial infections (viral, bacterial, or fungal) can stimulate the different TLRs and result in CD4 and CD8 T-cell activation and apoptosis. This rampant immune activation is thought to lead to the eventual collapse of the immune system.

Anti-HIV Cellular Immunity

In individuals with HIV infection, the initial control of viremia occurs with the initial expansion of HIV-specific CD8 T cells. The importance of anti-HIV cytotoxic T cells is suggested by the development of viral escape mutations that are driven by immune selection pressure and evade the immunological response. In addition to their cytotoxicity, CD8 T cells secrete chemokines, such as RANTES and MIP-1β, which inhibit viral entry via CCR5, and IFN-γ, which activates immune cells and increases HLA expression (Chapter 10). Although severely reduced in number, IFN-γ–secreting HIV-specific CD4 T cells are present in individuals with HIV infection. The role of adaptive cellular immunity in protection against natural infection is currently challenged, since about 40–60% of noninfected but exposed individuals do not present with detectable anti-HIV responses.

Mechanisms of T-Cell Depletion

Three major mechanisms of T-cell depletion have been reported: direct lytic infection, apoptosis, and autophagy. Direct lytic infection is the result of massive activation of the immune system with production of inflammatory cytokines, as in the GALT (see Cytokines in HIV infection).

Apoptosis

Increased apoptosis of T cells in HIV infection was described early in the history of the epidemic. Several explanations have been offered: *(i)* HIV-induced apoptosis of infected cells (viral cytopathic effect); *(ii)* bystander effect from HIV-infected neighbor cells releasing viral proteins; *(iii)* death of HIV-specific effectors following their migration to infected sites; *(iv)* perturbation of proapoptotic signaling molecules on immune cells secondary to the chronic immune activation; and *(v)* destruction of HIV-infected cells by immune effectors.

Late HIV infection is associated with the dominance of a syncytia-inducing form of the virus. Syncytia formation (clustering of CD4 T cells and membrane fusion) occurs through the interaction of HIV Env and CD4/CXCR4 on neighboring cells. These cells are more prone to undergoing apoptosis through a Fas-dependent pathway.[15] Most of the HIV proteins have been implicated at one time or another in HIV-induced apoptosis.

Autophagy

Autophagy is a process by which cytoplasm and organelles are sequestered and directed toward lysosomal pathways and histone deacetylase.[16] This process has been implicated in both the prevention and induction of apoptosis, reflecting common regulatory factors shared by both (*e.g.,* tumor necrosis factor [TNF]–related apoptosis inducing ligand [TRAIL], FADD, DAPk, ceramide, and Bcl-2). HIV Env has been shown to be a stimulus for autophagy in uninfected CD4 T cells via its interaction with CXCR4. Inhibition of the autophagic pathway prevents Env-induced cell death in uninfected cells *in vitro,* demonstrating that this mechanism contributes to the loss of uninfected cells.

Anti-HIV Humoral Immunity

After establishment of HIV infection, neutralizing antibodies are produced; however, the virus quickly mutates to avoid them, such that the host continues to respond to evolving viral mutants. The fact that antibodies can be effective in protecting hosts from HIV infection has been demonstrated in macaques and in a humanized mouse model of HIV infection.

INNATE IMMUNITY

Innate immunity mechanisms are characterized by their lack of antigen specificity and include epithelial barriers, the complement system, phagocytes, and antigen-presenting cells (APCs). Their role in HIV pathogenesis is complex, as their function is of most importance during acute infection but their contribution to immune activation may be deleterious in the chronic stage. Studies including individuals who were exposed to HIV but did not develop infection have suggested that an increased innate immune response generated in the mucosal microenvironment may explain the failure to develop a specific anti-HIV response.[17]

NK Cells in HIV Infection

NK cells are dramatically altered in HIV infection.[18] A small subset of NK cells has been found to express both CD4 and CCR5 or CXCR4. These NK cells can be infected by HIV and may serve as one of the sites of latent infection or can be activated and contribute to the state of immune chronic activation. CCR5 and inhibitory receptor expression are higher on NK cells in patients with HIV infection than in HIV-seronegative subjects. In contrast, there is reduced expression of surface receptors on NK cells that induce cytotoxic activity, such as NK-cell protein 30 (NKp30), NKp44, and NKp46. Antibody-dependent cellular cytotoxicity (ADCC) is reduced in patients with HIV infection, as well as the NK-cell responsiveness to IL-2. The expression of HLA-C–specific inhibitory NK-cell receptors has been found to be increased in patients with HIV. Interestingly, studies of Vietnamese patients who were exposed to HIV through IDU but were seronegative showed that the NK cells of these patients not only secreted greater quantities of chemokines compared with those of the controls but they also had greater direct cytotoxicity. Genetic studies have suggested that inheritance of particular killer inhibitor receptor (KIR) alleles with their HLA ligands delays disease progression.[19] The two subsets of DCs, myeloid and plasmacytoid, play an additional role in HIV infection by stimulating innate immune responses, such as the type 1 IFNs, against HIV via plasmacytoid DCs. DCs may have opposing roles, with myeloid DCs acting as reservoirs and favoring HIV dissemination and plasmacytoid DCs inhibiting HIV replication by cytokine secretion. In fact, low levels of plasmacytoid DCs correlate with high viral loads, low CD4 T-cell counts, opportunistic infections, and disease progression.

Cytokines in HIV Infection

The dysregulation of the immune system produced by HIV infection includes significant perturbation in the balance of Th1 and Th2 cytokine levels. The Th1 cytokines IL-2, TNF-α, IFN-γ, and IL-12 decrease during HIV infection, whereas the Th2 cytokines, IL-10, and IL-4 increase or remain normal; levels of proinflammatory cytokines, such as IL-1, IL-6, and IL-8, also increase. Viral replication in HIV-infected T cells and monocytes is induced by IL-2, IL-7, and IL-15, as well as by the proinflammatory cytokines. These cytokines appear to induce viral replication by activating the host cell, a requirement for HIV productive replication. Some cytokines, such as IL-10, decrease HIV production, likely by inhibiting the synthesis of the activating cytokines and by decreasing the expression of CCR5 and other chemokine receptors. In addition, the HIV long terminal repeat (LTR) promoter contains sequences that bind cellular factors that are activated as a response to cytokine binding, such as nuclear factor (NF)-κB and AP1. IFN-α and IFN-β have activity against HIV, although their role of these cytokines *in vivo* is not well established and may have only an adjuvant effect.[20] Their activity might favor HIV infection, as shown by the upregulation of CCR5 by IFN-α, which can facilitate viral entry. Chemokines inhibit HIV infection by competing for viral binding sites. The α chemokine stromal cell–derived factor 1 (SDF-1) competes with lymphotropic HIV strains for binding to CXCR4, whereas the β chemokines macrophage inflammatory protein 1-α (MIP1-α), MIP1-β, and RANTES compete with macrophage-tropic strains for binding to CCR5.

CLINICAL FEATURES

If the HIV infection is left untreated, its natural history involves the progression through four clinical phases: acute retroviral syndrome, asymptomatic or latent infection, symptomatic HIV infection, and finally AIDS. Each clinical phase correlates with specific events in the interaction between HIV and the host immune system. A small percentage of patients become long-term nonprogressors, and an even smaller percentage become elite controllers (see Long-term nonprogressors/elite controllers).

Acute HIV Infection

Soon after infection, unopposed by effective host immune responses, HIV rapidly replicates and disseminates to lymphoid tissues (see Immunopathogenesis: gastrointestinal system) and to the systemic circulation, with viremia reaching as high as 10 million copies per milliliter. Plasma viremia typically peaks in 3–4 weeks after transmission, and then, as a result of the depletion of susceptible CD4 T cells and HIV-specific immune responses, the virus load precipitously declines, followed by more gradual decline for several weeks before reaching the set point.

Clinically, the acute phase of HIV infection is manifested by a flu-like illness, referred to as *acute retroviral syndrome.*[21] Two to 4 weeks after transmission, coinciding with the period of high plasma viremia and dissemination of virus to lymphoid organs, the majority of infected individuals experience a nonspecific infectious mononucleosis–like illness that lasts from a few days to several weeks. As the host develops HIV-specific immunity, the virus load decreases, CD4 and CD8 T cells recover, and the symptoms of the acute infection resolve. Although up to 90% of patients seek medical care for this illness, the nonspecific nature of the symptoms makes diagnosis of acute infection difficult, and most newly infected individuals are not diagnosed until much later. The public health implications of the acute HIV infection are enormous because the risk of transmission from individuals with acute infection appears to be much higher than that from those with established infection, in part because of the high viral load in the former.

Asymptomatic HIV Infection

The acute infection is followed by a prolonged asymptomatic or latent period that may last 8–10 years in adults but is much shorter in children. During this time, the HIV viral load fluctuates around a relatively stable set point.[22] The viral set point is a major determinant of infectivity and risk of disease progression, with higher viral loads being associated with more likely viral transmission, more rapid disease progression, and greater risk of death. The host immune response is insufficient to eradicate the infection but is enough to contain viral replication for many years. Although commonly thought to represent a stalemate between viral replication and CD4 T-cell production, this period is actually characterized by a steady and inexorable decline of CD4 T cells (50–75 cells per year).

CLINICAL PEARLS

Acute Infection Is an Opportunity for Early Human Immunodeficiency Virus (HIV) Diagnosis

- Acute HIV infection is a nonspecific viral syndrome, often described as being similar to infectious mononucleosis.
- Many patients present to a clinician at this stage, but most are not recognized as HIV infection.
- Irreparable damage to the host immune system occurs during this stage of HIV infection resulting in chronic immune activation and the eventual collapse of the immune system.
- HIV viral latency is established during the acute infection as HIV DNA integrates with the host genome resulting. Integrated viral genetic material makes cure of HIV impossible even after prolonged viral suppression with antiretroviral therapy (ART).
- ART initiated during the acute infection can halt the destruction of the body's memory T cells in the gut-associated lymphoid tissue (GALT) and lead to better long-term outcomes for the patient.
- The practical importance of early therapy of acute infection remains in question because patients who are able to achieve and maintain undetectable viral loads on ART do well despite having commenced treatment long after their acute infection.

Symptomatic HIV Infection (pre-AIDS)

As the infection progresses, most individuals develop clinical symptoms. The ability of the immune system to contain viral replication is overcome, and the viral load begins to increase. There is usually an inflection point in the CD4 T-cell curve marking the start of a period of more rapid decline in CD4 T-cell counts. As these counts fall, immunodeficiency, symptomatic disease, and AIDS eventually occur (Fig. 39.4).

End-Stage HIV Infection: AIDS

As the CD4 T-cell count drops to <200 cells/μL, the immune system's ability to fight infection is compromised to the extent

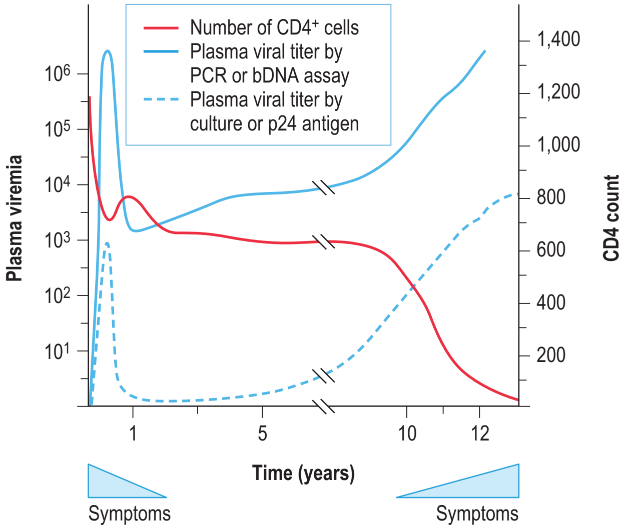


FIG 39.4 Natural course of human immunodeficiency virus (HIV) infection. The relationship between CD4 T-cell counts and viral loads over the course of infection in the absence of treatment. (From Baliga CS, Shearer WT. HIV/AIDS. In: Fireman P, ed. Atlas of allergy. 3rd ed. Philadelphia, PA: Elsevier Science (USA); 2005: p. 351–67.)

TABLE 39.1 **Opportunistic Infection Prophylaxis and Treatment in Adolescents and Adults***

Risk Factor	Agent	Prophylactic Medication
CD4 cell count <200 cells/μL	*Pneumocystis jiroveci*	Trimethoprim–sulfamethoxazole (TMP-SMX) or dapsone plus or minus pyrimethamine and leucovorin or aerosolized pentamidine or atovaquone
	Coccidioidomycosis	In endemic areas: fluconazole or itraconazole
CD4 T-cell count <100 cells/μL	*Toxoplasma gondii*	TMP-SMX or dapsone plus pyrimethamine plus leucovorin or atovaquone plus pyrimethamine plus leucovorin
	Histoplasmosis	In endemic areas: itraconazole
CD4 T-cell count <50 cells/μL	*Mycobacterium avium* complex (MAC)	Macrolide (clarithromycin or azithromycin) or rifabutin
	Cryptococcosis	In endemic areas: fluconazole or itraconazole
Purified protein derivative (PPD) >5 mm induration or recent tuberculosis (TB) contact but no active TB and no history of treatment for active or latent TB	*Mycobacterium tuberculosis*	Isoniazid (INH) + pyridoxine for 9 months; if unlikely to complete 9 month course and on highly active antiretroviral therapy (HAART): rifabutin plus pyrazinamide for 2 months
Contact with chickenpox or shingles in varicella-zoster seronegative individuals	Varicella-zoster	Varicella-zoster immunoglobulins (VZIGs)
Human immunodeficiency virus (HIV)–infected	*Streptococcus pneumoniae*	Pneumovax
	Meningococcus—for youth attending the military or college and consider for unvaccinated adults	Menactra
Negative anti-hepatitis B core antibodies (HBc) and previously unimmunized or underimmunized to hepatitis B	Hepatitis B	Recombivax-HB or Energerix-B
Negative anti-hepatitis A serology	Hepatitis A	Havrix

*For additional information see the current US guidelines at http://AIDSinfo.nih.gov.
From Baliga CS, Shearer WT. HIV/AIDS. In: Fireman P, ed. Atlas of allergy. 3rd ed. Philadelphia, PA: Elsevier Science (USA); 2005: p. 351–67.

that AIDS-defining illnesses begin to appear. On the basis of the CD4 T-cell count, prophylaxis for opportunistic infections is administered (Table 39.1). Without treatment, most patients will succumb to opportunistic infections within 2 years of developing AIDS.

Long-Term Nonprogressors/Elite Controllers

Some HIV-infected patients are called *long-term nonprogressors* (LTNPs) because they do not progress to AIDS after a defined period. One explanation for this phenomenon is the deletion of 32 nucleotides in the *CCR5* gene (*CCR5-δ32)*. This deletion, postulated to have originated from selective pressure exerted by the bubonic plague, is found in people of European Caucasian ancestry. The deletion renders the CCR5 receptor nonfunctional and offers protection against viral infection. Other mutations have been identified, such as the *CCR2-64I* mutation, the *SDF1-3'A* mutation, and the *RANTES-28 G* mutation; however, these mutations are found in only a minority of LTNPs.

Elevated levels of messenger RNA (mRNA) for the ARV cytidine deaminase APOBEC3G have been found in LTNPs compared with noninfected controls and most patients with HIV infection. An explanation for the finding of APOBEC3G in LTNPs is that these individuals are infected with a defective or less fit virus. More conventional mechanisms have been proposed for LTNPs, such as strong and broadly neutralizing antibody and/or cytotoxic T-cell responses. HLA-B27 and HLA-B57 have also been correlated with LTNPs. In children, CD8 HLA-DR$^+$ T-cell percentages of <5% at 1–2 months of age have been implicated in predicting LTNPs, presumably as a result of decreased T-cell activation and inflammation. In a subset of LTNPs, circulating autoantibodies to a conformational epitope in the extramembrane loop of CCR5 that induce downregulation of the coreceptor have been found. Interestingly, if the levels of those antibodies waned, the LTNPs would be more likely to progress to symptomatic disease.

TABLE 39.2 **Factors Affecting Human Immunodeficiency Virus (HIV) Disease Progression**

Inoculum size	Higher the inoculum the faster the disease progression
Age	Infected infants have a higher risk of rapidly developing acquired immunodeficiency syndrome (AIDS) at any given CD4 percentage Less common in adults
Viral set point	The higher the viral set point the faster the disease progression
Broad and robust cellular and humoral immune responses	These correlate with lower set points, but what parts of these and which one is more important are unknown
Coreceptor mutations: δ32 CCR5, CCR2-64 I, SDF 1-3' A	Homozygotes for the δ32 CCR5 mutation fail to develop disease; heterozygotes have a much longer period of clinical latency
T-helper 1 (Th1) responses: interleukin-2 (IL-2) and interferon (IFN)-γ levels	Found in subsets of long-term nonprogressors (LTNPs)
Autoantibodies to CCR5 HLA-B27, -B57, -DLR	Found in subsets of LTNPs to also correlate with Autoantibodies
APOBEC3G levels	Higher levels are associated with slower progression
Viral fitness	Less fit viruses are associated with slower disease progression

A rare group of HIV-infected individuals are able to control the viral load at very low levels without ART. These elite controllers display high expression of the T-bet transcription factor perforin, and granzyme B.[22] Table 39.2 summarizes factors that can affect HIV progression.

DIAGNOSIS AND MONITORING OF HIV INFECTION

Diagnostic Tests

The diagnosis of HIV infection depends on the detection of biological markers, produced either by the virus or by the infected host, which appear in a chronology that is typically consistent among individuals. The earliest virological markers, HIV RNA and HIV p24, can be measured 9–12 days and 14–19 days after infection, respectively. The period between infection and detection of HIV RNA is called the "eclipse period" (Fig. 39.5).

Enzyme Immunoassays

Various generations of screening enzyme immunoassays (EIAs) are able to detect anti-HIV antibody 2–6 weeks after HIV RNA is detected. Traditionally, the period between infection and detection of HIV antibody by a particular test is referred to as the "window period" for that assay. The window period for new generation combination EIAs has been shortened to about 2 weeks because in addition to detecting antibody to both HIV-1 and HIV-2, they also detect p24 antigen.[23]

Rapid HIV EIA Tests

Rapid enzyme immunoassays are the most commonly used HIV screening tests. Rapid HIV tests are self-contained EIAs that detect HIV in whole blood or oral fluid specimens and can be used at the point of care, yielding results within 30 minutes. Because results of rapid tests are available quickly, individuals can learn of their results in a single appointment.

For both EIAs and rapid tests, a negative result is conclusive and generally requires no follow-up testing. A common exception is that in individuals with possible recent exposure, antibody tests may be negative in the window period before seroconversion, so testing should be repeated within 1–2 months. A reactive EIA or rapid test requires further testing to confirm the diagnosis (Fig. 39.6). The Western blot assay identifies specific HIV antigens to which IgG antibodies in the patient's serum react and historically has been used to confirm positive screening tests results. However, because Western blots do not become diagnostically reactive until approximately 5–6 weeks after infection, nucleic acid amplification tests are now commonly used to confirm infection because they detect HIV infection well before the Western blot can show positive results. The current laboratory HIV testing algorithm as recommended by the CDC for serum or plasma specimens is shown in Fig. 39.6.

Nucleic Acid Amplification Tests

Nucleic acid amplification tests are used to detect HIV RNA or DNA in biological samples to diagnose and/or monitor HIV

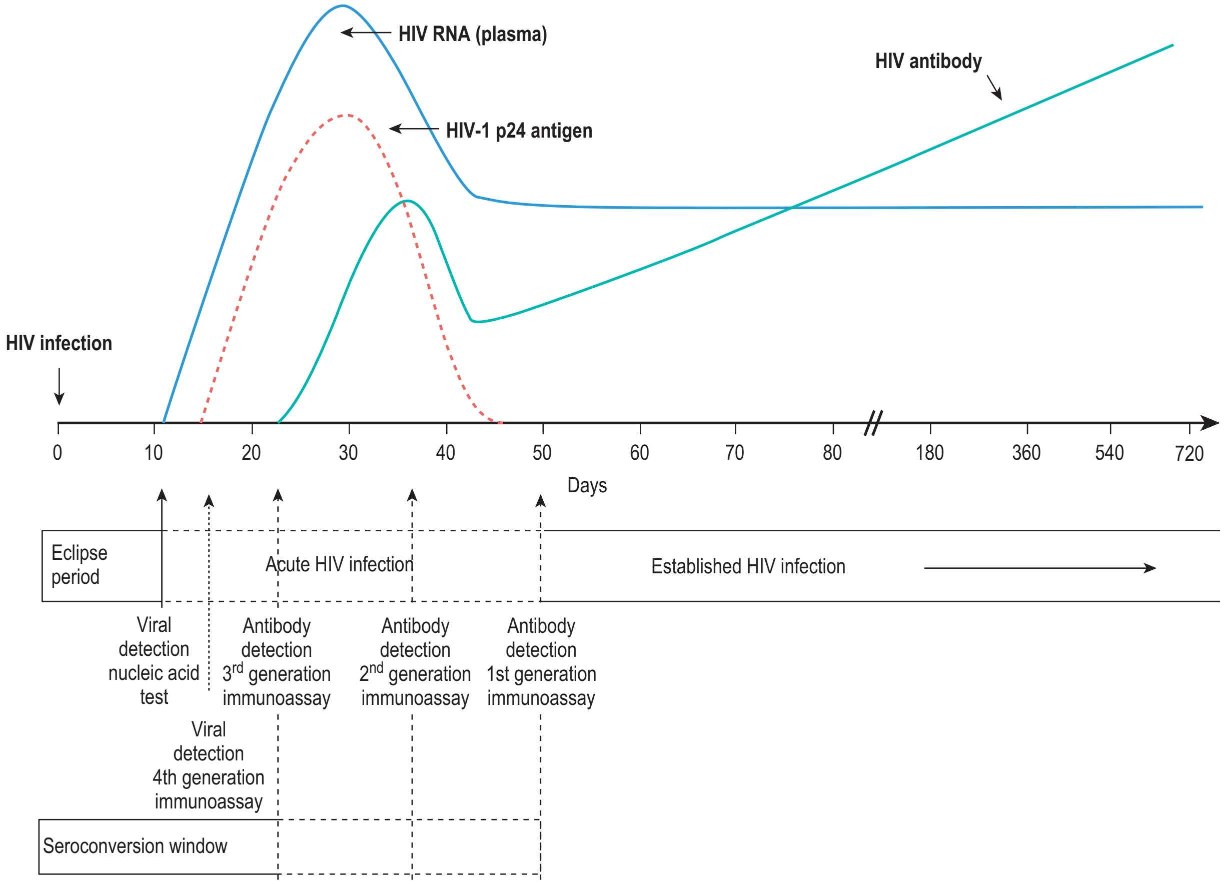


FIG 39.5 Sequence of appearance of laboratory markers for human immunodeficiency virus type 1 (HIV-1) infection. (From: Centers for Disease Control and Prevention and Association of Public Health Laboratories. Laboratory testing for the diagnosis of HIV infection: Updated recommendations. Published June 27, 2014. Available at: http://dx.doi.org/10.15620/cdc.23447. Accessed March 28, 2016.)

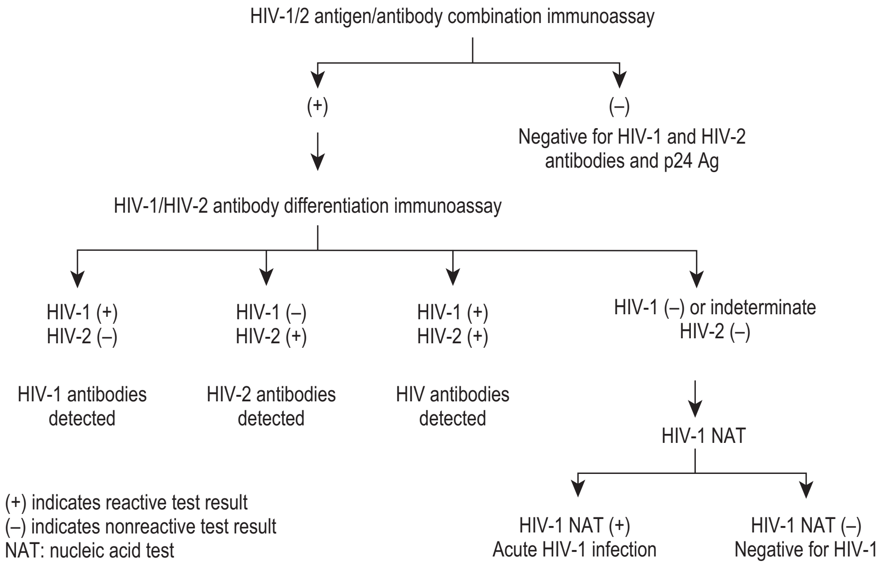

HIV-1/2 antigen/antibody combination immunoassay
(+)
(–)
Negative for HIV-1 and HIV-2 antibodies and p24 Ag
HIV-1/HIV-2 antibody differentiation immunoassay
HIV-1 (+) HIV-2 (–)
HIV-1 antibodies detected
HIV-1 (–) HIV-2 (+)
HIV-2 antibodies detected
HIV-1 (+) HIV-2 (+)
HIV antibodies detected
HIV-1 (–) or indeterminate HIV-2 (–)
HIV-1 NAT
HIV-1 NAT (+) Acute HIV-1 infection
HIV-1 NAT (–) Negative for HIV-1
(+) indicates reactive test result
(–) indicates nonreactive test result
NAT: nucleic acid test

FIG 39.6 Recommended laboratory human immunodeficiency virus (HIV) testing algorithm for serum or plasma specimens. (From http://stacks.cdc.gov/view/cdc/23447, p. 7.)

infection. HIV RNA polymerase chain reaction (PCR; Chapter 96) is used to quantify the virus load or the extracellular viral RNA in plasma and is also measured by using qualitative assays. HIV RNA detection plays a valuable role in identifying early infection before seroconversion and in confirming reactive screening tests. Quantification of HIV RNA is also used to monitor the effectiveness of ART in suppressing viral replication. HIV DNA PCR is a qualitative assay used to detect HIV viral DNA in peripheral blood mononuclear cells. Detecting HIV DNA has allowed the early diagnosis of HIV in perinatally exposed infants and is still used for this purpose in international settings. HIV DNA PCR is not typically available for diagnostics in the United States, although more and more physicians are ordering this test.

Monitoring Tests

Once infection is confirmed, specific laboratory tests are performed at baseline and then periodically to monitor disease status and progression, to inform treatment decisions, and to identify end-organ toxicity. The CD4 T-cell count is used to assess immune function and is the most important factor in the discussion of short-term prognosis and determining whether prophylactic medications for opportunistic infections are needed. Once a patient is on ART, an increasing CD4 T-cell count also helps confirm the efficacy of the therapy.[24] In children younger than 5 years of age, CD4 percentage is preferred because it typically remains stable in the setting of age-related changes in absolute CD4 count in this age group.[25]

HIV Viral Load

Particularly for patients on ART, plasma HIV RNA (viral load) is the most important indicator of response to therapy. Optimal viral suppression is generally defined as a viral load persistently below the level of detection (<20–75 copies/mL, depending on the assay used) and is usually achieved in 12–24 weeks of effective ART.[24] Failure to achieve maximal viral suppression or detectable virus after a period of maximal suppression may indicate virological failure attributable to drug resistance.

KEY CONCEPTS

Other Human Immunodeficiency Virus (HIV) Tests

- *Genotyping:* helps to guide the choice of antiretroviral medications in patients with resistant virus by sequencing the viral genetic code and identifying mutations that confer resistance to specific agents or classes of agents.
- *Phenotyping:* similar information to genotyping but not widely used; based on growing engineered viruses with a patient's virus' genes in the presence of antiretrovirals to determine their resistance.
- *Viral fitness:* this least utilized assay is based on the phenotype assay except that it grows the virus without antiretrovirals to measure its intrinsic replicative capacity relative to a defined wild-type virus.

All three of these assays present data as if there is only one virus strain in the body; in reality, there are numerous viral strains at any one time, with many more archived in cells; these assays detect the dominant strain in the circulation, neglecting the other strains, which may account for up to 20% of the circulating viral particles. Given this limitation, when genotyping results are utilized for changing therapy, old treatment regimens and previous genotype results must be taken into consideration.

Drug Resistance: HIV Genotype Versus Phenotype

Viral resistance to ARV agents can be assessed by either HIV genotype or phenotype assays. Genotype assays, based on PCR and genomic sequencing, identify the presence of key mutations that confer anti-HIV drug resistance. Phenotype assays assess the ability of HIV to replicate *in vitro* in the presence of ARV agents. The assay is performed by isolating certain key regulatory genes from HIV, usually protease and reverse transcriptase, inserting them into standardized viral constructs containing an indicator cassette, and infecting cell lines in the presence of ARV agents. The results are compared against control viral isolates and expressed as a fold-change in viral susceptibility. Phenotyping assays remain very expensive, and large studies have failed to conclusively prove a clinical advantage of phenotype assays over genotype assays. The phenotype assay quantifies susceptibility and is typically used by experts evaluating individuals who have accumulated resistance and failed multiple regimens.

Testing for Viral Tropism and Abacavir Hypersensitivity

Additional assays should be performed before the initiation of specific ARV medications. A viral tropism assay should be performed before initiation of a CCR5 antagonist. HLA-B*57:01 testing is indicated before initiation of abacavir, as this HLA phenotype is associated with abacavir hypersensitivity in 5–8% of patients early in the course of treatment.

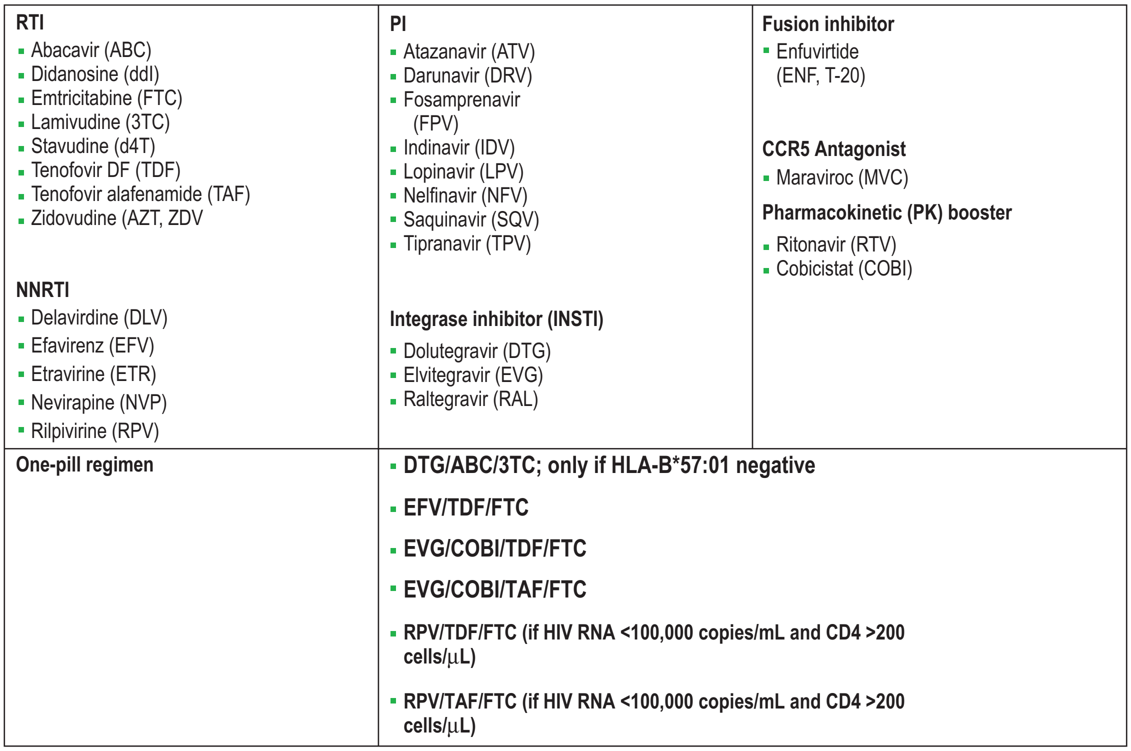
TREATMENT

Antiretroviral Therapy: Attacking HIV's Lifecycle

Combinations of ARV medications are used to maximally inhibit HIV replication and to reduce HIV-associated morbidity and mortality. Combination ART refers specifically to a combination of at least three ARV medications inhibiting the HIV lifecycle (see Fig. 39.3; Fig. 39.7). Current US guidelines recommend initiation of ART-naïve persons with a combination of two nucleoside reverse transcriptase inhibitors (NRTIs) plus an integrase strand transfer inhibitor, a nonnucleoside reverse transcriptase inhibitor (NNRTI), or a protease inhibitor (PI) with a pharmacokinetic enhancer (cobicistat or ritonavir). In ART-experienced patients, modification of ARV regimens and use of other classes of ARVs is guided, in part, by consideration of a number of factors, including viral resistance patterns, potential side effects, available medication formulations, pill burden, frequency of dosing, tolerability, short-term and long-term adverse event profiles, desire for pregnancy, and desire to preserve subsequent treatment options.[24,25]

When to Start Therapy

There are significant data to support treating all HIV-infected individuals with ART and starting treatment as soon as possible after diagnosis to reduce morbidity and mortality and to prevent transmission of HIV (see also HIV prevention). Randomized controlled trials have demonstrated that ART should be initiated in all patients with HIV infection, regardless of disease stage. The urgency to initiate ART is greatest for patients with lower CD4 counts, in whom the absolute risk of opportunistic infections, non-AIDS morbidity, and death is highest. However, the START (Strategic Timing of Anti-Retroviral Therapy) and the French National Agency for Research on AIDS and viral hepatitis (ANRS) 12136 "Temprano" trials provide the evidence that morbidity and mortality is reduced in individuals with CD4 counts >500 cells/mm^3 at the start of ART, resulting in strong recommendations, from both the US HIV treatment panel and the World Health Organization (WHO), to initiate ART in all patients regardless of CD4 cell count. In the START trial—a large, multinational, randomized controlled clinical trial with 4685 participants—ART-naive adults with CD4 counts >500 cells/mm^3 were randomized to initiate ART soon after randomization or to wait to initiate ART until their CD4 counts declined to <350 cells/mm^3, or until they developed a clinical indication for therapy. Serious AIDS or non-AIDS events endpoint was reported in 42 (1.8%) early ART participants and in 96 (4.1%) deferred ART participants (hazard ratio [HR] 0.43, favoring early ART [95% confidence interval (CI), 0.30–0.62; p < 0.001]). The majority (59%) of clinical events in the delayed ART arm occurred in participants whose CD4 counts were still >500 cells/mm^3, evidence for a benefit of immediate ART even before CD4 count declines below this threshold. The TEMPRANO ANRS 12136 study—a study of more than 2000 participants in Cote d'Ivoire, who were randomized to immediate or deferred ART—showed the benefit of immediate treatment over deferred treatment in reducing a combination of all-cause deaths, AIDS diseases, non-AIDS malignancies, and non-AIDS invasive bacterial diseases,



RTI ▪ Abacavir (ABC) ▪ Didanosine (ddI) ▪ Emtricitabine (FTC) ▪ Lamivudine (3TC) ▪ Stavudine (d4T) ▪ Tenofovir DF (TDF) ▪ Tenofovir alafenamide (TAF) ▪ Zidovudine (AZT, ZDV **NNRTI** ▪ Delavirdine (DLV) ▪ Efavirenz (EFV) ▪ Etravirine (ETR) ▪ Nevirapine (NVP) ▪ Rilpivirine (RPV)	**PI** ▪ Atazanavir (ATV) ▪ Darunavir (DRV) ▪ Fosamprenavir (FPV) ▪ Indinavir (IDV) ▪ Lopinavir (LPV) ▪ Nelfinavir (NFV) ▪ Saquinavir (SQV) ▪ Tipranavir (TPV) **Integrase inhibitor (INSTI)** ▪ Dolutegravir (DTG) ▪ Elvitegravir (EVG) ▪ Raltegravir (RAL)	**Fusion inhibitor** ▪ Enfuvirtide (ENF, T-20) **CCR5 Antagonist** ▪ Maraviroc (MVC) **Pharmacokinetic (PK) booster** ▪ Ritonavir (RTV) ▪ Cobicistat (COBI)
One-pill regimen	▪ **DTG/ABC/3TC; only if HLA-B*57:01 negative** ▪ **EFV/TDF/FTC** ▪ **EVG/COBI/TDF/FTC** ▪ **EVG/COBI/TAF/FTC** ▪ **RPV/TDF/FTC (if HIV RNA <100,000 copies/mL and CD4 >200 cells/µL)** ▪ **RPV/TAF/FTC (if HIV RNA <100,000 copies/mL and CD4 >200 cells/µL)**	

FIG 39.7 Antiretroviral (ARV) medications approved by the U.S. Food and Drug Administration (FDA).

even among participants who had baseline CD4 counts >500 cells/mm^3. Earlier ART initiation appears to increase the probability of restoring normal CD4 counts, a normal CD4/CD8 ratio, and lower levels of immune activation and inflammation, as summarized in the US adult treatment guidelines.[24] Individuals initiating ART during the first 6 months after infection also appear to achieve lower immune activation levels and better immune function during ART-mediated viral suppression compared with those who delay therapy.[25]

Recommendations for initiating therapy in children have always been relatively more aggressive compared with those for adolescents and adults because most infected children were infected perinatally and experience rapid disease progression.[26] The decision to initiate ART should always include consideration of a patient's comorbid conditions, willingness, and readiness to initiate therapy. ART may be deferred because of clinical or psychosocial factors, with the goal of initiating therapy as soon as possible once barriers to successful treatment have been reduced.

Antiretroviral Agents

There are more than 20 approved ARV drugs that are classified into six classes based on their chemical structure or the viral lifecycle step that they inhibit (see Fig. 39.7). The classes of ARV agents currently available include NRTIs, NNRTIs, PIs, fusion inhibitors, integrase inhibitors, and CCR5 antagonists.

Reverse Transcriptase Inhibitors, Protease Inhibitors, and Integrase Inhibitors

Modified versions of cellular nucleosides, NRTIs, once triphosphorylated *in vivo,* are incorporated into the proviral DNA by HIV reverse transcriptase (RT) and induce premature chain termination, thereby inhibiting successful conversion of the viral RNA to DNA. NNRTIs bind to RT and induce a conformational change such that RT is unable to bind with nucleotides. PIs act on viral protease, preventing the cleaving of the posttranslational viral polyproteins necessary for the maturation and infectivity of viral particles. Integrase inhibitors prevent strand transfer of viral DNA and thus block the incorporation of the completed HIV DNA copy into the host-cell DNA.

Fusion Inhibitors, CCR5 Blockers, and Low-Molecular-Weight Inhibitors

Fusion inhibitors and CCR5 antagonists inhibit HIV entry into host cells. Fusion inhibitors bind to viral gp41 and block the conformational changes necessary to induce fusion of the viral particle with the host cell. CCR5 antagonists bind to the CCR5 chemokine coreceptor on host cells, inducing a conformational change that impedes CCR5 interaction with HIV gp120, thereby preventing HIV entry into host cells.

Low-molecular-weight inhibitors of HIV binding to the CD4 molecule are a new approach. Epigallocatechin gallate, by blocking gp120 binding to the CD4 molecule, has been shown to inhibit the infections of CD4 T cells in culture by wild-type viruses, setting the stage for a phase I/II study in humans (Fig. 39.8).[27]

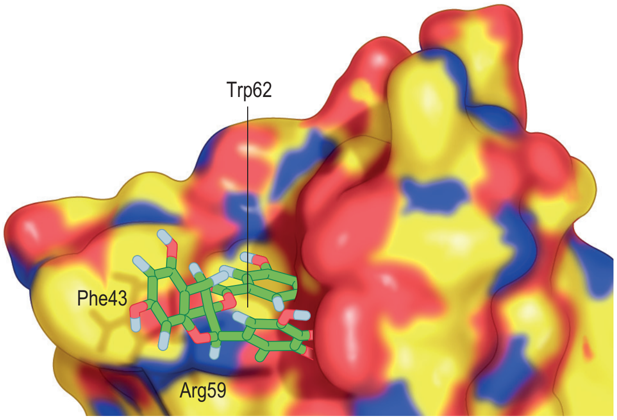


FIG 39.8 Epigallocatechin gallate (EGCG) binding to CD4. Recent studies indicate that polyphenolic EGCG, a component of green tea, can bind to the human immunodeficiency virus (HIV) binding site of CD4 (KD = 10 nM, as measured by nuclear magnetic spectroscopy) and potentially interfere in viral infection by interaction with three key amino acids in the D1 domain of the CD4 molecule (tryptophan 62, phenylalanine 43, and arginine 59). (From Williamson MP, et al. Epigallocatechin gallate, the main polyphenol in green tea, binds to the T-cell receptor, CD4. J Allergy Clin Immunol 2006; 118: 1369–74.)

IMMUNORECONSTITUTION AFTER THERAPY

Return of T Cells: Memory T Cells, Then Naïve T Cells

To varying degrees, the immune system is able to recover following initiation of therapy, a process called *immunoreconstitution.*[28] Upon start of ART in patients who are compliant and able to tolerate the regimen, the initial CD4 T-cell count is the best predictor of a successful outcome. Rapid reduction in the viral load, often to an undetectable level, is one of the earliest changes following initiation of ART, reflecting the ability of combination ART to rapidly suppress viral replication. Lagging behind the drop in viral load is the rise in CD4 T-cell values. An initial increase in circulating cells occurs in 3–6 months as a result of a decrease in immune activation and subsequent migration of memory T cells (CD4$^+$, CD45RO$^+$) out of the lymphoid compartment. A more gradual rise in total CD4 T cells occurs over the course of 3–5 years with the appearance of new naïve (CD4$^+$, CD45RA$^+$, CD62L$^+$) and memory T cells. Interestingly, a substantial minority of patients never reach a normal level of CD4 T cells but, instead, reach a plateau at lower levels. Primary drug prophylaxis and some secondary drug prophylaxis for opportunistic infections may be discontinued in patients once the CD4 T-cell count reaches >200 cells/mm^3 and is maintained for more than 3–6 months. Cellular and humoral responses to most pathogens also recover with rising CD4 T-cell counts. Of interest, a low CD4 T-cell count at the time of initiating therapy predicts a poor response to bacterial vaccines even after recovery of CD4 T-cell levels, suggesting a lag in the return of naïve CD4 T cells.

CLINICAL PEARLS

Immunoreconstitution Inflammatory Syndrome (IRIS)

- IRIS is typically found 2–3 weeks after initiating antiretroviral therapy (ART).
- Patients often become profoundly ill and require hospitalization.
- Steroids are sometimes useful in the treatment.
- Clinicians may prevent IRIS by initiating ART only after treating opportunistic infections.

Immune Reconstitution Inflammatory Syndrome

Immune reconstitution inflammatory syndrome (IRIS) is a well-known, if incompletely understood, response in patients with AIDS after initiating ART.[29] IRIS is characterized by an acute paradoxical worsening of inflammatory symptoms of treated opportunistic infections or the unmasking of previously subclinical, untreated infections related to the recovery of immune responses to opportunistic pathogens. IRIS occurs within weeks of ART initiation as the memory and effector antigen-activated CD4 T-cell population recovers. A recent systematic review found that IRIS developed in 13% of patients after initiation of ART.[28] The most predictive risk factor for the development of IRIS was a low CD4 T-cell count at the start of ART, with the incidence of IRIS increasing exponentially as the CD4 T-cell count declined. IRIS develops more commonly in patients with cytomegalovirus (CMV) retinitis, cryptococcal meningitis, progressive multifocal leukoencephalopathy, and tuberculosis. Studies reported that as many as 4% of patients with IRIS died, but the proportion was much higher if the syndrome was associated with cryptococcal meningitis.[29]

Hyperallergenic State Associated With Immunoreconstitution

Another complication possibly associated to IRIS is the appearance of asthma in children who were perinatally infected with HIV and received combination ART since infancy.[30] This condition may be mediated by CD4 T-cell activation, release of Th2-type cytokines, and loss of Tregs and tolerance. In support of this concept, Gingo et al.[31] reported at least a 20% prevalence of asthma in adults with HIV infection compared with that of 8.8% in the general population. In a subsequent pediatric study in which pulmonary function testing was objectively measured by spirometry, the following findings emerged: Nonreversible obstructive pulmonary disease was present in youth who had been perinatally infected with HIV and possibly in those exposed to HIV but were uninfected. This pulmonary disorder contains elements of asthma and chronic obstructive pulmonary disease (COPD) and closely resembles the asthma–COPD overlap syndrome. Chronic infections and immune dysregulation appears to play a significant role in this complication of HIV infection.[32]

PREVENTION

Prevention of Mother-to-Child Transmission

More than 90% of children living with HIV worldwide were infected through mother-to-child transmission during pregnancy, around the time of birth, or through breastfeeding.[33] Efforts to prevent this transmission hold the most promise in reducing the number of children infected with HIV, and these efforts include *(i)* early identification of HIV infection in pregnant women through routine antenatal testing; *(ii)* provision of ARV medications to both the pregnant woman and her infant[34,35]; *(iii)* delivery by elective cesarean section, when indicated; *(iv)* complete avoidance of breastfeeding when safe and sustainable alternatives are available; *(v)* widespread availability of educational programs addressing HIV infection; *(vi)* prevention of HIV perinatal transmission; and *(vii)* HIV counseling and testing services.[35,36]

Prevention of Sexual Transmission

Several biomedical interventions have the potential for radically changing the patterns and rates of HIV transmission. These include male circumcision and expanded use of ART in infected individuals to prevent ongoing infection (treatment as prevention) or prophylactically in uninfected individuals either before or after potential exposure to HIV to prevent acquisition of infection (preexposure prophylaxis [PrEP] and postexposure prophylaxis).

Male Medical Circumcision

The penile foreskin contains HIV-susceptible cells and is a potential portal of viral entry. Randomized controlled trials in several African countries have indicated that male medical circumcision reduces the risk of heterosexually acquiring HIV infection by 50–60%.[37] The WHO recommends male circumcision as part of a comprehensive HIV prevention package; however, condom use and other prevention modalities remain important in HIV prevention. There was only a modest benefit for the female partners of the circumcised men. In fact, in some circumstances, the risk of transmission to women from men with HIV infection who have undergone circumcision may be increased, perhaps through exposure to infected blood attributed to resumption of sexual activity before the circumcision site has fully healed.[37]

Preexposure Prophylaxis

Recently, two different modalities of ART have been found to decrease acquisition of HIV when used by uninfected individuals before their exposure to the virus as PrEP: topically applied ARV medications and systemic ART. Topically applied ARVs have been shown to block the acquisition of HIV in women from their partners with HIV infection by 35–54%. The Centre for the AIDS Program of Research in South Africa (CAPRISA 004) trial was the first double-blind randomized controlled trial demonstrating the efficacy of vaginally inserted gel containing the ARV medication tenofovir disoproxil fumarate, compared with a placebo gel, in preventing HIV transmission from men with HIV infection to their uninfected partners. Vaginal PrEP represents a promising intervention for preventing heterosexual transmission of HIV to women. This method may be particularly effective because women are able to initiate and control its use to protect themselves from infection.

Prophylactic Antiretroviral Therapy

ARVs taken prophylactically have also been shown to be effective in reducing the transmission of HIV both in MSM and in heterosexual individuals and couples. The multinational Preexposure Prophylaxis Initiative trial demonstrated that daily oral combination therapy with emtricitabine and tenofovir disoproxil fumarate (FTC-TDF) reduced the incidence of HIV among MSM by 44% compared with placebo. Daily oral FTC-TDF was also shown to reduce HIV incidence among heterosexual individuals by 63–78% and by 73% among heterosexual discordant couples. Despite these successful trials, preliminary results of one PrEP trial have failed to show similar efficacy for preventing HIV acquisition among heterosexual women.

Expanded Treatment With Antiretroviral Therapy

The expanded use of ART in infected individuals has been shown to reduce HIV transmission to uninfected partners. Both "Treatment as Prevention" and "Test and Treat" strategies involve the use of ART in individuals who do not otherwise meet the criteria for ART initiation with the express purpose of reducing transmission to others. In both approaches, ART is initiated regardless of CD4 count or viral load to reduce the viral load in the genital secretions of infected partners and thereby reduce HIV transmission to uninfected partners. The efficacy of the "Treatment as Prevention" model is demonstrated in the multinational HIV Prevention Trials Network (HPTN 052) clinical trial that examined the effectiveness of ART to prevent the sexual transmission of HIV in serodiscordant couples. Serodiscordant couples (1763 in number) were randomly assigned to have the infected partner either start ART immediately upon enrollment or to defer ART until immunological or clinical criteria were met. Of 28 genetically linked infections that occurred during the trial, only 1 infection occurred in couples assigned to receive immediate treatment, representing a 96% reduction in the risk of HIV transmission. There were also fewer morbidity and mortality events in the early-treatment group, suggesting a therapeutic benefit from early treatment as well.

The fundamental paradigm for "Test and Treat" programs is the same as that for "Treatment as Prevention" programs. "Test and Treat," however, emphasizes the need for universal voluntary routine HIV testing and initiating ART immediately in those found to be positive regardless of CD4 count or viral load. Although the approach remains controversial, with concerns regarding drug resistance, increased sexual risk taking in treated individuals, and societal cost-effectiveness, this approach may, over time, prove to be an effective prevention modality.

HIV VACCINES: CLINICAL TRIALS

Preventive Vaccines

The production of an effective HIV vaccine has been thwarted by the extreme rate of mutation in the virion and the sequestration of the virus in impenetrable reservoirs, predominantly the nonreplicating CD4 T cell. More than 30 HIV vaccines have been tested in human trials, including those with recombinant env gp120 proteins with adjuvants, HIV DNA plasmids, viral vectors, and prime-boost designs.[38] These vaccines have, for the most part, yielded disappointing results. Phase III trials of VAX003 in Thailand (AIDSVAX B/E plus alum) and VAX004 in North America and Europe (AIDSVAX B/E plus alum) showed no protection against HIV infection. Naked HIV DNA with cytokines IL-15, IL-12, and chemokine receptor 005 with and without electroporation has not been effective. Pox virus HIV vaccines in the STEP trial of North and South Americas and the Phambili trial in South Africa (Ad5-HIV trivalent vaccine) were ineffective in stopping HIV entry or postinfection viremia. However, the phase III trial RV144 in Thailand (ALVAC-HIV vCP1521 + AIDSVAX gp120 B/E) showed possible protection against HIV infection in heterosexual men and women.[38]

In addition to the goal of developing an HIV vaccine that elicits neutralizing antibodies, the search for a vaccine that stimulates a protective CD8 cytotoxic T-cell response continues. Association of certain major histocompatibility complex (MHC) molecules with HIV disease progression is clearly linked to the cytotoxic T-cell responses. Unfortunately, the HIV vaccine-induced cytotoxic CD8 T-cell response is insufficient to halt the progression of acute or chronic HIV disease. This was clearly indicated in the STEP clinical trial, in which the CD8 T-cell effects were the same between HIV-infected vaccinees and sham HIV-infected vaccinee controls. Qualities of effector and central memory CD8 T cells that would be protective include *(i)* production of cytotoxic cytokines (*e.g.,* IFN-γ and IL-2); *(ii)* rapidly replicating capacity; *(iii)* cytotoxic potential; *(iv)* high affinity for HIV antigens; *(v)* inhibition of HIV replication; *(vi)* recognition of specific HIV epitopes restricted by protective HLA-B antigens; *(vii)* central memory cells with long life spans; and *(viii)* rapid-attack memory cells at mucosal HIV entry sites. A high-level collaboration of preeminent clinicians and scientists has proposed the modification of the partially successful RV144 vaccine with the goal of producing a new HIV vaccine that will broadly neutralize HIV and variants of HIV that emerge under selective pressure (Table 39.3 and Table 39.4).[39] These modifications include changes in the viral epitopes, vaccine adjuvants, and use of a different clade as the construct of the virus. The goal of this new proposal is to use the information of many previous HIV trials to produce an ideal HIV vaccine that will prevent the spread of HIV infection in children and adults.

Therapeutic Vaccines

A therapeutic vaccine is one in which the vaccine is used after infection occurs, aiming to induce antiviral immunity to alter the course of disease. This would be accomplished by controlling viremia or reducing the viral set point in infected patients. Primate models suggest that just such a result is possible, especially with cellular immunity–inducing vaccines. To date, however, data from human studies have not shown any conclusive benefit in using therapeutic vaccines alone. Using a therapeutic vaccine in combination with ART is another approach currently under investigation. A small study of a therapeutic vaccine in 25 individuals with HIV infection produced a 1 $\log_{10}$ reduction in viral load compared with placebo-treated individuals with HIV infection.[40] However, HIV disease progression was not seen with this therapeutic vaccine in limited follow-up, indicating the need for longer follow-up and larger clinical trials before any measure of success can be claimed.

Future for HIV Vaccines

The general pessimism that prevailed in the scientific community after the failure of the STEP vaccine trial is gradually being replaced by cautious optimism as a result of the limited success of the RV144 clinical trial in Thailand. In addition, there is general belief that more basic research exploring vaccine design and

TABLE 39.3 Post RV144 Vaccine Combinations

Improve Frequency, Magnitude, and Polyfunctionality of CD4$^+$ T-Cell Response to HIV-1 Envelope Products	Improve Clade-Specific Antibody Responses to gp120	Improve Durability of Antibody Response Through Adjuvants
ALVAC (ZM96) +/–DNA	Bivalent clade C recombinant gp120 protein (strains 1086 and TV-1) construct	Alum, MF59, AS01$_B$
NYVAC (ZM96) +/– DNA		
Ad26+/– MVA	gp140 clade C trimeric protein +/– gp140 mosaic trimeric protein	Alum, AS01$_B$

Taken with permission from Corey et al. Immune correlates of vaccine protection against HIV-1 acquisition. Sci Transl Med 2015; 7: 310rv7. Review.

TABLE 39.4 **Comparison of Immune Responses That Influence VE in RV 144 Versus Past and Ongoing Human Immunodeficiency Virus (HIV) Vaccine Trials**

Immune Response Identified as a Correlate in RV144	Responses in Non-RV144 Efficacy Trials	Responses in Vaccines in Current Development
Total IgG to V1V2 scaffold	Lower in HVTN 505 compared with RV144	Higher titers in DNA/NYVAC/gp120 than RV144. Correlates of protection in heterologous NHP challenge studies with Ad26/Ad35 and Ad26/MVA studies
Serum IgA to gp120 (higher IgA = lower VE)	Higher IgA (including IgA to A1ConEnv) in HVTN 505, compared with RV144	Administration of gp120 at onset of priming with NYVAC and DNA/NYVAC markedly lowers serum IgA to gp120 DNA/MVA priming has lower IgA
IgG3 to V1V2	Lower in HVTN 505 and VAX003 than RV144	Under evaluation
ADCC activity	Minimal ADCC in HVTN 505	ADCC correlates with protection in NHP using Ad26 +/– trimeric gp120 High ADCC in DNA/MVA regimen
Tier 1 nAbs	Higher frequency in RV144 compared with HVTN 505	Clade C regions under study
High avidity to gp120	Not measured in HVTN 505 program Env IgG avidity with low IgA correlated with decreased risk of infection	DNA/MVA containing regimens have high avidity. Other products under study
CD4+ T cells with polyfunctional response	Different cytokine profile in HVTN 505 vs. RV144	DNA/NYVAC and Ad26/MVA increase prevalence and magnitude of Env specific CD4+ T cells

ADCC, antibody-dependent cellular cytotoxicity; Ig, immunoglobulin; nAbs, neutralizing antibodies.
Taken with permission from Corey et al. Sci Transl Med 2015 Oct 21;7(310):310rv7. doi:10.1126/ scitranslmed.aac7732.

trials in animals will lead to important clues for human study. Illustration of clinical trials in antiretroviral-treated macaques may be applied to humans in future by Byrareddy et al and Nishimura et al. Monoclonal antibody specific for CD4 T cell surface integrin ($\alpha4\beta7$) disrupts cellular trafficking of CD4 T cells with gastrointestinal tissue mucosal vascular addressin cell adhesion molecule (MAdCAM1). The CD4 T cell counts remained steady, CD8 T cell immunity sharply increased, and HIV replication became undetectable for up to 2 years.[41,42] As illustrated by these HIV vaccine studies in animals, understanding the immune correlates of vaccine efficacy is an absolute requirement for judging the success of an HIV vaccine.

ON THE HORIZON

- Production of newer antiretroviral drugs, including those used for postexposure prophylaxis with greater specificity for interrupting events in the viral lifecycle and with fewer side effects for patients.
- Investigation of new microbicidal drugs that can be safely applied before exposure for protection against human immunodeficiency virus (HIV) transfer.
- Development of preventive and therapeutic vaccines for HIV/acquired immunodeficiency syndrome (AIDS) that induce strong viral neutralizing antibody power and strong CD8 T-cell cytotoxic responses.
- Testing of gene construct-modified autologous hematopoietic stem cells capable of halting HIV replication.

TRANSLATIONAL RESEARCH NEEDS AND CONCLUSIONS

The quest for better ARV therapeutic agents continues, with the goal of greater selectivity and fewer side effects. However, the eradication of HIV/AIDS can only be approached practically with a preventive vaccine that elicits strong HIV-neutralizing ability and generates a strong cytotoxic CD8 T-cell response specific for HIV antigens. The approach to cure with gene therapy is perhaps the most sophisticated translational research venture. Using viral vectors to place gene constructs within nuclear DNA that prevent HIV replication is the goal of such research. Arguably, the most advanced form of this genetic engineering to halt HIV replication is the zinc finger endonuclease approach to disrupt specific genes necessary for the lifecycle of HIV. Adoptive transfer of autologous zinc finger–treated stem cells with infinite replication capacity may be an attractive future for individuals already infected with HIV.[43] The extraordinary experiment of HLA-matched and CCR5-δ35 deletion of hematopoietic stem cell immunoreconstitution of an patient with HIV infection ("Berlin Man") is a proof of concept of molecular and genetic engineering to cure HIV infection, a technique totally impractical to the millions of patients with HIV infection worldwide.[44] Nevertheless, this "one in a million" chance experiment has demonstrated the survival advantage of lymphocytes that cannot become infected with HIV.

HIV, a type 1 retrovirus, contains merely 9 genes, but those 9 genes have so far thwarted all scientific efforts toward finding a cure of its infection in humans. Optimism is warranted, however, because of the enormous knowledge base the study of HIV has generated in understanding the many arms of innate and adaptive immunity protecting humans and the promise of a curative treatment or vaccine for HIV. The HIV/AIDS pandemic has also brought the sobering realization that other new and potentially deadly pathogens may yet emerge to strike at humanity. Perhaps no other disease has caused so much to be learned so fast. In the developed world, HIV causes chronic infection, rather than certain death, thanks, in large part, to the use of ARV drugs. More novel drugs are in development as a result of the new-found understanding of the molecular biology of HIV. HIV continues to perplex and fascinate virologists and immunologists, and it continues to teach humility to clinicians.

ACKNOWLEDGMENTS

Supported in part by NIH Grants HD052102, AI069536 AI36211, and AI082978, and the Pediatric AIDS Fund of Texas Children's Hospital.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Joint United Nations Programme on HIV/AIDS (UNAIDS). Fact Sheet 2015. [3/29/2016]. Available at: http://www.unaids.org/en/resources/campaigns/HowAIDSchangedeverything/factsheet.
2. UNICEF. Children & AIDS; 2015 Statistical update. New York: UNICEF; 2015.
3. Centers for Disease Control and Prevention. HIV Surveillance Report. Atlanta, GA: CDC; 2014. p. 26.
4. Centers for Disease Control and Prevention. HIV among pregnant women, infants and children. 2016. [3/29/2016]. Available at: http://www.cdc.gov/hiv/pdf/group/gender/pregnantwomen/cdc-hiv-pregnant-women.pdf.
5. Richer LE, Halperin J. Acute human immunodeficiency virus infection. Am J Med Sci 2013;345:136–42.
6. Keele BF, Estes JD. Barriers to mucosal transmission of immunodeficiency viruses. Blood 2011;118:839–46.
7. Maartens G, Celum C, Lewin SR. HIV infection: epidemiology, pathogenesis, treatment, and prevention. Lancet 2014;384:259–71.
8. Chevalier MF, Weiss L. The split personality of regulatory T cells in HIV infection. Blood 2013;121:29–37.
9. Martin AR, Siliciano RF. Progress toward HIV eradication: Case reports, current efforts, and challenges associated with cure. Annu Rev Med 2016;67:215–28.
10. Bonsignori M, Zhou T, Sheng Z, et al. Maturation pathway from germline to broad HIV-1 neutralizer of a CD4-mimic antibody. Cell 2016;165:449–63.
11. Centers for Disease Control and Prevention. HIV and injection drug use. 2015. [3/29/2016]. Available at: http://www.cdc.gov/hiv/pdf/risk/cdc-hiv-idu-fact-sheet.pdf.
12. De Cock KM, Fowler MG, Mercier E, et al. Prevention of mother-to-child HIV transmission in resource-poor countries: translating research into policy and practice. JAMA 2000;283:1175–82.
13. Meier A, Alter G, Frahm N, et al. My D88-dependent immune activation mediated by human immunodeficiency virus type 1-encoded toll- like receptor ligands. J Virol 2007;81:8180–91.
14. Gringhius SI, van der Vlist M, van den Berg LM, et al. HIV-1 exploits innate signaling by TLR8 and DC-SIGN for productive infection of dendritic cells. Nat Immunol 2010;II:419–28.
15. Heigele A, Joas S, Regensburger K, et al. Increased susceptibility of CD4+ T cells from elderly individuals to HIV-1 infection and apoptosis is associated with reduced CD4 and enhanced CXCR4 and FAS surface expression levels. Retrovirology 2015;12:86.
16. Campbell GR, Bruckman RS, Chu YL, et al. Autophagy induction by histone deacetylase inhibitors inhibits HIV type 1. J Biol Chem 2015;290:5028–40.
17. Tomescu C, Abdulhaqq S, Montaner LJ. Evidence for the innate immune response as a correlate of protection in human immunodeficiency virus (HIV)-1 highly exposed seronegative subjects (HESN). Clin Exp Immunol 2011;164:158–69.
18. Scully E, Alter G. NK Cells in HIV Disease. Curr HIV/AIDS Rep 2016;13:85–94.
19. Singh KK, Qin M, Brummel SS, et al. Killer cell immunoglobulin-like receptor alleles alter HIV disease in children. PLoS ONE 2016;11:e0151364.
20. Zhao L, Gao J, Li Y, et al. Disrupted homeostatic cytokines expression in secondary lymph organs during HIV infection. Int J Mol Sci 2016;17(3):pii: E413.
21. Fidler S, Fox J. Primary HIV infection: a medical and public health emergency requiring rapid specialist management. Clin Med (Lond) 2016;16:180–3.
22. Hersperger AR, Martin JN, Shin LY, et al. Increased HIV-specific CD8+ T-cell cytotoxic potential in HIV elite controllers is associated with T-bet expression. Blood 2011;117:3799–808.
23. Branson BM. The future of HIV testing. J Acquir Immune Defic Syndr 2010;55:S102–5.
24. Panel on Antiretroviral Guidelines for Adults and Adolescents. Guidelines for the use of antiretroviral agents in HIV-1-infected adults and adolescents. Bethesda, MD: NIH Office of the AIDS Research Advisory Council; 2011.
25. Violari A, Cotton M, Gibb D, et al. Early antiretroviral therapy and mortality among HIV-infected infants. N Engl J Med 2008;359:2233–44.
26. Panel on Antiretroviral Therapy and Medical Management of HIV-Infected Children. Guidelines for the use of antiretroviral agents in pediatric HIV infection. Available at: http://aidsinfo.nih.gov/contentfiles/lvguidelines/pediatricguidelines.pdf. Accessed 2 July 2016.
27. Nance CL, Siwak EB, Shearer WT. Preclinical development of the green tea catechin, epigallocatechin gallate, as an HIV-1 therapy. J Allergy Clin Immunol 2009;123:459–65.
28. Azzoni L, Barbour R, Papasavvas E, et al. Early ART results in greater immune reconstitution benefits in HIV-infected infants: working with data missingness in a longitudinal dataset. PLoS ONE 2015;10:e0145320.
29. Boulougoura A, Sereti I. HIV infection and immune activation: the role of coinfections. Curr Opin HIV AIDS 2016;11:191–200.
30. Siberry GK, Leister E, Jacobson D, et al. Increased risk of asthma and atopic dermatitis in perinatally HIV-infected children and adolescents. Clin Immunol 2012;142:201–8.
31. Gingo MR, Wenzel SE, Steele C, et al. Asthma diagnosis and airway bronchodilator response in HIV-infected individuals. J Allergy Clin Immunol 2012;129:708–14.
32. Shearer WT, Jacobson D, Yu W, et al. Long-term pulmonary complications of perinatally hiv-infected youth. J Allergy Clin Immunol 2017;pii: S0091-6749(17):30336–6.
33. White AB, Mirjahangir JF, Horvath H, et al. Antiretroviral interventions for preventing breast milk transmission of HIV. Cochrane Database Syst Rev 2014;(10):CD011323.
34. Townsend CL, Byrne L, Cortina-Borja M, et al. Earlier initiation of ART and further decline in mother-to-child HIV transmission rates, 2000-2011. AIDS 2014;28:1049–57.
35. Panel on Treatment of HIV-Infected Pregnant Women and Prevention of Perinatal Transmission. Recommendations for use of antiretroviral drugs in pregnant HIV-1-infected women for maternal health and interventions to reduce perinatal HIV transmission in the United States. 2016. Available at: https://aidsinfo.nih.gov/guidelines/html/3/perinatal-guidelines/0. Accessed 3 May 2016.
36. Luzuriaga K, Mofenson LM. Challenges in the elimination of pediatric HIV-1 infection. New Engl J Med 2016;374:761–70. PubMed PMID: 26933850.
37. Tobian AA, Kacker S, Quinn TC. Male circumcision: a globally relevant but under-utilized method for the prevention of HIV and other sexually transmitted infections. Annu Rev Med 2014;65:293–306.
38. Huang Y, Follmann D, Nason M, et al. Effect of rAd5-vector HIV-1 preventive vaccines on HIV-1 acquisition: a participant-level meta-analysis of randomized trials. PLoS ONE 2015;10:e0136626.
39. Corey L, Gilbert PB, Tomaras GD, et al. Immune correlates of vaccine protection against HIV-1 acquisition. Sci Transl Med 2015;7:310rv7.
40. Mylvaganam GH, Silvestri G, Amara RR. HIV therapeutic vaccines: moving towards a functional cure. Curr Opin Immunol 2015;35:1–8.
41. Byrareddy SN, Arthos J, Cicala C, et al. Sustained virologic control in SIV+ macaques after antiretroviral and α4β7 antibody therapy. Science 2016;354(6309):197–202.
42. Nishimura Y, Gautam R, Chun TW, et al. Early antibody therapy can induce long-lasting immunity to SHIV. Nature 2017;543(7646):559–63.
43. June CH, Levine BL. T cell engineering as therapy for cancer and HIV: our synthetic future. Philos Trans R Soc Lond B Biol Sci 2015;370:20140374.
44. Hutter G, Nowak D, Mossner M, et al. Long-term control by CCR5 Delta 32/Delta 32 stem-cell transplantation. N Engl J Med 2009;360:692–8.

MULTIPLE-CHOICE QUESTIONS

1. Globally the segment of the population with the highest human immunodeficiency virus (HIV) infection rate is that of the:
   A. Adult male
   B. Adult female
   C. Adolescent male
   D. Adolescent female

2. The rapid HIV diagnostic screening test (results in 30 minutes) most commonly used is:
   A. HIV RNA
   B. Enzyme immune assay
   C. HIV DNA
   D. Western blot

3. Worldwide, a cure for HIV/AIDS will most likely happen with which treatment?
   A. Therapeutic vaccines
   B. Combination antiretroviral therapy
   C. Preventive vaccines
   D. Male circumcision

4. The first lymphocyte subpopulation to return in the immunoreconstitution observed after antiretroviral treatment of immunosuppressed patients with HIV infection is:
   A. CD4 memory cells
   B. CD19 memory cells
   C. CD4 naïve T cell
   D. CD8 naïve T cell

40

Autoantibody-Mediated Phenocopies of Primary Immunodeficiency Diseases

Sarthak Gupta, Sarah K. Browne

Anticytokine autoantibodies have been identified in health and disease, with their role in pathogenesis ranging from none to directly causal. In fact, it is becoming increasingly recognized that neutralizing, high-titer autoantibodies cause a variety of potentially life-threatening illnesses. Their manifestations are diverse, and their clinical presentation is generally a reflection of the cytokine pathways that are rendered functionally deficient.

Examples include pulmonary alveolar proteinosis (PAP) caused by anti–granulocyte macrophage–colony-stimulating factor (GM-CSF) autoantibodies[1]; disseminated nontuberculous mycobacterial (NTM) and other opportunistic infections and anti–interferon-γ (IFN-γ) autoantibodies[2]; chronic mucocutaneous candidiasis (CMC) and anti–interleukin (IL)-17A, anti–IL-17F, or anti–IL-22 autoantibodies[3,4]; and staphylococcal skin infection and anti–IL-6 autoantibodies[5] (Table 40.1). In each of these cases a high-titer, biologically active autoantibody has been identified in association with a unique clinical syndrome. Although this list is not exhaustive, these diseases are of particular interest in that the clinical manifestations are often similar to those in patients with genetic defects in the same cytokine-associated pathways. The mendelian defects that confer a similar phenotype to their anticytokine autoantibody counterpart also provide a strong biological rationale for establishing the autoantibodies as causal. Although a number of other anticytokine–autoantibody associated syndromes such as pure red-cell aplasia (antierythropoietin autoantibodies)[6] and severe osteoporosis in celiac disease (antiosteoprotegerin autoantibodies)[7] have been described, this chapter focuses on those that increase susceptibility to infection.

KEY CONCEPTS

Anticytokine Autoantibody–Associated Primary Immunodeficiency Diseases

- Anticytokine autoantibodies have been identified in a group of patients with immunodeficiency diseases characterized by the development of high-titer, neutralizing autoantibodies to cytokines.
- The clinical manifestations in patients with anticytokine autoantibodies are similar to those seen in patients with genetic defects in the pathway of the target cytokine.
- Immunodeficiency syndromes described include those associated with anti– granulocyte macrophage–colony-stimulating factor (GM-CSF) autoantibodies and pulmonary alveolar proteinosis (PAP); anti–interferon (IFN)-γ autoantibodies and severe immunodeficiency; anti–interleukin (IL)-17A, anti–IL-17F, and anti–IL-22 autoantibodies and chronic mucocutaneous candidiasis (CMC); and anti–IL-6 autoantibodies and staphylococcal skin infection.
- Management of these syndromes involves treating the consequences of the autoantibody (*e.g.,* infectious manifestations, lung disease) or targeting the autoantibody itself.

There are numerous commonalities and distinctions worth highlighting among anticytokine autoantibody–associated immune deficiency. Shared in common is the presence of high-titer, neutralizing anticytokine autoantibodies of the immunoglobulin G (IgG) isotype. PAP caused by anti–GM-CSF autoantibodies appears to be primarily a lung disease largely resulting from disruption of GM-CSF–dependent pulmonary surfactant catabolism. *In vitro* studies using primary human cells and mouse work have demonstrated that anti–GM-CSF autoantibodies cause immune dysfunction, largely through transcription factor PU.1, which may explain pulmonary and extrapulmonary infections observed in PAP, particularly those opportunists known to be controlled by neutrophils or macrophages.[1] PAP caused by anti–GM-CSF autoantibodies is similar to severe immune deficiency caused by anti-IFN-γ autoantibodies in that both diseases have adult onset. On the basis of broad screening for other autoantibodies, it appears they make high-titer neutralizing autoantibody against only one cytokine and do not demonstrate an increased frequency of other autoantibodies or forms of autoimmunity.[1,2] Patients with autoantibodies against IL-17A, IL-17F, IL-22, and CMC are different in that, so far, all have had an underlying primary diagnosis of either autoimmune polyendocrinopathy with candidiasis and ectodermal dysplasia (APECED) or thymic neoplasia.[3,4] Thus in the case of APECED, a mendelian defect in the gene *AIRE,* which is responsible for negative selection for autoreactive T cells in the thymus, the onset of anticytokine autoantibodies and CMC occurs earlier in life. Occasionally, CMC is the initial presentation of thymoma (just as myasthenia gravis caused by antiacetylcholine receptor autoantibodies must prompt evaluation for thymoma) and highlights that for all anticytokine autoantibody–associated syndromes, the time line for development of anticytokine autoantibodies relative to the observation of clinical disease is largely unknown. Additionally, patients with thymoma and APECED can develop multiple anticytokine autoantibodies, although they have a proclivity for certain ones, including type I IFNs, IL-17A, IL-17F, IL-22, and, in the case of thymoma (but not APECED), IL-12. Given the presence of multiple neutralizing autoantibodies, and the role of the thymus in T-cell development that can lead to intrinsic T-cell defects that could also confer infection susceptibility, it is challenging to prove that a particular anticytokine autoantibody is the necessary and sufficient agent of disease pathogenesis. In the case of anti–IL-6 autoantibodies, high-titer, neutralizing anti–IL-6 IgG was demonstrated in one 11-month-old child with recurrent staphylococcal skin infections and low C-reactive protein-reactive protein (CRP), an IL-6–driven inflammatory marker.[5] The opportunity to identify common clinical features around this entity is limited, since other cases have yet to be identified.

TABLE 40.1 Anticytokine Autoantibody–Associated Syndromes and Their Phenotypically Similar Genetic Counterparts

Cytokine Target	Genetic Phenocopies	Clinical Manifestations	Laboratory and Radiological Manifestations	*In vitro* Evidence for Biological Activity	Infections	Comments
GM-CSF	GM-CSF receptor α or β subunits	PAP Insidious and progressive respiratory failure Newly described cases of isolated cryptococcal meningitis and *Nocardia* infection in individuals without human immunodeficiency virus (HIV) infection[17-19]	BAL fluid contains large foamy macrophages or monocyte-like macrophages and elevated levels of surfactant proteins Characteristic computed tomography (CT) of chest demonstrates ground-glass opacities with thickening of intralobular septae, "crazy paving." PFTs demonstrate restrictive and diffusion defects[1]	Anti–GM-cerebrospinal fluid autoantibodies inhibit pSTAT-5 production,[17] PU.1 expression , MIP-1α protein production	Pulmonary and extrapulmonary infections with *Nocardia, Aspergillus, Proteus, Histoplasmosis,* and *Cryptococcosis*[1]	Mutations in surfactant proteins have also been described SP-B, SP-C, or ABCA3, but are considered clinically distinct
IFN-γ^h	Mutations in IFN-γR1; IFN-γR2; STAT1; IL-12Rβ1; IL-12Rβ2; IL-12p40; NEMO; IRF8; ISG15	Chronic infections with intracellular pathogens, particularly lymphadenitis, skin, soft tissue, and bone infections; can be multiple organisms simultaneously or sequentially; reactive dermatoses; constitutional symptoms common	Elevated erythrocyte sedimentation rate, C-reactive protein-reactive protein (CRP), β_2 microglobulin, anemia, hypergammaglobulinemia. CT imaging may demonstrate abscess formation or osteomyelitis	Anti–IFN-γ autoantibodies inhibit pSTAT1 production(2, 11), IFN-γ inducible gene expression[29] and IL-12p70 and TNF-α protein production[2,11]	Nontuberculous mycobacteria, tuberculosis, nontyphoidal *Salmonella, Histoplasma, Penicillium, Cryptococcus, Burkholderia pseudomallei,* varicella-zoster virus[2]	Functional testing to evaluate downstream inhibitory effects of autoantibody important, (*i.e.,* plasma inhibition of IFN-γ–induced pSTAT1 production)
IL-17A IL-17F IL-22	IL-17RA IL-17F	Recurrent candidal infections of mucosal surfaces, nails, and skin Infection may become resistant to antifungals	None	Anti–IL-17 autoantibodies inhibit IL-17–induced IL-6[4]	*Candida* spp.	Reported only in the context of APECED and thymoma so far
IL-6	STAT3	Recurrent staphylococcal skin infections.	Undetectable CRP	Anti–IL-6 autoantibodies prevent IL-6–induced CRP mRNA[5]	*Staphylococcus aureus*	One case reported to date

GM-CSF, Granulocyte macrophage–colony stimulating factor; PAP, pulmonary alveolar proteinosis; BAL, bronchoalveolar lavage; PFT, pulmonary function test; pSTAT, phosphorylated signal transducer and activator of transcription test; MIP, macrophage inflammatory protein; SP, surfactant protein; ABC, adenosine triphosphate-binding cassette; IFN, interferon; NEMO, nuclear factor (NF)-κB essential modulator; IRF, interferon regulatory factor; TNF, tumor necrosis factor; APECED, autoimmune polyendocrinopathy with candidiasis and ectodermal dysplasia.

OVERVIEW OF PATHOPHYSIOLOGY

The pathophysiology of infection susceptibility is generally thought to result from a functional deficiency in the cytokine that is being neutralized (Chapter 9). It is believed that a high-titer autoantibody binds its respective cytokine target, thereby blocking downstream signaling and biological activity. For each anticytokine–autoantibody pair, it has been demonstrated that plasma or purified IgG from a patient with the anticytokine autoantibody prevents the activity of the targeted cytokine at the levels of signal transduction, gene transcription, and/or protein expression. In the case of anti–IFN-γ autoantibodies, it appears that antibody levels may track with disease activity[8]; however, for anti–GM-CSF autoantibodies, the results have been conflicting, and in neither case has the question been rigorously studied. It is also possible, but not yet proven, that antibody-binding avidity may influence the degree of disease severity as well. Thus it may be possible to have high-titer, lower avidity anticytokine autoantibodies leading to a similar disease phenotype to low-titer, high-avidity anticytokine autoantibodies.

The events that lead to the generation of anticytokine autoantibodies are poorly understood and are likely disease specific. Again, by comparing and contrasting these diseases, we may begin to understand some key factors. In the case of pulmonary alveolar proteinosis, it appears to be linked to tobacco use and also demonstrates a male predominance, which may be an artifact of the historically higher rates of smoking in men, particularly because no gender predilection has been observed among nonsmokers.[1] Although a large cohort of patients has been described in Japan, this disease is seen worldwide across all ethnicities and not within families, suggesting that if there is a genetic component, it is a complex one. No familial clustering has been identified in over 130 cases of anti–IFN-γ autoantibodies and opportunistic infection reported[2,9-12]; however, the disease is mostly seen in Asian-born Asians, suggesting that there may be an environmental trigger in the context of a common genetic background.

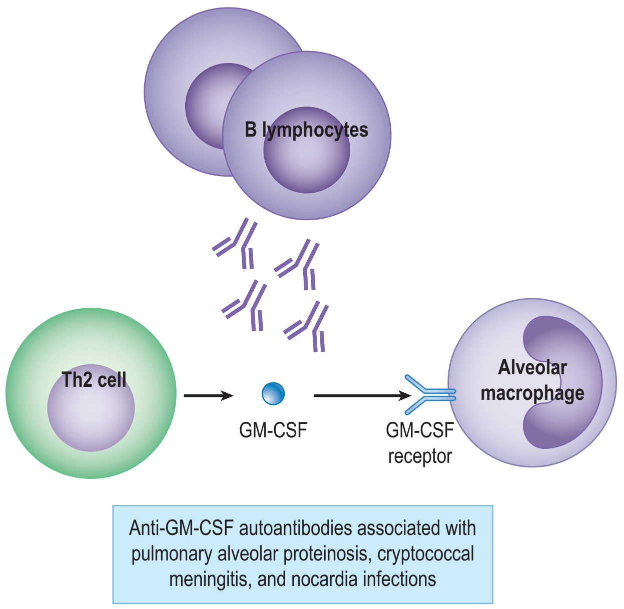
The fact that anticytokine autoantibodies are both IgG and high-affinity autoantibodies implicates the T-helper (Th) lymphocyte–dependent processes of class switching and affinity maturation. Interestingly, anti–IL-17A, –IL-17F, and –IL-22 autoantibodies appear directly linked to either the genetic AIRE deficiency of APECED or the acquired AIRE deficiency observed in patients with thymoma.[13] In both cases, thymic-driven disease appears to be leading to extensive B-lymphocyte dysregulation in the form of many autoantibodies, beyond just anticytokine autoantibodies. However, given that B cells may play a primary role in the development of autoimmunity in AIRE deficiency, the mechanisms underlying B-cell autoreactivity are likely complex.[13] Furthermore, evidence in rheumatological mouse models suggests that peripheral B-lymphocyte lineages leading to autoantibodies may fundamentally differ from those leading to development of protective antibodies.[14] Thus a common phenomenon of anticytokine–autoantibody production may, in fact, be a reflection of a convergence of multiple differing mechanisms.

ANTI–GM-CSF AUTOANTIBODIES AND PULMONARY ALVEOLAR PROTEINOSIS

GM-CSF is a hematopoietic stem cell (HSC) growth factor that binds the GM-CSF receptor, which is widely expressed on many cell lineages, including neutrophils, macrophage precursors, dendritic cells (DCs), erythrocyte progenitors, and megakaryocytes. The GM-CSF receptor is composed of two α and two β subunits, which together bind two GM-CSF molecules with high affinity and induce signal transduction and activator of transcription (STAT)–5 phosphorylation, nuclear translocation, and induction of master transcription factor PU.1. PAP was first described in 1958 by Rosen et al. as an idiopathic syndrome of respiratory failure, histopathologically characterized by alveolar filling with acellular periodic acid-Schiff–positive proteinaceous material.[1] The pathogenesis of PAP has since been linked to congenital or acquired defects in the GM-CSF signaling pathway.

The first clues to the etiological mechanism of PAP surfaced in 1994 and 1995 when a GM-CSF$^{-/-}$ and GM-CSF receptor β$^{-/-}$ mice, respectively, demonstrated pulmonary disease that was virtually identical to human PAP. Shortly thereafter, mechanisms involving disruption of GM-CSF signaling were linked to PAP in humans.[1] Primary PAP results from mutations in either the GM-CSF receptor subunits α or β and generally leads to severe respiratory failure and usually presents early in life. Autoimmune PAP results from neutralizing anti–GM-CSF autoantibodies, can also cause respiratory failure, and shares the same pulmonary histopathology as the primary form (Fig. 40.1). In contrast to primary PAP, the median age of diagnosis of the autoimmune form is 39 years, and its clinical course and severity is highly variable ranging from progressive respiratory decline to spontaneous resolution. A secondary form of PAP caused by qualitative or quantitative deficiency of alveolar macrophages, generally in the context of hematologic malignancies, iatrogenic immunosuppression, or inhaled toxins, has also been recognized.

Although the primary pathological process relates to impairment of GM-CSF–dependent catabolism of surfactant and its subsequent overaccumulation in pulmonary alveoli, there has



Anti-GM-CSF autoantibodies associated with pulmonary alveolar proteinosis, cryptococcal meningitis, and nocardia infections

FIG 40.1 Anti– Granulocyte Macrophage–Colony-Stimulating Factor (GM-CSF) Autoantibody Associated Pulmonary Alveolar Proteinosis (PAP). Autoimmune PAP results from impairment of GM-CSF–dependent catabolism of surfactant and its subsequent overaccumulation in pulmonary alveoli due to neutralizing anti–GM-CSF autoantibodies, leading to respiratory failure. Th2 cell, T-helper cell-2.

long been an association with opportunistic infections.[1] Beyond its role in pulmonary surfactant homeostasis, the GM-CSF receptor is widely expressed on immune cells, including neutrophils, monocytes, DCs, megakaryocytes, and erythrocyte progenitor cells, and influences cell differentiation, proliferation, and immune activation. GM-CSF has been shown in both humans and mice to facilitate not only terminal differentiation of monocytes to alveolar macrophages but also innate immune responses, primarily via induction of transcription factor PU.1. Cells isolated from bronchoalveolar lavage (BAL) fluid from patients with autoimmune PAP show decreased PU.1 messenger RNA (mRNA), which is thought to be central to the pathogenesis of PAP. PU.1 has been shown to regulate Toll-like receptor (TLR) signaling, adhesion, phagocytosis, and microbicidal activity, thus providing a mechanism for the increased infection susceptibility that is observed in PAP. Beyond the macrophage, defects have also been shown in neutrophil adhesion, phagocytosis, oxidative burst, and bacterial killing from both the blood of human patients with PAP and GM-CSF$^{-/-}$ mouse bone marrow.[15]

Patients with acquired PAP are vulnerable to typical respiratory pathogens as well as opportunistic infections. Although the high incidence of respiratory infections may be partially attributable to their underlying chronic lung disease, the opportunistic infections are generally caused by organisms controlled by macrophages, including *Nocardia, Histoplasma*,[1] nontuberculous mycobacteria,[16] and *Cryptococcus.*[17] Although Witty et al. reported eight cases of PAP and *Mycobacterium avium* complex cultured from BAL fluid,[16] these patients did not receive antimycobacterial drugs and appeared to fare no worse. Furthermore, most of these infections were described before the recognition of anti–GM-CSF autoantibodies as being causal, so these reports may be confounded by heterogeneity within the underlying diagnosis.

Extrapulmonary infections have been observed with some frequency, which could be ascribed to the systemic effects of anti–GM-CSF autoantibodies identified in serum. In support of this systemic effect, several case reports describe patients with PAP who had extrapulmonary infections, including cases of central nervous system (CNS) *Nocardia infection,* septic arthritis, perinephric abscess, and *Nocardia* dissemination to skin; CNS *Aspergillus* and *Proteus;* and disseminated histoplasmosis.[1] Recently, a few case series have reported opportunistic infections associated with high-titer neutralizing anti–GM-CSF autoantibodies without concurrent PAP. Anti–GM-CSF antibodies have been reported in previously healthy adults uninfected by HIV who had cryptococcal meningitis specifically caused by *Cryptococcus gattii.*[17,18] Neutralizing anti–GM-CSF autoantibodies were also found in five of seven patients with disseminated/extrapulmonary *Nocardia* infections.[19] However, it remains to be seen if these infections are the only manifestations in these patients or they may eventually develop PAP as seen in the two of seven cases of cryptococcal meningitis associated with anti–GM-CSF autoantibodies.[17] Why the same autoantibody may have different clinical phenotypes is unknown. Opportunistic infections have not been reported as a complication of congenital PAP. Reasons for this could include the extreme rarity of this condition or limited opportunity for exposure to environmental opportunists as a result of medical debilitation.

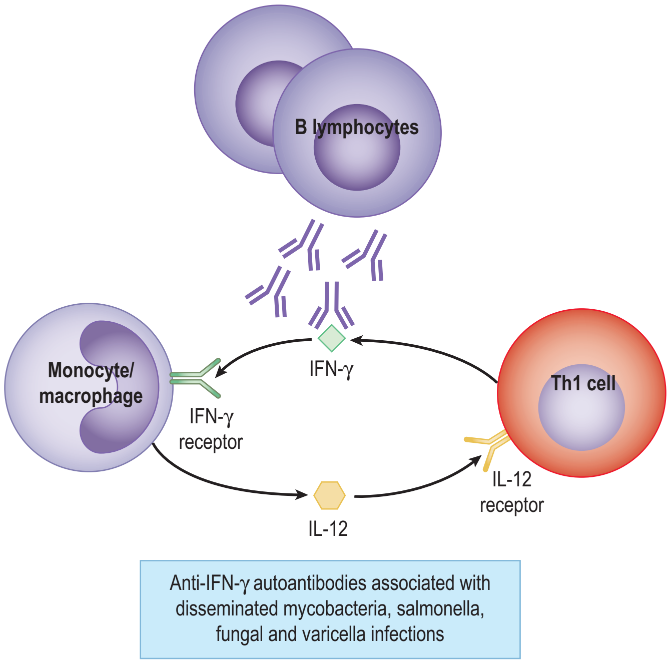
Treatment includes whole lung lavage to evacuate the proteinaceous material contained in the alveoli. Inhaled and subcutaneous GM-CSF has also been effective either by saturating the antibody or inducing tolerance.[20,21] In two large studies, one using subcutaneous GM-CSF and one using inhaled GM-CSF, a clinical response was not associated with a reduction in plasma or BAL concentrations of anti–GM-CSF autoantibodies, thereby providing possible support for the former mechanism.[20,21] B-cell–targeted therapy using the anti-CD20 chimeric monoclonal antibody (mAb) rituximab has been used to treat a small number of patients and has shown encouraging clinical results.[22]

ANTI–IFN-γ AUTOANTIBODIES AND SUSCEPTIBILITY TO INTRACELLULAR PATHOGENS

IFN-γ, produced predominantly by activated Th1 cells and natural killer (NK) cells, is central to host defense against intracellular pathogens (Chapter 26). The IFN-γ receptor (IFN-γR), expressed primarily on monocytes, is composed of two subunits in duplicate, IFN-γR1 and IFN-γR2. Binding of IFN-γ to its receptor leads to Janus kinase 2 (JAK2) and then JAK1 phosphorylation on the intracellular portions of IFN-γR2 and -1, respectively. Subsequent STAT-1 docking, phosphorylation, homodimerization, and nuclear translocation lead to transcription of IFN-γ responsive genes. Macrophage activation, differentiation, and elaboration of inflammatory mediators, such as tumor necrosis factor-α (TNF-α) and IL-12, ensue. Defects in the IFN-γ–IL-12 axis lead to mendelian susceptibility to mycobacterial disease as well as those caused by other intracellular pathogens, including listeriosis, salmonellosis, histoplasmosis, melioidosis, and penicilliosis.[23] The list of genetic mutations involving this pathway that result in increased susceptibility to mycobacteria or other intracellular pathogens continues to expand and, to date, includes mutations in IFN-γR1, IFN-γR2, STAT1, IL-12p40, IL-12Rb1, nuclear factor-κB (NF-κB) essential modulator (NEMO), IFN regulatory factor (IRF) 8, and IFN-stimulated gene (ISG) 15.[23] Neutralizing autoantibodies against IFN-γ represent an alternative mechanism by which the IFN-γ–IL-12 metabolic pathway is disrupted, with the first cases described in 2004[9] (Fig. 40.2). We reported 85 patients identified in a 6-month period,[2] with over 130 patients reported to date,[8-12] and this suggests that immunodeficiency caused by anti–IFN-γ autoantibodies is probably underappreciated.

The infections in patients with anti–IFN-γ autoantibodies mimic many of those seen in patients with inborn errors in the IL-12–IFN-γ signaling pathways and include mycobacterial, particularly nontuberculous environmental mycobacteria, nontyphoidal *Salmonella, Penicillium, Histoplasma, Cryptococcus, Burkholderia pseudomallei,* and additionally varicella-zoster virus (VZV), both dermatomal and disseminated.[2] Infections have been noted in all organ systems, although lymph nodes, skin, soft tissue, and bone appear to be preferentially affected. Up to 50% of patients develop sterile reactive dermatoses, most commonly neutrophilic dermatosis, but also erythema nodosum, pustular psoriasis, and exanthematous pustulosis.

Although patients with mendelian defects tend to present in childhood, all reported cases of patients with anti–IFN-γ were previously healthy adults, the vast majority of whom were Asian-born Asians. A recent association with human leukocyte antigen–antigen D related (HLA-DR)*15:02/16:02 and HLA-DQ*05:01/05:02 was shown in 78 patients with anti–IFN-γ autoantibodies,[24] further implicating a genetic predisposition to development of disease. However, no familial clustering has been observed, and Asians born outside of Asia have yet to be reported with this syndrome, suggesting complex genetics and possibly an environmental contribution to autoantibody pathogenesis. The recent report of an American-born, Caucasian female with this syndrome also adds to the complexity of its pathogenesis.[12]



Anti-IFN-γ autoantibodies associated with disseminated mycobacteria, salmonella, fungal and varicella infections

FIG 40.2 Anti-Interferon (IFN)γ Autoantibody Associated Susceptibility to Intracellular Pathogens. Neutralizing autoantibodies against IFN-γ disrupt the IFN-γ–IL-12 metabolic pathway. Infections in patients with anti–IFN-γ autoantibodies mimic many of the infections seen in patients with inborn errors in the interleukin (IL)-12–IFN-γ signaling pathways and include mycobacterial, nontyphoidal *Salmonella, Penicillium, Histoplasma, Cryptococcus, Burkholderia pseudomallei,* and varicella zoster virus. Th1 cell, T-helper cell-1.

Notable laboratory features include anemia of chronic disease and elevation in inflammatory markers, such as erythrocyte sedimentation rate (ESR), CRP, and β_2 microglobulin. Immunologically, patients commonly demonstrate polyclonal hypergammaglobulinemia, but they otherwise have normal lymphocyte phenotyping, including CD4 T lymphocytes, monocyte numbers, and IFN-γR1 expression.

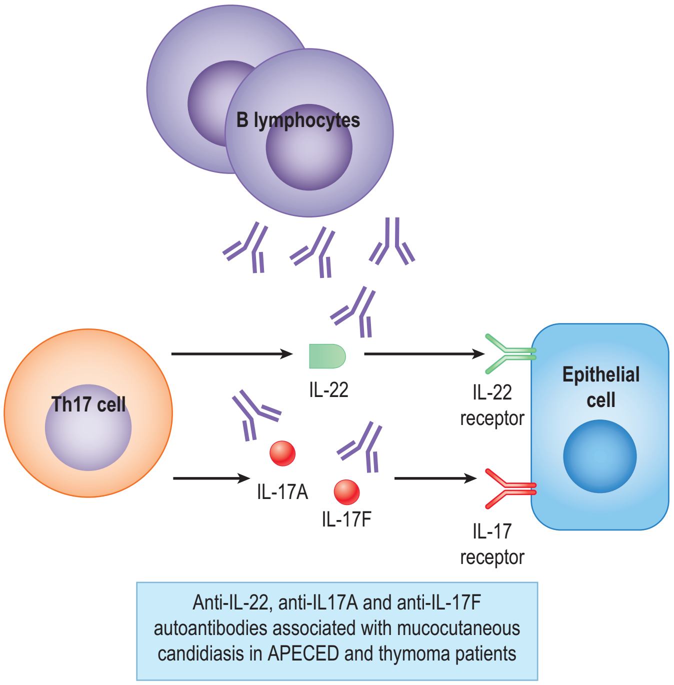
Treatment has focused mainly on managing the infections with targeted antimicrobials. Interestingly, a severe paradoxical inflammatory reaction similar to immune reconstitution syndrome, manifesting as lymphadenopathy, cavitary lung lesions, and lytic bone lesions during antituberculosis treatment, has been reported and should be distinguished from true failure of antimicrobial treatment.[25] In cases refractory to antiinfective agents, some have attempted to overcome the antibody with IFN-γ administration or drive down antibody levels with plasmapheresis and cyclophosphamide[26] or rituximab.[8] It is unclear what factors predict response to antimicrobials alone versus a need for further immunomodulation; the efficacy of these immune modulatory approaches remains to be evaluated formally in clinical trials.

ANTI–IL-17 AND ANTI–IL-22 AUTOANTIBODIES AND CHRONIC MUCOCUTANEOUS CANDIDIASIS

IL-17A and IL-17F are proinflammatory cytokines that can combine as either homodimers or heterodimers with each other (Chapter 9). These dimerized combinations signal via IL-17RA and IL-17RC heterodimeric receptor complexes, ultimately activating NF-κB. IL-22, produced by T lymphocytes and NK cells, also signals via a heterodimeric receptor composed of IL-22R1 and IL-10R2 subunits, expressed mainly on epithelial and other nonimmune cells. IL-17A/F and IL-22 cooperate to induce proinflammatory cytokines involved in granulopoiesis and neutrophil recruitment as well as antimicrobial peptides, such as β defensins and S100 proteins, which are thought to be important in mucosal immunity. Clinical evidence for a protective role of IL-17 in CMC arose from the observation of this infectious complication in diseases with varying degrees of Th17 impairment, including *STAT3*-deficient hyper-IgE syndrome (HIES, or Job's syndrome), dectin-1 deficiency, CARD9 deficiency, and, to a lesser degree, IL-12 receptor β_1 deficiency.[27] Strong support of this hypothesis came via instances of two families that demonstrated inherited susceptibility to mucosal candidiasis: one with an autosomal recessive mutation in IL-17RA, and another with an autosomal dominant negative mutation in IL-17F.[27]

APECED syndrome is another mendelian disorder associated with CMC. APECED, caused by mutations in the *AIRE* gene, leads to a clinical triad of hypoparathyroidism, adrenal insufficiency, and CMC.[13] Other endocrine glands, including gonads, thyroid, endocrine cells in the gut, and pancreatic islet cells, are variably affected. Many other autoimmune phenomena, including Sjögren syndrome, rheumatoid arthritis, hepatitis, keratitis, vitiligo, pernicious anemia, and alopecia, as well as autoantibodies to type I IFNs, have been described.[13] The biological consequences of these anti–IFN-α autoantibodies are unclear, but autoantibodies to other tissue antigens in autoimmune polyendocrinopathy syndrome type 1 (APS-1), such as autoantibodies to glutamic acid decarboxylase, thyroid peroxidase, and 21-hydroxylase, are clearly pathological. Providing explanation for the preponderance of autoimmunity seen in APECED was the observation that the transcriptional activity of AIRE in medullary thymic epithelial cells (mTECs) promotes expression of tissue-specific genes, thereby facilitating intrathymic destruction of autoreactive T cells and fostering self-tolerance.[13]

Unlike the above-described mendelian disorders, the mechanism of CMC in APECED is not directly linked to the genetic deficiency itself but, rather, to the strong genetic predisposition to autoimmunity, which, in this case, includes production of neutralizing anti–IL-17A, anti–IL-17F and anti–IL-22 autoantibodies[3,4] (Fig. 40.3). Puel et al. identified antibodies to IL-17A, IL-17F, or IL-22 in 33 patients with APECED, 29 of whom also had CMC, compared with healthy controls who had neither autoantibodies nor CMC.[4] Kisand et al. found anti–IL-17A, –IL-17F, or –IL-22 autoantibodies in up to 90% of 162 cases, also strongly associated with CMC.[3] They also identified anti–IL-17 and anti–IL-22 autoantibodies in two patients with thymoma and CMC, but in none of the 33 patients with thymoma who did not have CMC. There were a few instances of autoantibodies without CMC, and there are many examples of CMC independent of APECED, and this suggests that autoantibodies to IL-17 and IL-22 are not absolutely necessary for development of CMC. However, their identification, particularly in light of the CMC disease seen in patients with IL-17RA and IL-17F mutations, provides support for a causal relationship. Interestingly, thymoma tissue also demonstrates decreased AIRE expression, further linking these diseases beyond a shared tendency toward autoimmunity and anticytokine–autoantibody production. What is less clear however, is the exact role AIRE is playing in pathogenesis of organ-specific autoimmunity (including anticytokine autoantibody



Anti-IL-22, anti-IL17A and anti-IL-17F autoantibodies associated with mucocutaneous candidiasis in APECED and thymoma patients

FIG 40.3 **Anti–Interleukin (IL)-17 and Anti–IL-22 Autoantibodies Associated Chronic Mucocutaneous Candidiasis (CMC).** IL-17A/F and IL-22 induce proinflammatory cytokines involved in granulopoiesis and neutrophil recruitment as well as important mucosal antimicrobial peptides, such as β defensins and S100 proteins. Neutralizing anti–IL-17A, anti–IL-17F, and anti–IL-22 autoantibodies in autoimmune polyendocrinopathy with candidiasis and ectodermal dysplasia (APECED) syndrome is the likely mechanism of CMC in these patients.

formation), since the syndromes diverge considerably in this regard.[13]

ANTI–IL-6 AUTOANTIBODIES AND RECURRENT STAPHYLOCOCCAL SKIN INFECTION

IL-6 is produced by many immune and nonimmune cells, including B cells, T cells, macrophages, synovial cells, endothelial cells, and hepatocytes, and is involved in both acute and chronic inflammation, ranging from sepsis to rheumatoid arthritis. IL-6 binds a heterodimeric receptor composed of IL-6Rα and a shared receptor chain gp130. IL-6Rα confers ligand specificity, whereas gp130 mediates signal transduction. IL-6 regulates the acute phase response in the liver, with its hallmark induction of CRP and elevated ESR. Anti–IL-6 autoantibodies were identified in a Haitian boy who had two episodes of severe staphylococcal cellulitis, one complicating chickenpox infection, the other complicating mosquito bites.[5] Treatment included supportive care and antiinfective agents. Undetectable CRP, despite severe infection, suggested impairment in IL-6 activity. STAT3 is the critical signal transduction molecule for IL-6 and IL-10, and autosomal dominant mutations in *STAT3* also lead to recurrent staphylococcal skin abscesses,[28] suggesting a common potential mechanism for this phenotypic profile of infection susceptibility. The limitations to this association are that the infections in the Haitian boy resolved without any apparent change in the anti–IL-6 autoantibody titers, and this is the only reported case to date.

MANAGEMENT

Therapies directed at the syndromes of pathogenic autoantibodies have focused either on treating the disease consequences or targeting the autoantibody itself. Therapeutic BAL removes proteinaceous alveolar material from the lungs of patients with PAP,[1] and antimycobacterials treat the disseminated NTM infection in patients with autoantibodies to IFN-γ.[9,29] Approaches ranging from physical removal of the antibody, to immunomodulatory therapy, to induction of immune tolerance or suppression of the population of cells that produce the pathogenic autoantibodies have also been employed.

PAP and disseminated NTM associated with anti–IFN-γ autoantibodies have been treated with exogenous GM-CSF[20,21] and IFN-γ,[29] respectively, resulting in clinical improvement, although PAP has been studied more rigorously in this regard. Autoantibody levels after therapy were not routinely evaluated across diseases, nor are they performed in a standardized fashion, although, at least in the PAP cohorts, it appears they do not change in response to exogenous cytokine.[20,21]

In cases that are refractory to treatment, attempts have been made to alleviate the blockade by reducing anticytokine–autoantibody levels. One patient with anti–IFN-γ autoantibodies underwent plasmapheresis and cyclophosphamide therapy in addition to antimycobacterials.[26] Rituximab, the mouse–human chimeric mAb that targets the human B-cell marker CD20, is currently approved for treatment of B-cell lymphoma and rheumatoid arthritis. Since B cells ultimately differentiate into antibody-producing cells, it has been applied to a number of autoantibody-mediated diseases, such as myasthenia gravis (Chapter 65) and pemphigus vulgaris (a blistering skin disease caused by autoantibodies that recognize desmoglein 3, a keratinocyte cell-surface protein; Chapter 63). Successful rituximab therapy has also been reported in PAP[22] and immunodeficiency caused by anti–IFN-γ autoantibodies[8] with a specific reduction of anticytokine–autoantibody titers.

A new approach in pure red blood cell (RBC) aplasia caused by antierythropoietin autoantibodies is to bypass the autoantibody with a erythropoietin receptor synthetic peptide agonist peginesatide (Hematide; Affymax) that does not share homology with the eosinophil peroxidase (EPO) ligand.[30] Although this approach has not been used in cases of anticytokine autoantibodies and immune deficiency, it underscores the range of novel treatment approaches that might be explored for anticytokine autoantibody–associated syndromes.

CONCLUSIONS

Anticytokine autoantibodies have been identified in healthy normal adults as well as in those with different diseases, suggesting that their occurrence may range from a normal homeostatic mechanism, to epiphenomena, to being directly pathogenic. In PAP, the identification of anti–GM-CSF autoantibodies came over 40 years after the initial description of the syndrome, suggesting that other currently idiopathic diseases may someday be explained by the identification of neutralizing or agonizing autoantibodies. Furthermore, it was not intuitive that systemic autoantibodies to a hematopoietic growth factor should result in disease confined mainly to the lung. The identification of anti–GM-CSF autoantibodies in cryptococcal meningitis implicates the GM-CSF pathway in host defense of this infection, much as mendelian disorders have done for NTM and the

ON THE HORIZON

- Anticytokine autoantibodies are an emerging mechanism underlying the pathogenesis of immune deficiency in previously “healthy” adults.
- The clinical manifestations of these diseases, like pulmonary involvement in anti–granulocyte macrophage–colony-stimulating factor (GM-CSF) associated pulmonary alveolar proteinosis (PAP), may provide insight to signaling pathways and target cells of the involved cytokine.
- Recent discovery of many anticytokine autoantibody–associated diseases suggests that other idiopathic immunodeficiency diseases may be the result of an autoantibody affecting an as-yet unknown signaling pathway or cell type.
- Much research needs to be done on the management of patients with these syndromes, since besides just treating the symptomatic consequences of the autoantibody, the autoantibody or its generation itself can now be targeted.
- Profiling anticytokine autoantibodies in patients with immunodeficiency could help personalize their management by potentially predicting disease manifestations and optimizing treatment options.

IFN-γ–IL-12 pathway. The observation of high-titer, neutralizing anticytokine autoantibodies in an expanding number of diseases, beyond just those characterized by immune deficiency,[7] combined with the opportunity for novel therapeutic approaches to their diagnoses, mandates that their presence be not merely considered but rigorously sought.

ACKNOWLEDGEMENTS

This work was supported by the Divisions of Intramural Research of National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS) and National Institute of Allergy and Infectious Diseases (NIAID) of the National Institutes of Health (NIH), Bethesda, MD.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Seymour JF, Presneill JJ. Pulmonary alveolar proteinosis: progress in the first 44 years. Am J Respir Crit Care Med 2002;166:215–35.
2. Browne SK, Burbelo PD, Chetchotisakd P, et al. Adult-onset immunodeficiency in Thailand and Taiwan. N Engl J Med 2012;367:725–34.
3. Kisand K, Boe Wolff AS, Podkrajsek KT, et al. Chronic mucocutaneous candidiasis in APECED or thymoma patients correlates with autoimmunity to Th17-associated cytokines. J Exp Med 2010;207:299–308.
4. Puel A, Doffinger R, Natividad A, et al. Autoantibodies against IL-17A, IL-17F, and IL-22 in patients with chronic mucocutaneous candidiasis and autoimmune polyendocrine syndrome type I. J Exp Med 2010;207:291–7.
5. Puel A, Picard C, Lorrot M, et al. Recurrent staphylococcal cellulitis and subcutaneous abscesses in a child with autoantibodies against IL-6. J Immunol 2008;180:647–54.
6. Casadevall N, Dupuy E, Molho-Sabatier P, et al. Autoantibodies against erythropoietin in a patient with pure red-cell aplasia. N Engl J Med 1996;334:630–3.
7. Riches PL, McRorie E, Fraser WD, et al. Osteoporosis associated with neutralizing autoantibodies against osteoprotegerin. N Engl J Med 2009;361:1459–65.
8. Browne SK, Zaman R, Sampaio EP, et al. Anti-CD20 (Rituximab) therapy for anti-IFN-γautoantibody-associated nontuberculous mycobacterial infection. Blood 2012;119:3933–9.
9. Hoflich C, Sabat R, Rosseau S, et al. Naturally occurring anti-IFN-γautoantibody and severe infections with Mycobacterium cheloneae and Burkholderia cocovenenans. Blood 2004;103:673–5.
10. Kampitak T, Suwanpimolkul G, Browne S, et al. Anti-interferon-γautoantibody and opportunistic infections: case series and review of the literature. Infection 2011;39:65–71.
11. Patel SY, Ding L, Brown MR, et al. Anti-IFN-γautoantibodies in disseminated nontuberculous mycobacterial infections. J Immunol 2005;175:4769–76.
12. O’Connell E, Rosen LB, LaRue RW, et al. The first US domestic report of disseminated *Mycobacterium avium* complex and anti-interferon-γautoantibodies. J Clin Immunol 2014;34:928–32.
13. Kisand K, Peterson P. Autoimmune polyendocrinopathy candidiasis ectodermal dystrophy. J Clin Immunol 2015;35:463–78.
14. Huang H, Benoist C, Mathis D. Rituximab specifically depletes short-lived autoreactive plasma cells in a mouse model of inflammatory arthritis. Proc Natl Acad Sci USA 2010;107:4658–63.
15. Uchida K, Beck DC, Yamamoto T, et al. GM-CSF autoantibodies and neutrophil dysfunction in pulmonary alveolar proteinosis. N Engl J Med 2007;356:567–79.
16. Witty LA, Tapson VF, Piantadosi CA. Isolation of mycobacteria in patients with pulmonary alveolar proteinosis. Medicine (Baltimore) 1994;73:103–9.
17. Rosen LB, Freeman AF, Yang LM, et al. Anti-GM-CSF autoantibodies in patients with cryptococcal meningitis. J Immunol 2013;190:3959–66.
18. Saijo T, Chen J, Chen SC, et al. Anti-granulocyte-macrophage colony-stimulating factor autoantibodies are a risk factor for central nervous system infection by Cryptococcus gattii in otherwise immunocompetent patients. MBio 2014;5:e00912–14.
19. Rosen LB, Rocha Pereira N, Figueiredo C, et al. Nocardia-induced granulocyte macrophage colony-stimulating factor is neutralized by autoantibodies in disseminated/extrapulmonary nocardiosis. Clin Infect Dis 2015;60:1017–25.
20. Kavuru MS, Sullivan EJ, Piccin R, et al. Exogenous granulocyte-macrophage colony-stimulating factor administration for pulmonary alveolar proteinosis. Am J Respir Crit Care Med 2000;161:1143–8.
21. Tazawa R, Trapnell BC, Inoue Y, et al. Inhaled granulocyte/macrophage-colony stimulating factor as therapy for pulmonary alveolar proteinosis. Am J Respir Crit Care Med 2010;181:1345–54.
22. Kavuru MS, Malur A, Marshall I, et al. An open-label trial of rituximab therapy in pulmonary alveolar proteinosis. Eur Respir J 2011;38:1361–7.
23. Bustamante J, Boisson-Dupuis S, Abel L, et al. Mendelian susceptibility to mycobacterial disease: genetic, immunological, and clinical features of inborn errors of IFN-gamma immunity. Semin Immunol 2014;26:454–70.
24. Ku CL, Lin CH, Chang SW, et al. Anti-IFN-γautoantibodies are strongly associated with HLA-DR*15:02/16:02 and HLA-DQ*05:01/05:02 across Southeast Asia. J Allergy Clin Immunol 2016;137:945–8.e8.
25. Xie YL, Rosen LB, Sereti I, et al. Severe paradoxical reaction during treatment of disseminated tuberculosis in a patient with neutralizing anti-IFN-γ autoantibodies. Clin Infect Dis 2016;62:770–3.
26. Baerlecken N, Jacobs R, Stoll M, et al. Recurrent, multifocal Mycobacterium avium-intercellulare infection in a patient with interferon-γautoantibody. Clin Infect Dis 2009;49:e76–8.
27. McDonald DR. TH17 deficiency in human disease. J Allergy Clin Immunol 2012;129:1429–35, quiz 36–7.
28. Holland SM, DeLeo FR, Elloumi HZ, et al. STAT3 mutations in the hyper-IgE syndrome. N Engl J Med 2007;357:1608–19.
29. Kampmann B, Hemingway C, Stephens A, et al. Acquired predisposition to mycobacterial disease due to autoantibodies to IFN-gamma. J Clin Invest 2005;115:2480–8.
30. Macdougall IC, Rossert J, Casadevall N, et al. A peptide-based erythropoietin-receptor agonist for pure red-cell aplasia. N Engl J Med 2009;361:1848–55.

MULTIPLE-CHOICE QUESTIONS

1. High-titer, neutralizing autoantibodies against which cytokine is commonly seen in autoimmune polyendocrinopathy candidiasis ectodermal dystrophy (APECED) syndrome?
   - **A.** Interferon (IFN)-α
   - **B.** IFN-γ
   - **C.** Granulocyte macrophage–colony-stimulating factor (GM-CSF)
   - **D.** Interleukin (IL)-6

2. A 45-year-old female who moved from Taiwan to the United States at the age of 20 years presents with disseminated *Mycobacterium abscessus* with pulmonary and bone lesions and cervical lymphadenopathy. Autoantibodies to which cytokine may be associated with her disease?
   - **A.** IL-17
   - **B.** IL-22
   - **C.** IFN-α
   - **D.** IFN-γ

3. A previously healthy HIV-negative 34-year-old woman presented with few days of fever, headache, neck pain, and vomiting. She also reported progressively increasing exertional dyspnea since the past 6 months. Lumbar puncture showed increased opening pressure. Cerebrospinal fluid (CSF) analysis revealed increased white blood cell (WBC) count with lymphocytic predominance, low glucose, and elevated protein. The CSF cryptococcal antigen was positive, and the culture grew *Cryptococcus neoformans.* Brain magnetic resonance imaging (MRI) was normal. Chest computed tomography (CT) showed ground-glass opacities in both lungs. Bronchoalveolar lavage (BAL) revealed periodic acid-Schiff (PAS)–positive material on cytology, consistent with a diagnosis of pulmonary alveolar proteinosis (PAP). Autoantibodies against which cytokine may be associated with her manifestations?
   - **A.** GM-CSF
   - **B.** IL-6
   - **C.** G-CSF
   - **D.** Erythropoietin
   - **E.** IFN-γ

41

Immunological Mechanisms of Airway Diseases and Pathways to Therapy

David B. Corry, Farrah Kheradmand, Amber Luong, Lavannya Pandit

The immunological airway diseases comprise a large and disparate group of respiratory disorders characterized by airway and parenchymal inflammation that impairs sinus and lung function. The physiological importance of the airways, combined with the need to respond immunologically to an extremely broad range of particulate and gaseous aerosols, explains much of the diverse nature of airway immune disorders and their disproportionately large effect on human health. The allergic respiratory tract immune disorders covered in this chapter are among the most common of all human afflictions.

Allergic disorders have a common immune phenotype comprising highly characteristic cellular, humoral, biochemical, and molecular components, although individual variability means that not all these immunological features are equally expressed. The most studied and visually characteristic allergic immune cells are eosinophils and tissue mast cells, which are easily seen on conventional hematoxylin and eosin (H&E) staining of pathological specimens. Eosinophils may far outnumber all other inflammatory cells. Less obvious on histochemical staining, but equally important to allergic diseases, are B cells that secrete the antibody isotypes immunoglobulin E (IgE) and IgG4, and T-helper 2 (Th2) cells, which secrete a restricted repertoire of cytokines, including interleukin-4 (IL-4), -5, -9, -10, and -13, which coordinate and activate other allergic effector cells, such as innate lymphoid cells (ILCs).

The allergic airway diseases are typically chronic and occasionally fatal; although spontaneous remissions are not uncommon, the conditions are rarely curable. Recent insights into pathophysiological mechanisms have, however, opened up new prospects for improved therapy.

CLINICAL PRESENTATION OF ALLERGIC AIRWAY DISEASE

Although the diverse effector cells and molecules that characterize allergic inflammatory exudates may be seen at any point along the respiratory tract, the functional impact of allergic disease is quite different in the upper and lower airways.

Chronic Rhinitis and Rhinosinusitis

Epidemiology and Clinical Presentation

The major upper airway inflammatory disorders are rhinitis and chronic rhinosinusitis. Allergic rhinitis (AR) involves inflammation of the nasal mucosa: it is the most common type of chronic rhinitis affecting approximately 50 million Americans. In different surveys, AR accounts for 30–70% of chronic perennial rhinitis, with the balance being nonallergic. Typical symptoms of AR include clear rhinorrhea, sneezing, postnasal drip, nasal pruritus, and congestion, often in association with ocular symptoms, such as conjunctivitis and tearing. Based on symptom duration, AR is now classified as "intermittent" or "persistent," with severity of symptoms noted as "mild" or "moderate to severe."[1] Consequently, AR can be categorized into four groups based on the combination of above two categories of frequency and symptom severity. Patients with persistent AR typically have daily symptoms all year round but can also show seasonal fluctuations in the severity of their symptoms. The impact of symptoms on quality of life and on productivity is reflected in the "mild" and "moderate–severe" categorization.

KEY CONCEPTS

Classification of Allergic Rhinitis

Frequency of Symptoms	Severity of Symptoms
Intermittent	**Mild**
Symptoms present for <4 days a week	No presence of sleep disturbance
Or for <4 weeks	Or impairment of daily activities, leisure, and/or sport
	Or impairment of school or work
	Or troublesome symptoms
Persistent	**Moderate-Severe**
Symptoms present for >4 days a week	Presence of sleep disturbance
Or for >4 weeks	And/or impairment of daily activities, leisure, and/or sport
	And/or impairment of school or work
	And/or troublesome symptoms

In contrast to rhinitis, which involves inflammation and symptoms confined to the nasal passages and mucosa, rhinosinusitis (RS) affects both the sinuses and the nasal mucosa. Rhinosinusitis can be subdivided into acute and chronic rhinosinusitis (CRS), based on duration of symptoms. CRS is reported to affect approximately 29 million Americans and is defined by symptoms persisting for longer than 12 weeks. Clinically, CRS is subdivided on the basis of the presence or absence of nasal polyps. Within CRS with nasal polyps (CRSwNP), there are several recognized clinical subtypes, including allergic fungal rhinosinusitis (AFRS), aspirin-exacerbated respiratory disease (AERD), and cystic fibrosis.[2]

Signs and symptoms associated with RS in adults include facial pain and pressure, headaches, nasal congestion with or without obstruction, frontal or post nasal drainage, generalized malaise, loss of sense of smell (hyposmia), and cough. Of these, nasal obstruction is the most common (81–95%), followed by facial pain and pressure (70–85%), discolored nasal drainage (51–83%), and hyposmia (61–69%). In contrast, symptoms in children vary with age and require the parent or caregiver to recognize them. Young children often present with a chronic cough and irritability, rather than facial pain. Parents often also report the child has halitosis and purulent nasal discharge. Although it is more difficult to determine the prevalence of RS in children because of overlapping symptoms with AR and viral upper respiratory tract infections, its prevalence is inversely related to age.[2]

Diagnosis

The diagnosis of AR is based on a history of typical symptoms and physical examination findings. Common symptoms include postnasal drainage, sneezing, itchy nose and eyes, and clear rhinorrhea. The frequency and effect of symptoms on sleep and productivity should be assessed to classify the AR as intermittent or persistent and mild or moderate–severe.

Patients suffering from AR can present with an "allergic shiner," a darkening of the infraorbital skin resulting from chronic venous pooling. In some children, wiping the front of the nose with the back of the hand in an upward motion (the allergic salute) creates a persistent horizontal crease across the nasal bridge that is a hallmark of chronic anterior rhinorrhea. Bilateral conjunctivitis may be present in patients along with ocular involvement.

On anterior rhinoscopy with a hand-held otoscope, engorged, boggy, and pale inferior turbinates also suggest AR. In addition, there is often a clear discharge coating the nasal cavity. Examination of the oropharynx often reveals cobblestoning of the mucosa, a sign of chronic postnasal drip.

Although not necessary to make the diagnosis of AR, two tests commonly used to demonstrate IgE-mediated allergic reactions are immediate-hypersensitivity skin testing and measurement of serum allergen–specific IgE levels. Skin prick tests correlate well with the symptoms of allergic rhinitis and with airway responsiveness to allergens, and measuring serum allergen-specific IgE levels provides an *in vitro* means of supporting a diagnosis of AR. Compared with skin prick testing, the *in vitro* test is more specific but less sensitive and can be more expensive.[3] However, serum tests may be the only practical way of detecting allergen-specific IgE in some patients, especially those with urticaria and eczema.

The diagnosis of CRS requires the presence of at least two major symptoms or one major and at least two minor clinical symptoms persisting for longer than 12 weeks, in conjunction with objective evidence of inflammation within the sinus cavity. The major symptoms include facial pain or pressure, nasal obstruction, nasal drainage, and hyposmia or anosmia. Minor symptoms are headaches, halitosis, fatigue, dental pain, cough, and ear pain or pressure. The most specific symptom for RS is discolored nasal discharge.[4] Since most symptoms of rhinosinusitis are nonspecific, evidence of inflammation on nasal endoscopy or imaging is also necessary to make a diagnosis of CRS. On nasal endoscopy, inflammation is suggested by edema and/or drainage from the middle meatus; CRSwNP is diagnosed when nasal polyps are visualized (Fig. 41.1). Inflammation within the sinuses may not be apparent on nasal endoscopy: in patients with a strong history of CRS, computed tomography (CT) of the sinuses is necessary to detect mucosal thickening and/or fluid within the sinuses (Fig. 41.2).

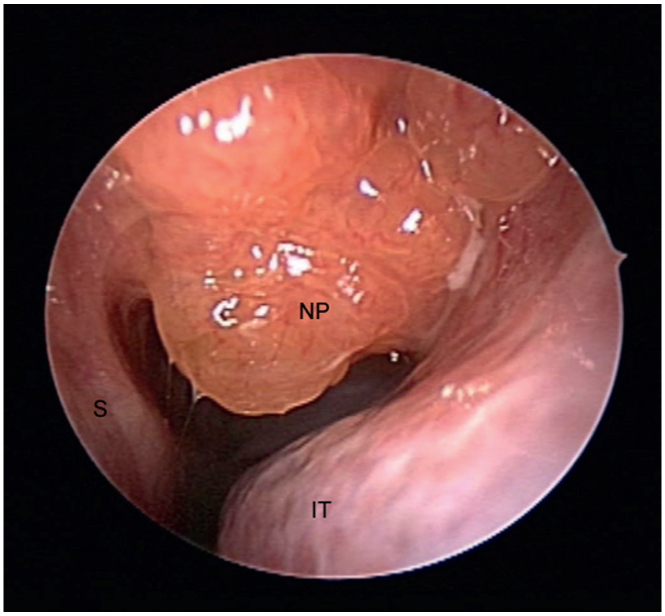


FIG 41.1 Nasal Polyposis. Seen on this nasal endoscopy are nasal polyps (NP) emanating from the sinus cavity into the nasal cavity between the septum (S) and the inferior turbinate (IT) of a patient with chronic rhinosinusitis.

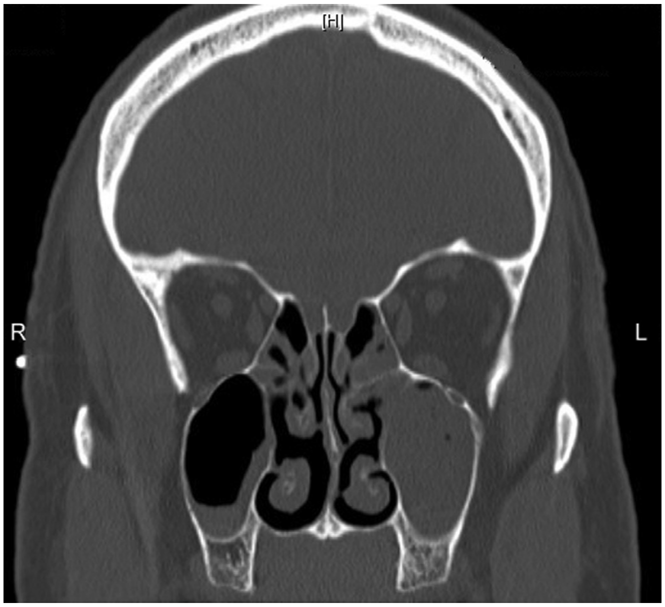


FIG 41.2 Coronal Sinus Computed Tomography (CT) Image From a Patient With Chronic Rhinosinusitis. The maxillary sinuses (lateral to the nasal cavity) and ethmoid sinuses (medial to the orbital cavities) exhibit mucosal thickening and accumulation of obstructed secretions consistent with inflammatory changes within the paranasal sinuses.

To make the diagnosis of AFRS, the most widely accepted diagnostic criteria include five characteristics: nasal polyps; type I (immediate) hypersensitivity to fungi; radiographic imaging consistent with AFRS; eosinophilic mucin with evidence of fungi; and lack of evidence of fungal invasion into the surrounding sinus tissue. Type I hypersensitivity to fungi can be detected by positive skin testing to fungal antigens or elevated serum fungal-antigen-specific IgE titers. Imaging consistent with AFRS is obtained by CT and magnetic resonance imaging (MRI). Fulfilling the remaining criteria for AFRS requires histological examination and microbiology for fungi within the mucin and tissue obtained at the time of surgery.[5]

KEY CONCEPTS

Diagnosis of Chronic Rhinosinusitis (CRS)

Chronic Rhinosinusitis

- Presence of at least two of the following symptoms:
  - Facial pressure or pain
  - Nasal obstruction or congestion
  - Anterior or posterior nasal drainage
  - Hyposmia or anosmia
- Edema or discolored drainage within the sinus cavity or middle meatus
- *Or* CT sinus showing fluid or mucosal thickening within sinus cavities

CRS Without Nasal Polyps	CRS With Nasal Polyps
No evidence of nasal polyps within the middle meatus as noted on nasal endoscopy in a patient with no history of previous sinus surgery	Presence of nasal polyps within the middle meatus as noted on nasal endoscopy
	History of nasal polyps within sinus cavity in a previously operated patient with CRS diagnosis

The diagnosis of AERD requires meeting diagnostic criteria for nasal polyps and asthma, along with a history of respiratory symptoms exacerbated by oral intake of aspirin or other cyclooxygenase-1 (COX-1) inhibitors on at least two occasions, or a positive reaction on aspirin challenge.[6]

Cystic fibrosis is typically diagnosed in early childhood as a result of pulmonary symptoms. However, the presence of nasal polyps in anyone under 18 years should prompt an evaluation for cystic fibrosis, which is confirmed either by sweat test or by demonstrating a mutation in the cystic fibrosis transmembrane conductance regulator gene.

Therapy

The most effective and widely used drugs for controlling the nasal and ocular symptoms of AR are glucocorticosteroids and antihistamines. Intranasal glucocorticosteroids (steroids) are the most efficacious medication class available for the management of AR symptoms, particularly nasal congestion, and are considered first-line medical therapy. They may also improve eye symptoms. Of the numerous formulations available, mometasone furoate demonstrates the highest antiinflammatory potency and lowest systemic absorption compared with other commonly used intranasal steroids. However, no studies have shown any differences in clinical efficacy among the currently available intranasal steroids.[7]

Oral antihistamines are also highly effective against the histamine-driven symptoms of sneezing, pruritus, and rhinorrhea. Second-generation antihistamines are recommended over first-generation formulations for their lack of sedating effects and equivalent ability to reduce AR symptoms. Intranasal antihistamines can be as effective as oral antihistamines for control of allergic and nonallergic rhinitis, but neither is as effective as intranasal steroids for nasal congestion. Other pharmacological agents potentially useful in the management of rhinitis include oral steroids, oral leukotriene (LT) receptor antagonists, intranasal chromones, and intranasal ipratropium bromide.[8]

In addition to pharmaceutical approaches, allergen immunotherapy offers a means of providing long-term relief from rhinitis (Chapter 91). Two types of immunotherapy are available: subcutaneous immunotherapy (SCIT) and sublingual immunotherapy (SLIT). An important concern associated with the use of SCIT is anaphylaxis, whereas this risk is probably much lower with SLIT. The antigens for SLIT are available in both liquid formulation, which allows for customized combinations, and preformed tablets. Finally, intensive home and workplace remediation directed at limiting allergen exposure is potentially beneficial in those with demonstrable environmental allergies.[8]

The treatment of CRS is more controversial than that of AR. In general, treatment can consist of either medical or surgical therapy. CRS is currently viewed as primarily an inflammatory, rather than an infectious, disorder such that antibiotics are less frequently advocated and have largely been replaced with intranasal and oral antiinflammatory agents, such as steroids. In addition, daily sinus irrigations with physiological saline can counteract mucociliary dysfunction and promote antigen clearance.

Future therapies for CRS are likely to include biologics aimed at the key cytokines, molecules, and cells that are important in the type 2 inflammatory responses which characterize CRSwNP. These potential therapeutic targets include IL-33, IL-25, thymic stromal lymphopoietin (TSLP), IL-5, IL-4, IL-13, siglec-8, and IgE. However, these treatment options will require a more sophisticated classification of CRS beyond the basis of nasal polyps. Endotyping based on immunological, molecular, or histological characteristics represents the future of CRS classification. We also need to better understand the role that environmental fungi play in triggering the inflammatory cascade leading to particular CRSwNP subtypes and allergic asthma. The use of antifungals in CRS seems logical, given the strong association between CRS and fungi, including yeasts and molds.[9] This has been investigated through many clinical trials, most of which are methodologically weak, failing to specify which patients actually have fungal infection or at least fungal sensitivity and lacking adequate control populations. Despite these limitations, the existing studies suggest that topical antifungal antibiotics (*e.g.*, amphotericin B) are not effective, whereas systemic antifungals may significantly inhibit recurrent disease.[10]

Patients with persistent symptoms despite medical therapy may benefit from functional endoscopic sinus surgery (see Video 1 online). Outcomes of medical therapy versus sinus surgery combined with intranasal steroids showed similar improvements, but patients with CRSwNP achieved improved asthma control with medical therapy. Endoscopic sinus surgery is mandatory for patients with AFRS in whom the sinuses are impacted with copious thick mucus requiring manual extraction and patients presenting with serious complications of CRS, such as vision loss or intracranial extension of disease.

Asthma

Epidemiology and Clinical Presentation

After several decades of rising incidence, asthma is now the most common chronic disease of childhood and one of the most common disorders of adults in the United States. Although most frequently diagnosed initially in childhood, asthma can be first diagnosed at any age. The widely prevalent and incurable nature of asthma will continue to assure its position as one of the most costly of medical afflictions both in terms of total medical expenditure and in time lost from work and school. Asthma is a lower respiratory tract disease that is characterized by dyspnea and other symptoms, including cough, chest tightness, chest pain, and wheezing. Persons with mild disease often only have a mild, chronic cough. In contrast to other obstructive lung diseases, asthma symptoms are present intermittently and are characteristically relieved by bronchodilator and antiinflammatory therapy.

Patients with asthma are classified into distinct clinical subtypes according to characteristic environmental or occupational exposures that elicit disease symptoms, the presence or absence of concomitant atopy, temporal expression of symptoms, and responsiveness to antiinflammatory therapy. Respiratory viruses are the most frequently implicated causes of asthma attacks, especially in children, and tobacco smoke and air pollution are other major inciting agents. A large minority of patients with asthma have *atopy,* a term that reflects the production of IgE, a genetic predisposition toward immediate-type immune reactions, and symptoms on exposure to causative environmental agents, such as pollens, dust mites, fungi, and insects. If atopy is present, patients are referred to as having extrinsic, atopic, or allergic asthma, whereas those without atopy are referred to as having intrinsic or nonallergic asthma. In general, airway constriction occurs and symptoms of asthma are provoked when triggering agents are inhaled from the environment, representing the clinical expression of *airway hyperresponsiveness* (AHR)—the exaggerated tendency of the asthmatic airway to constrict in response to exposure to a wide variety of provocative agents. Some of these agents, such as viruses and pollens, are only intermittently present, causing seasonal asthma, whereas other agents are encountered continuously (*e.g.,* fungi) and cause persistent (or perennial) asthma. *Occupational asthma* is defined as asthma acquired in the workplace, where dozens of potentially toxic agents have been identified (Chapter 49). Numerous additional clinical subsets of asthma can be defined according to the factor or factors that most often elicit attacks of dyspnea. A final category of asthma, *steroid-resistant asthma,* refers to the condition in patients who are relatively unresponsive to antiinflammatory steroid therapy.

Diagnosis

Asthma is often recognized on clinical grounds alone, with acute attacks marked by obvious dyspnea, wheezing, cough, and use of accessory muscles of respiration. Such attacks typically resolve with bronchodilator therapy. Spirometry can provide more objective evidence of airway obstruction as assessed by reversible decrements in the forced expiratory volume per second (FEV_1) and other measures of air flow. Nonetheless, a uniformly acceptable disease definition has remained elusive, at least partly because of the nonspecific nature of symptoms and a clinical spectrum that blends with many other disease processes. When the clinical presentation is uncertain, bronchial provocation tests can be used to determine the presence of AHR and thereby establish the diagnosis.[11] Additional laboratory data that support a diagnosis of allergic asthma include peripheral blood eosinophilia, elevated serum total and antigen-specific IgE levels, and positive skin prick test results against one or more allergens.

Therapy

As with AR and RS, the therapy of asthma is generally nonspecific and directed at improving air flow through bronchodilation and reducing inflammation. Immediate relief of bronchoconstriction and dyspnea is achieved with bronchodilating agents that activate the β_2 adrenergic receptor on airway smooth muscle (β-agonists). For long-term control of asthma, the most effective agent class is glucocorticosteroids, which reduce inflammation and suppress airway constriction and dyspnea. For mild to moderate disease, bronchodilating agents and steroids are typically administered by inhalation, which reduces but does not eliminate systemic side effects. A secondary class of agents used for controlling bronchospasm comprises anticholinergics, which are antagonists of the muscarinic acetylcholine receptor. Severe disease may also require treatment with steroids, which are given orally or intravenously for relatively brief periods to minimize the often-severe side effects, and high-dose inhaled β-agonists, often given by nebulizer.

Additional antiinflammatory agents available for the treatment of asthma include LT receptor antagonists, chromones, theophylline, omalizumab (a monoclonal antibody [mAb] that reduces circulating and mast cell-bound IgE), and most recently mepolizumab (an anti–IL-5 antibody).[12] Bronchial thermoplasty (BT) is a relatively new bronchoscopic technique, in which heat is applied to the airways via a radiofrequency catheter to ablate smooth muscle cells. Although early clinical trials have been encouraging with regard to the ability of this technique to reduce exacerbation rates, further studies are required to understand the patient subgroups that are most likely to respond to these novel treatments.[13]

OTHER AIRWAY ALLERGIC DISEASE SYNDROMES

In addition to the common allergic disorders discussed above, several other allergic airway diseases have been described: although not as common, these often cause profound morbidity. As with asthma, these disorders are clinically heterogeneous but are believed to share a similar pathophysiology related to the inhalation of antigens that provoke airway eosinophil and Th2 responses. No single clinical, pathological, or radiographic feature is pathognomonic for these diseases, and diagnosis relies on a constellation of findings, especially antigen exposures, radiographic details, and histopathology. Nonetheless, all of these disorders are prominently linked by the presence of eosinophils in peripheral blood and airway tissues.

Under the trophic influence of IL-5, eosinophils develop from precursor cells present largely in bone marrow. Early mature eosinophils are then released into blood, where they circulate for a brief period before entering the interstitium of airway tissues, where they can reside for long periods. Chemotactic factors that enhance the extravasation of eosinophils include complement components, eosinophil chemotactic factor-A, LTs, tumor-associated factors, and chemokines. Putative functions of tissue resident eosinophils range from host defense against parasites, such as nematode helminths, in diseases, such as tropical pulmonary eosinophilia, to mediators of end-stage tissue destruction and irreversible lung damage and fibrosis in other disorders, such as acute and chronic eosinophilic pneumonia.

The eosinophilic disorders discussed below are organized according to whether there is an *extrinsic* or *intrinsic* cause of the eosinophilia (Table 41.1). Inhaled or ingested extrinsic factors, including medications and infectious agents (*e.g.,* parasites, fungi, mycobacteria), can trigger an eosinophilic immune response. This may be mild and self-limiting, as in Loeffler syndrome. Intrinsic pulmonary eosinophilic syndromes are, by definition,

TABLE 41.1 Eosinophilic Lung Disorders

Disease	Causative Agent	Proposed Immune Mechanism
Loeffler syndrome	Inhaled food, infection, or medication	T cell–mediated hypersensitivity reaction
Drug rash with eosinophilia and systemic symptoms (DRESS) syndrome	Drugs: sulfonamides, phenobarbital, sulfasalazine, carbamazepine, and phenytoin	Hypersensitivity reaction to drug
Parasitic infections	*Strongyloides* spp., *Wuchereria bancrofti, Brugia malayi*	T-cell and B-cell clonal activation in response to parasite antigens and adjuvant factors
Allergic bronchopulmonary aspergillosis	*Aspergillus*	Immunoglobulin E (IgE) and immune complex deposition
Acute eosinophilic pneumonia	Fungal infections, cigarette smoking, post–stem cell transplantation	Hypersensitivity response to inhaled antigen (infectious or otherwise)
Chronic eosinophilic pneumonia	Unknown systemic-mediated process	Unknown, but chronic nature evident with T cell–mediated granuloma production
Idiopathic hypereosinophilic syndrome	Infections, systemic diseases, and drugs that drive peripheral eosinophilia	Systemic responses caused, in part, by excess interleukin-5 (IL-5) production from clonal expansion of Th2 cells as well as fusion gene *FIP1L1–PDGFR*
Churg-Strauss syndrome	Autoimmune vasculitis to unknown antigen, associated with asthma	Decreased regulatory T-cell function with diminished IL-10 production

idiopathic and represent a varied group of diseases, often systemic in nature.

Extrinsic Eosinophilic Syndromes

Tropical Eosinophilic Pneumonias

The tropical eosinophilic syndromes are a group of clinically similar eosinophil-predominant inflammatory disorders characterized by chest pain, wheezing, cough, and AHR, often in the setting of a debilitating, but transient, febrile illness. Fleeting lung infiltrates may be seen on chest radiographs, and laboratory studies often demonstrate strikingly high peripheral blood, lung, and airway eosinophilia, as well as elevated serum IgE levels. Migrating parasites traversing the lungs are thought to be responsible for most cases of tropical eosinophilic pneumonia. Embolization of microfilariae (*e.g., Dirofilaria* spp.) or helminth eggs within the pulmonary microvasculature leads to antigen release and induction of a typically granulomatous allergic immune reaction. Persistent or chronic, recurrent infection with etiological organisms (*e.g., Strongyloides* spp. *Wuchereria bancrofti, Brugia malayi*) leads to chronic inflammation that may cause parenchymal necrosis and irreversible fibrosis. In the United States, *Strongyloides* spp. are the most common cause of parasitical infection and tropical eosinophilic pneumonia. Patients who are immunocompromised, including those recently prescribed systemic steroids, may develop *Strongyloides* hyperinfection syndrome, in which large numbers of recently released larvae burrow through the intestine and migrate to the lungs, causing a severe and potentially fatal lung disease that is frequently complicated by sepsis (Fig. 41.3). A very similar disease, termed *Loeffler syndrome,* was originally reported to be caused by the larvae of *Ascaris* spp., but other helminths and hypersensitivity responses to medications have since been etiologically implicated. Therapy of parasite-related pulmonary eosinophilia syndromes is directed at relieving symptoms and eliminating the parasites.

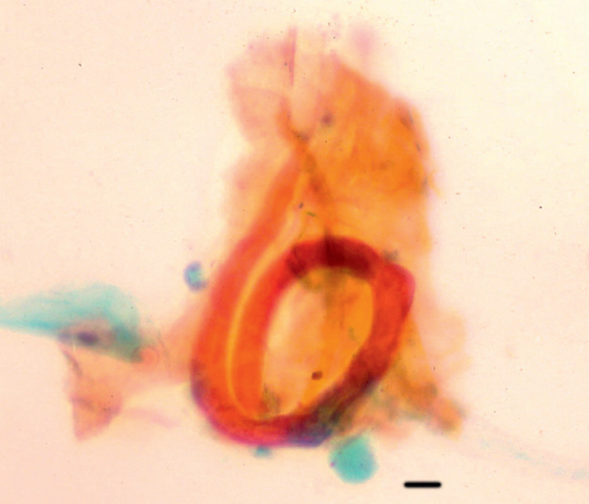


FIG 41.3 Strongyloidiasis. The coiled larva of *Strongyloides stercoralis* is seen on this Papanicolau stain of a bronchoalveolar lavage sample from a patient with *Strongyloides* hyperinfection. Original magnification ×400; bar = 10 µm.

DRESS Syndrome

The drug rash with eosinophilia and systemic symptoms (DRESS) syndrome is a severe drug hypersensitivity reaction that has a constellation of systemic signs and symptoms, including skin rash, fever, lymphadenopathy, and inflammation of the liver, lung, and heart (Chapter 48). Numerous drugs have been reported to cause DRESS syndrome, including sulfonamides, phenobarbital, sulfasalazine, and antiseizure medications, such as carbamazepine and phenytoin. Importantly, symptom onset may be delayed long after initiation of the drug.[14] The therapy of DRESS syndrome involves discontinuing the offending medication and providing supportive care in the setting of severe organ involvement.

Allergic Bronchopulmonary Aspergillosis

Allergic bronchopulmonary aspergillosis (ABPA) is a severe pulmonary allergic reaction to *Aspergillus* antigens that is seen almost exclusively in the setting of preexisting asthma or cystic fibrosis.[15] Diagnostic criteria include asthma with wheezing, peripheral blood eosinophilia, detection of precipitating anti-*Aspergillus* antibodies, elevated serum total IgE levels, and radiographic evidence of fleeting pulmonary infiltrates often accompanied by central bronchiectasis. *Aspergillus* spp. and other filamentous fungal species can frequently be isolated from airway secretions of patients with ABPA, suggesting that active fungal growth within the airways is responsible for the disease. Complications of chronic ABPA include severe AHR, severe bronchiectasis, eosinophilic pneumonia, pulmonary fibrosis, and invasive fungal disease.

Treatment of ABPA aims to suppress the inflammatory response to the fungus and to control bronchospasm with glucocorticoid therapy, the duration of which may be shortened by concomitant use of oral antifungal agents, such as itraconazole.[15]

Acute Eosinophilic Pneumonia

Acute eosinophilic pneumonia (AEP) is an acute, often debilitating eosinophilic inflammatory syndrome that exclusively involves the lungs and is marked by pulmonary infiltrates, dyspnea progressing to frank respiratory failure, and fever. The diagnosis is dependent on eosinophils exceeding 25% of all inflammatory cells within the bronchoalveolar lavage (BAL) fluid. Increasing evidence suggests an association between AEP and respiratory fungal infections and recent-onset cigarette smoking.[16,17] AEP has also been reported following allogeneic hematopoietic stem cell transplantation (HSCT) in the setting of graft-versus-host disease (GvHD).[18] Prompt recognition of the disease and initiation of treatment with glucocorticoids usually results in rapid radiographic and clinical improvement.

Intrinsic Eosinophilic Syndromes

Chronic Eosinophilic Pneumonia

This disorder presents similarly to AEP, but in a more chronic manner (greater than 6 weeks duration). Chronic eosinophilic pneumonia (CEP) can occur in isolation and/or in association with polyarteritis nodosa, rheumatoid arthritis, scleroderma, ulcerative colitis, breast carcinoma, and histiocytic lymphoma. Most patients have evidence of asthma and atopy. Like AEP, CEP can present with striking eosinophilic inflammation of the lung (Fig. 41.4). Granulomas are occasionally seen on biopsy specimens, suggesting that an antigen-driven, T cell–mediated process is involved in the chronicity of the disease. Treatment, as for AEP, is based on steroid therapy, but unlike AEP, CEP frequently relapses after discontinuation of treatment.[19]

Idiopathic Hypereosinophilic Syndrome

This multisystem disease is characterized by the massive accumulation of eosinophils in many tissues and almost always involves the lungs. Expansion of Th2 cells and local and systemic release of IL-4 and IL-5 are also frequently seen. The clonal expansion of Th2 cells in the absence of known antigen exposures, and the association of idiopathic hypereosinophilic syndrome (IHES) with a variety of chromosomal aberrations, strongly support the concept that IHES is a myeloproliferative disorder involving Th2 cells, although aberrant secretion of IL-5 by both solid and liquid tumors can produce very similar syndromes. Many organs can be affected, resulting in dysfunction or failure of the gastrointestinal (GI) tract, skeletal muscles (leading potentially to respiratory failure), endomyocardial fibrosis, myocarditis, and congestive heart failure. Pulmonary involvement manifests as obstructive airway disease, pulmonary edema, or pulmonary emboli caused by a hypercoagulable state. In addition to hypereosinophilia, patients with IHES may have evidence of polyclonal hypergammaglobulinemia. The diagnosis is based on the discovery of elevated peripheral blood eosinophilia in the setting of a multisystem disorder, with evidence of aberrant Th2 responses or elevated IL-5 secretion and defined genetic mutations. The most consistently effective therapy for IHES involves tyrosine kinase inhibition, using agents such as imatinib mesylate.[20]

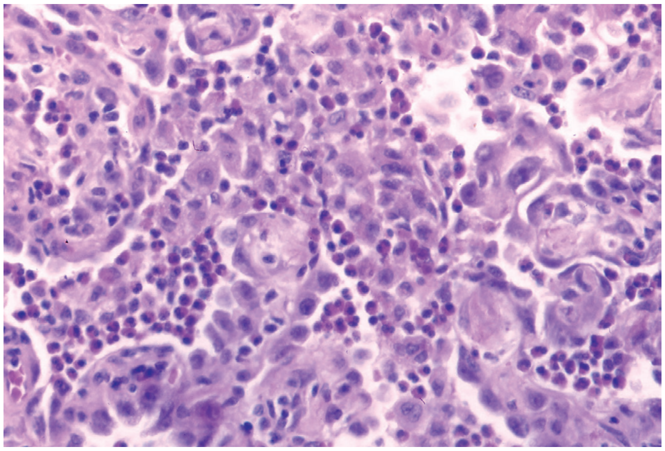


FIG 41.4 **Histology of Chronic Eosinophilic Pneumonia.** Lung biopsy specimen from a patient with chronic eosinophilic pneumonia demonstrates a confluent infiltrate with eosinophils filling alveoli together with large, multinucleate macrophages. Original magnification, ×200, hematoxylin and eosin stain.

Churg-Strauss Syndrome

Churg-Strauss syndrome (CSS), also termed *eosinophilic granulomatosis with polyangiitis* (EGPA), is a necrotizing vasculitis of medium and small caliber vessels (Chapter 58), which is characterized by airway obstruction and eosinophilia. The disease has an autoimmune nature, given the presence of circulating anti-myeloperoxidase and antineutrophil cytoplasmic antibodies (p-ANCA) as seen in 60–70% of affected individuals. Because CSS is seen in patients with a history of asthma and allergies and the prominent pathological feature is necrotizing vascular and tissue granulomas, the term “allergic granulomatosis and angiitis” is used synonymously. Inhaled or ingested antigens have been proposed as causative agents in susceptible individuals. Reports linking the syndrome with the LT inhibitors zafirlukast and montelukast in the setting of steroid withdrawal suggest that these agents may unmask preexisting CSS rather than directly causing the disorder. Similar observations have been made with omalizumab treatment. The vasculitis of CSS can affect the sinuses, the central and peripheral nervous systems, the GI tract, kidneys, and the heart. Treatment of CSS is based on the use of systemic steroids, which leads to disease resolution in the majority of patients. In severe steroid-resistant disease, cyclophosphamide and other immunosuppressants may be required.[21]

IMMUNOLOGICAL MECHANISMS OF ALLERGIC AIRWAY DISEASE

All patients with major allergic airway disease syndromes include patients who demonstrate evidence of allergic inflammation, with elevated blood and airway eosinophilia, elevated serum total and antigen-specific IgE levels, and positive skin prick tests. However, not all patients demonstrate all markers of allergic inflammation simultaneously, and some patients with apparent allergic disease express none of these markers. Specialized testing may also reveal predominant Th2 cytokine responses generated from peripheral blood mononuclear cells when stimulated with common allergens *in vitro* and the presence of Th2 cells, ILC2, eosinophils, and IgE-secreting B cells, and the secreted products of these cells, within affected sinus and airway tissues. As well as the conventional causes of Th2-mediated airway inflammatory

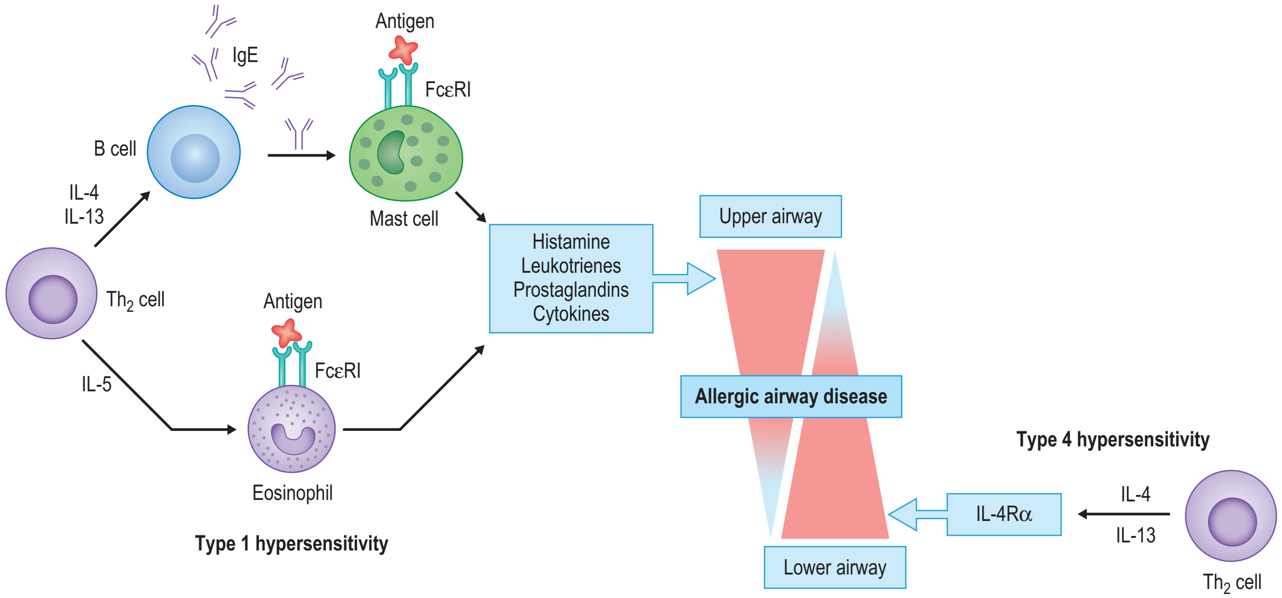


FIG 41.5 **Differential Importance of Allergic Immune Mechanisms According to Airway Level.** Type I hypersensitivity (left), mediated by immunoglobulin E (IgE)-primed mast cells and eosinophils, is ultimately driven indirectly by the cytokines secreted by T-helper 2 (Th2) cells. In contrast, type IV hypersensitivity (right) is mediated directly by Th2 cytokines, especially interleukin-4 (IL-4) and IL-13, acting through a similar receptor that includes IL-4Rα. Both immune mechanisms are important to the expression of allergic disease at all airway levels, but type I hypersensitivity predominates in the upper airway, whereas type IV hypersensitivity likely assumes a more important role in the lower airway.

responses, fungal and bacterial infections also serve as distinct causes of inflammation.

The major effector immune mechanisms of disease proposed by Gell and Coombs in the 1960s remain essential to understanding the pathogenesis of allergic airway diseases.[22] Although such mechanisms probably operate to some degree in all allergic diseases, their relative importance appears to vary, depending on whether the disease process predominantly affects the upper respiratory tract or the lower respiratory tract (Fig. 41.5).

KEY CONCEPTS

Immunopathogenesis of Allergic Airway Disease

- Gell and Coombs type 1 and type 4 hypersensitivity mechanisms contribute to disease expression, especially airway obstruction.
- Innate immune pathways involving thymic stromal lymphopoietin (TSLP), interleukin-25 (IL-25), and complement proteins critically contribute to the development of allergic airway inflammation.
- Environmental agents that are emerging as potentially important initiation factors for allergic airway inflammation include proteases, chitin, and endotoxin as derived from fungi, insects, and bacteria.
- Fungi and viruses are further emerging as potentially important infectious causes of allergic disease of the upper and lower airways.

Type I (Immediate) Hypersensitivity

This form of hypersensitivity involves the activation of basophils and mast cells that release histamine and other inflammatory mediators that drive the clinical features of the response. Antigen recognition is via IgE antibodies, which bind through their Fc portion to high affinity receptors (FcεRI) to arm effector cells. Th2 cells coordinate both the production of IgE antibodies and the activation and recruitment of allergic effector cells to the airway. Antigen-specific IgE bound to the surface of mast cells and basophils is cross-linked upon exposure to relevant antigens, causing cellular activation and release of preformed and newly formed mediators of inflammation, such as histamine, proteases, LTs, numerous cytokines, and other substances. IL-4, released primarily by Th2 cells, is an especially important regulator of type I hypersensitivity reactions, as it is required for B-cell maturation and IgE secretion.

There is evidence that IL-4 and IL-13 can mediate distinct effector phenotypes in airway and tissue macrophages and dendritic cells (DCs). Two major effector macrophage subtypes include conventionally activated (M1) macrophages that arise under the predominant influence of type I cytokines, especially IFN-γ, and alternatively activated macrophages (M2) that arise under the influence of IL-4 and IL-13 in the relative absence of IFN-γ. M2 macrophages express a distinct gene profile that includes high-level expression of arginase 1, Ym1, Fizz1 (RELM), and PD-L2. Current evidence suggests that M2 macrophages promote allergic responses and that redirecting them to the M1 phenotype might be therapeutically useful.[23]

The effect of mediator release in immediate hypersensitivity reactions is most dramatically illustrated by the clinical syndrome of anaphylaxis, in which patients may exhibit profound bronchoconstriction and dyspnea accompanied by the extravasation of vascular solutes and fluid into the interstitium, which produces diffuse edema, bronchoconstriction, and hypotension (Chapter 42). Type I hypersensitivity against inhaled allergens more commonly presents with symptoms of acute rhinitis.

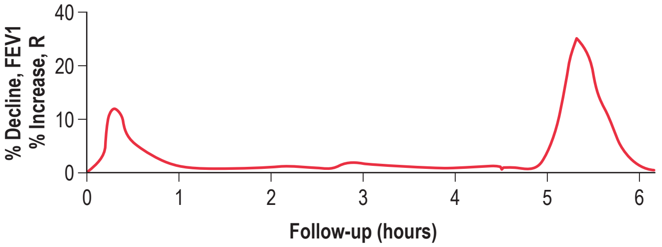


FIG 41.6 **Early-Phase and Late-Phase Airway Changes Following Allergen Challenge.** This schematic diagram depicts the two phases of bronchoconstriction typically seen after allergen inhalation in sensitized, asthmatic subjects. Within 20–30 minutes after allergen inhalation, the first (early) phase of bronchoconstriction as assessed by either a decline in forced expiratory volume per second (FEV_1) or increase in airway resistance (R) is seen. After quickly subsiding, approximately 4–6 hours (h) later, a second (late) phase of bronchoconstriction occurs.

Cell-Mediated Features of Immediate Hypersensitivity

Airway obstruction in allergen-sensitized asthma evolves over several hours after allergen exposure and is seen in two distinct phases. The early-phase response is marked by airway constriction that becomes maximal about 30 minutes after allergen exposure and is fully resolved after approximately 2 hours (Fig. 41.6). Approximately 50% of subjects with asthma who are tested will also develop a late-phase airway response in which airway obstruction again develops 4–6 hours after allergen exposure. Late-phase reactions are linked to the infiltration of the airways with Th2 cells and eosinophils, and the accompanying bronchoconstriction is less reversible with bronchodilating agents than the early-phase reactions.[24]

AHR is neurologically mediated through parasympathetic nerves, such as the vagus, and is fully reversible with bronchodilating agents that either interrupt muscarinic parasympathetic signaling directly (*e.g.*, ipratropium bromide) or activate receptors (*e.g.*, β_2-adrenergic) that antagonize muscarinic bronchoconstrictive pathways. Besides its formal demonstration in the clinical laboratory through bronchial provocation testing, AHR is recognized clinically as episodic bronchoconstriction that is reversed with bronchodilating agents. Late-phase responses after antigen challenge, therefore, represent a form of AHR and, in part, reveal the immunological nature of this ultimately neurological phenomenon.

More detailed analysis of the immunological mechanisms underlying AHR comes from experimental models of asthma. Studies from many species have demonstrated that AHR in the setting of allergic inflammation is critically dependent on Th2 cells and ILC2 that have specifically been recruited to the lung. Moreover, it is now clear that IL-13 is the major Th2 and ILC2 cytokine that mediates AHR by acting directly on constitutive airway cells, such as airway smooth muscle cells that express the IL-13 receptor. However, IL-13 does not directly induce bronchoconstriction. Bronchoconstriction in subjects with asthma, expressed as episodic bouts of dyspnea, is triggered by diverse exogenous factors in addition to allergens (*e.g.*, altered temperature and humidity, pungent odors, irritating aerosols) and endogenous stimuli (*e.g.*, extreme emotional states) with little apparent connection to immunological mechanisms. Thus rather than directly mediating airway obstruction, IL-13 establishes the basis for responding broadly to diverse agents with neurologically mediated bronchoconstriction. How IL-13 mediates these neurological changes is currently unknown.[11,25]

TABLE 41.2 **Chemoattractants Linked to the Recruitment of Allergic Inflammatory Cells**

Chemokine	Receptor
CCL1	CCR8
CCL11	CCR3
CCL17, CCL22	CCR4
CX_3CL1	CX_3CR1
Prostaglandin D2	CR T_H2
Leukotriene B4	BLT1

A second and more insidious form of airway obstruction is physical obstruction of the airways caused by mucus and fibrin clots that can accumulate in the airways as tenacious plugs, a phenomenon that is now termed plastic bronchitis.[26] IL-13, together with IL-9, is a trophic and differentiation factor for airway goblet cells and submucosal glands, which contribute to plastic bronchitis by promoting secretion of mucus by these cells. Airway obstruction caused by plastic bronchitis is not immediately reversible with bronchodilators or other pharmacological agents and is consequently the major cause of death, resulting from asphyxiation, in asthma.[27]

Finally, IL-4 and IL-13 further coordinate the recruitment and retention of allergic effector cells to airway epithelium and submucosa, an arrangement that facilitates rapid responses to inhaled allergens. Acting through a similar receptor, that includes the α chain of the IL-4 receptor (IL-4Rα), IL-4, and IL-13 signal on constitutive airway cells, such as airway epithelial cells, to induce secretion of a restricted repertoire of chemoattractants or chemotactic factors that promote the immigration of allergic cells expressing specific cognate receptors from the lung and airway microcirculation (Table 41.2).[28]

Contributing Immune Mechanisms in Allergic Airway Disease

In addition to its critical role in promoting IgE secretion, IL-4 is an important growth and differentiation factor for Th2 cells. Both IL-13 and IL-4 activate the transcription factor signal transducer and activator of transcription 6 (STAT6). In addition to these closely related cytokines, all components of the IL-4/IL-13 signaling pathway are required for expression of experimental allergic airway disease, especially AHR.

IL-5 also contributes to both immediate and cell-mediated hypersensitivity reactions through its role in promoting the growth and differentiation of eosinophils. Although widely viewed as pathogenic and contributing to the expression of allergic airway diseases, more recent studies indicate that eosinophils are important in tissue remodeling and in controlling allergic inflammation induced by pathogens, such as fungi.[29]

In addition to Th2 cells, mast cells, and eosinophils, innate lymphoid cells that include ILC2, natural killer (NK) cells (a type of ILC1), and $\gamma\delta$ T cells may also contribute to allergic disease expression through their ability to rapidly secrete Th2 and other cytokines.[25]

Numerous additional soluble mediators contribute to the expression of allergic disease. The complement system is especially important, and complement proteins C3a and C5a, the major proteolytically derived anaphylatoxins, are both essential for the expression and regulation of asthma in experimental systems. C3a signaling through the C3a receptor C3aR is required for robust Th2 responses, allergic inflammation, and AHR in response to airway allergen challenge. In contrast, C5a, which can signal through two receptors, C5aR and C5L2, appears to inhibit Th2 responses, perhaps acting as a physiological antagonist of the allergic disease-promoting activity of C3a.[30]

Lipid mediators of inflammation of importance to allergic airway disease include the LTs and prostaglandins (PGs). The cysteinyl leukotrienes (CysLTC4, -D4, and -E4) signal through at least two major receptors to mediate some of the same allergic disease features as IL-13, including airway inflammation and AHR. Full expression of experimental allergic disease, in fact, appears to require the concomitant expression of both IL-13 and the CysLTs. However, IL-13 appears to be the dominant allergic mediator, perhaps accounting, in part, for why LT antagonism alone is not as effective as inhaled steroids in asthma. Noncysteinyl LTs, such as LTB4, also contribute to the expression of allergic airway inflammation by controlling the recruitment of allergic effector cells, including Th2 cells.[31]

Similarly, PGD_2 is an important mediator of Th2 cell recruitment and allergic airway inflammation in rodent models, most likely acting cooperatively with LTB4 and chemokines. Most other PGs are thought to promote allergic airway disease, especially in aspirin-sensitive rhinitis and asthma. An important exception to this is PGE_2, which exhibits potent antiallergic activity. Other lipid mediators likely contributing to the expression of allergic inflammation and airway obstruction are the thromboxanes and lipoxins. Thromboxane A_2 is a potent proinflammatory lipid derived from platelets, whereas lipoxins have the opposite effect and are antiinflammatory.[32]

Environmental Factors and Allergic Disease Initiation

The earliest immunological events that initiate Th2 responses and the environmental factors that trigger them are incompletely understood. Currently, asthma and to a large extent RS and AR are believed to represent aberrant immunological responses to innocuous inhaled antigens. The very strong association between asthma and atopy supports this concept, with many allergens being relatively innocuous proteins derived from organisms that are not otherwise harmful or infectious, such as those from dust mites, cats, dogs, and plants. Atopy is further thought to represent a fundamental underlying condition that leads, in some cases, to overt allergic disease, but whether atopy represents primarily a genetically or an environmentally controlled condition remains unclear. Extensive analysis of allergens has revealed no consistent structural features, suggesting that physical properties are highly relevant to atopy and allergic disease expression.

A major exception to the intrinsically innocuous nature of allergens is, however, the role of filamentous fungi, such as *Aspergillus* spp., and other potentially infectious causes of allergic disease. Fungal sensitization is not different from sensitization to any other allergen ubiquitously present in human environments. However, fungi differ from all other allergen sources in that they are both ubiquitous and able to actively infect and grow within the respiratory tract.[33]

A fungal infectious basis for allergic disease is also suggested by *(i)* the high rate of isolation of filamentous fungi from the airways in CRS, especially AFRS, and ABPA; *(ii)* the universal presence of fungus-specific immunity in subjects with AFRS and ABPA; *(iii)* the efficacy of antifungal antibiotics when given to fungus-sensitized patients with asthma; and *(iv)* experimental validation that filamentous fungi are infectious for the mouse airway and readily produce allergic airway disease that is comparable with asthma. Moreover, fungal airway infection can induce atopy to innocuous bystander antigens, suggesting that fungal infection could underlie both atopy and respiratory tract allergic disease. However, not all patients with allergic airway disease demonstrate fungus-specific immunity; neither are antifungal agents effective in all such patients, especially those with CRS. Thus fungal airway infections may be etiologically relevant to only a subset of allergic disease patients.[33]

Respiratory viruses and particulate matter are also prominently linked to allergic airway disease. Approximately 70–80% of children and adults test positive for human rhinovirus (HRV) during acute disease exacerbations. Other respiratory viruses are likely to contribute to allergic disease pathogenesis, although the mechanisms remain obscure.[34] Particulate matter in the form of tobacco smoke, diesel exhaust particles, and other forms of smoke is strongly linked to asthma exacerbation and enhanced atopy, as is exposure to ozone (O_3). What links these various forms of air pollution to allergic disease may be the induction of oxidative stress, which ultimately leads to enhanced activation of nuclear factor kappa B (NF-κB), enhanced Th2 cytokine release, and increased allergic inflammation.[35]

Research from experimental systems has shed additional light on how allergens initiate allergic inflammation and disease. Although the structural features of allergens do not determine their allergic character, a common biochemical feature that is strongly associated with allergenicity is protease activity. Proteases, as single molecules, are as effective as any complex allergen or fungal infection in inducing allergic airway inflammation and AHR when administered to rodents or inhaled by humans. Household proteases are derived largely from fungi, suggesting again that airway infection caused by these organisms, which would result in *in situ* protease production, may be an important mechanism underlying allergic disease induction. Irritation of the airways through viral infection, ozone exposure, and other mechanisms further increases endogenous airway protease activity, especially through induction of thrombin activity.[36]

Analyses of diverse allergenic proteases suggest that they initiate a complex, airway epithelial-centered mechanism in which the epithelial cytokines TSLP, IL-33, and IL-25 are induced and lead to robust allergic responses. In part, this sequence is initiated by the action of exogenous and endogenous proteinases on the terminal coagulation protein fibrinogen, which is secreted by airway epithelial cells into the airway lumen. Cleavage of fibrinogen by proteases yields fibrinogen-cleaved products (FCPs) that signal through Toll-like receptor 4 (TLR4) to initiate both antifungal responses and innate allergic inflammation, including ILC2 responses.[37] Proteolytic activation of complement proteins (*e.g.,* C3) is also a crucial early event in the evolution of allergic airway disease.[30] Th2 responses arise, in part, through the action of proteases on pulmonary dendritic cells (Fig. 41.7).

NONALLERGIC RESPIRATORY TRACT INFLAMMATORY SYNDROMES

Although the majority of respiratory tract inflammatory disease is allergic, additional pulmonary immune-related disorders exist that are nonallergic and are responsible for considerable morbidity. The immunological mechanisms, etiological environmental factors, and therapies for these diseases are distinct from common

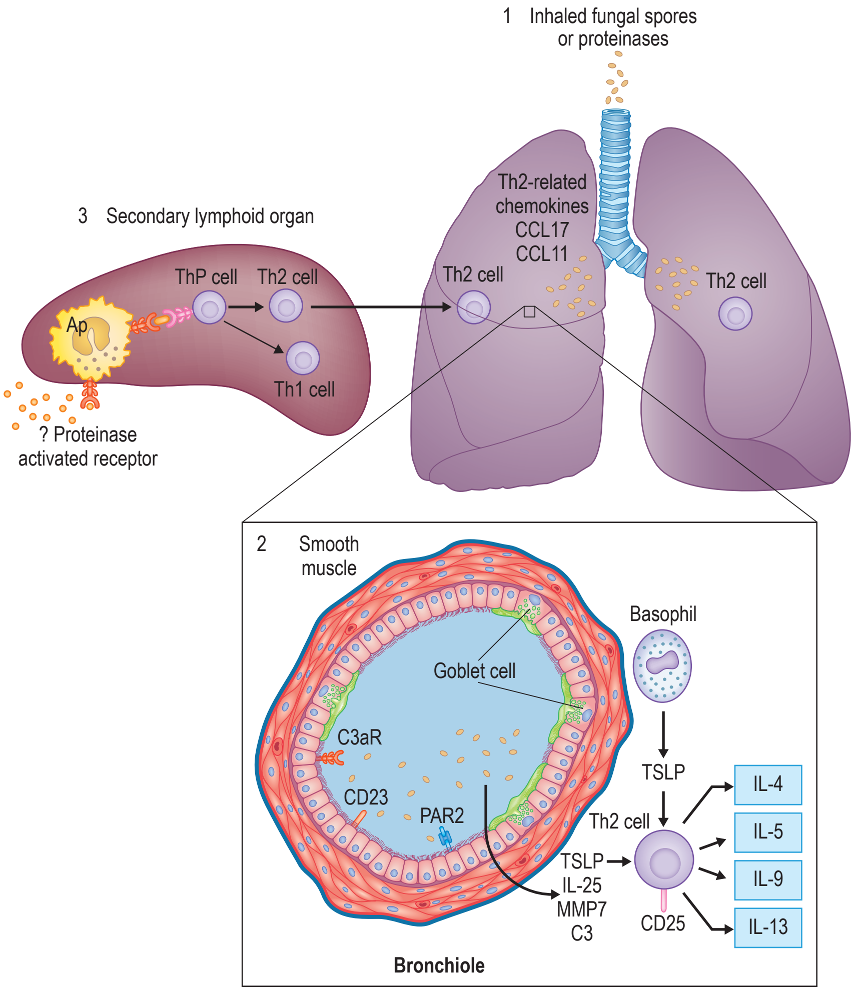

1 Inhaled fungal spores or proteinases
3 Secondary lymphoid organ
Th2-related chemokines
CCL17
CCL11
Th2 cell
Th2 cell
ThP cell
Th2 cell
Ap
Th1 cell
? Proteinase activated receptor
2 Smooth muscle
Basophil
Goblet cell
C3aR
CD23
PAR2
TSLP
Th2 cell
TSLP
IL-25
MMP7
C3
CD25
IL-4
IL-5
IL-9
IL-13
Bronchiole

FIG 41.7 Mechanisms of Allergenic Protease-Dependent Induction of Allergic Airway Disease. Inhaled proteases or protease sources (*e.g.,* fungal spores) initiate a series of molecular events in discrete lung compartments and involving distinct cell types that induce predominant airway T-helper 2 (Th2) responses that coordinate both the allergic inflammation and physiological changes that typify allergic respiratory tract disease. Initial innate immune responses induced by proteases include induction of airway chemokines that favor recruitment of allergic effector cells including Th2 cells (1). Likely airway cellular targets of proteinases include basophils, epithelial cells, smooth muscle cells, and macrophages (2). Activation of these cells by cleavage of cell surface receptors, such as PAR2 and CD23, potentially leads to activation of these cells to produce proallergic cytokines, such as thymic stromal lymphopoietin (TSLP), interleukin-33 (IL-33), and IL-25 that promote Th2 responses. Allergenic proteinases also likely act on soluble substrates, such as complement, especially C3 to generate C3a, the ligand for the C3aR, and fibrinogen, creating fibrinogen cleaved products (FCPs) that signal through Toll-like receptor 4 (TLR4) to coordinate antifungal immunity and promote Th2 responses. CD25 is another immune receptor present on T cells that can be cleaved by proteinases, potentially to favor Th2 cytokine secretion. Finally, proteinases act directly on antigen-presenting cells, such as dendritic cells (DCs), through an unknown mechanism in secondary lymphoid organs, such as lymph nodes, to promote their maturation in a manner that favors Th2 cell differentiation from naïve precursor (ThP) T cells (3). (Modified from Porter PC, Yang T, Luong A, et al. Proteinases as molecular adjuvants in allergic airway disease. Biochim Biophys Acta 2011;1810:1059–65.)

allergic disorders and hence warrant consideration in the differential diagnosis of airway inflammatory disorders. Two of the most important nonallergic airway obstructive immune disorders are hypersensitivity pneumonitis and chronic obstructive pulmonary disease.

Hypersensitivity Pneumonitis

Hypersensitivity pneumonitis (HP) is a non–IgE-mediated inflammatory lung disease that results from recurrent exposure to a wide variety of inhaled antigenic aerosols containing organic and possibly infectious matter. Pulmonary involvement is typically diffuse and consists predominantly of mononuclear inflammation of the terminal bronchioles, interstitium, and alveoli, with little involvement of the larger airways. Although the term “extrinsic allergic alveolitis” is commonly used for this disorder, low-grade eosinophilia is only seen in 1–3% of BAL specimens and is not a consistent or characteristic feature of disease. Over time, this pattern of inflammation leads to destruction of alveoli and to irreversible pulmonary fibrosis that may ultimately be fatal.

The etiology of HP is usually idiopathic, but many cases can be traced to exposure to organic aerosols that contain thermophilic bacteria (*e.g., Saccharopolyspora rectivirgula*), filamentous fungi (*e.g., Aspergillus* spp.), animal proteins and fecal matter (*e.g.,* pigeon breeder’s disease), and industrial chemicals, such as isocyanates. Symptoms, including chest tightness, chest pain, dyspnea, and fever, appear 4–6 hours after exposure. Cessation of exposure to the provoking antigen may prevent progression to chronic, irreversible disease. Other than oxygen therapy in the setting of profound hypoxemia, there is no defined medical therapy for HP; in particular, there is no role for glucocorticosteroids.

The immunopathogenesis of HP is complex and not well understood. During the acute presentation, the histological picture is an immune complex–mediated interstitial injury with a predominant neutrophilic infiltrate. Subsequently, the disease evolves into predominant mononuclear inflammatory infiltrates consisting of lymphocytes, plasma cells, and foamy macrophages, followed by granuloma formation. In advanced disease, the inflammatory infiltrates are replaced by fibrosis.

Acutely, HP is thought to be initiated by an immune-complex mediated hypersensitivity (type III) with *in situ* immune complex deposition in the lung interstitium as a result of interaction of the inhaled antigen and preexisting IgG antibodies in the alveolar spaces. Complement activation occurs and most likely contributes to the alveolitis and neutrophilia. However, many people with precipitating antibodies to putative HP antigens (precipitins) do not develop actual parenchymal or symptomatic disease. Therefore it is believed type IV, delayed-type hypersensitivity reactions involving Th1 and possibly Th17 cells also contribute to disease expression.[38]

Chronic Obstructive Pulmonary Disease

The disorder most frequently confused with asthma is chronic obstructive pulmonary disease (COPD), which encompasses chronic bronchitis with or without emphysema. COPD is currently the fourth leading cause of death in the world but is expected to reach third place by 2020. The most common cause of COPD is active tobacco smoking, but chronic exposure to coal dust, biomass fuels, second-hand smoke, and chronic respiratory infections have been linked to COPD in nonsmokers. Because not all smokers develop COPD, disease initiation is thought to reflect interaction of genetic susceptibility and environmental factors. Development of early-onset emphysema in highly susceptible smokers is sometimes caused by undetected mutations in the α_1 antitrypsin (*A1AT*) gene locus (*e.g.,* ZZ, MZ, MS). However, in the vast majority of individuals, the genetic factors responsible for increased susceptibility to emphysema remain unknown. Indeed, several large studies have shown that although multiple genes contribute to increased risk of emphysema, each individual gene shows only a modest and independent effect.[39]

In addition to cough and sputum production, COPD often presents clinically with insidious onset of dyspnea and progressive and persistent exercise limitation, with minimal reversibility of airway obstruction. Epidemiological studies have shown that lung cancer risk is strongly related to radiographic emphysema, independent of air flow obstruction. In contrast, nonsmoking patients with asthma have no greater risk of lung cancer than the general nonsmoking population.

The Global Initiative for Chronic Obstructive Lung Disease (GOLD) has in the past decade promoted the use of a fixed FEV_1/forced vital capacity (FVC) ratio of <70% to diagnose COPD. This latter guideline is now widely used to classify smokers with and without COPD, but it excludes up to 20% of smokers with emphysema who have normal lung function, and it underdiagnoses and overdiagnoses COPD in young and old smokers, respectively. Recognizing different COPD phenotypes based on lung function, symptom severity, and exacerbation frequency, smokers are categorized into one of four stages: (A) fewer symptoms, low risk; (B) more symptoms, low risk; (C) fewer symptoms, high risk; (D) more symptoms, high risk. Those in groups C and D are thought to benefit from inhaled corticosteroids in addition to bronchodilators.[40]

Smokers with emphysema have enlarged alveoli that contain increased numbers of macrophages containing nano-sized elemental carbon black (nCB). The key property underlying the initiation of destructive lung inflammation with nCB inhalation appears to be its chemical characteristics, namely, its insolubility and small particle size that activate the inflammasome pathway and promote differentiation of lung Th1 and Th17 cells. nCB can be readily detected in lung phagocytic cells (*e.g.,* DCs and macrophages) by using Raman spectroscopy and electron microscopy (Fig. 41.8).[41]

The treatment of emphysema is largely supportive, comprising bronchodilators and cholinergic antagonists, which suppress mucus production, and inhaled or systemic glucocorticosteroids,

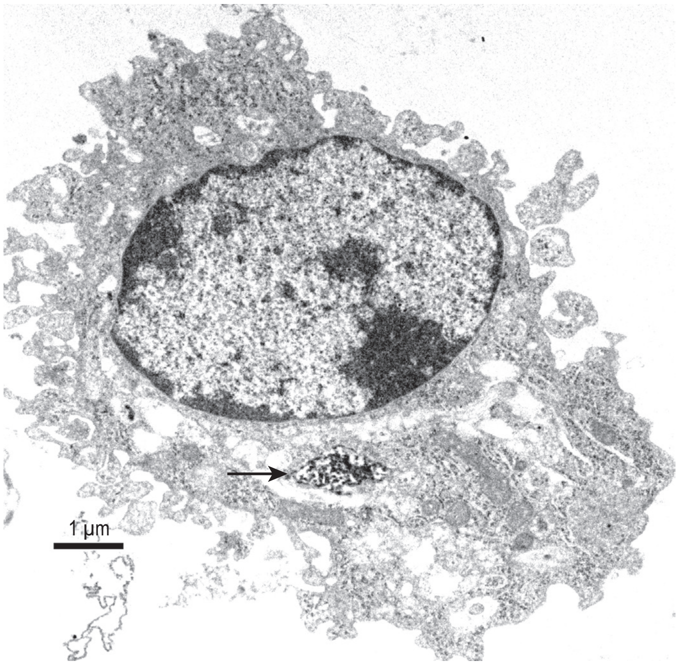


FIG 41.8 Nanoparticulate Carbon Black (nCB) Accumulation Within an Alveolar Macrophage. Electron micrograph of an alveolar macrophage isolated from the peripheral lung of a cigarette smoker with emphysema. Arrow points to a vacuole containing nCB.

https://archive.org/details/namdhari

which suppress inflammation. Additional therapies for moderate to severe disease include theophylline and the phosphodiesterase inhibitor roflumilast. During exacerbations of COPD, which become increasingly frequent with advancing disease and are often caused by bacterial and viral infections, tailored use of antibiotics is frequently helpful. For severe disease with resting hypoxemia, continuous nasal oxygen can both extend life and relieve dyspnea. Surgical interventions, such as lung volume reduction surgery and lung transplantation, are reserved for advanced disease.

The immunological basis of COPD differs markedly from asthma: smoking-related lung inflammation appears to be the fundamental underlying cause of both airway obstruction and lung destruction (Fig. 41.9). In contrast to asthma, where Th2 cells and other allergic effector cells are involved, the major immune effector cells in COPD are Th1 and Th17 cells. Together with macrophages and neutrophils, these innate and adaptive immune cells coordinate lung destruction in COPD by enhancing production of matrix metalloproteinases (MMPs) and other enzymes that degrade elastin (elastases), including MMP2, MMP9, MMP12, and neutrophil elastase (NE). The activity of most of these enzymes is controlled by endogenous inhibitors, such as tissue inhibitors of metalloproteinases (TIMPs) and A1AT, the principal antagonist of elastases. Cigarette smoke exposure is sufficient to enhance elastase secretion from macrophages and neutrophils and to reduce the activity of elastase inhibitors. In addition to elastin, MMP12 specifically inhibits A1AT, leading to unopposed NE activity. These biochemical events are critically important to the expression of emphysema because elastin is an important MMP that is partly responsible for both the structural integrity of the lung and its stretchiness. In genetically susceptible individuals, elastin becomes the antigenic target of pathological Th1 and Th17 cells. Emphysema, therefore, is the clinical expression of the autoimmune-based loss of lung integrity resulting from enhanced expression of antielastin immunity that promotes secretion of elastases.[42]

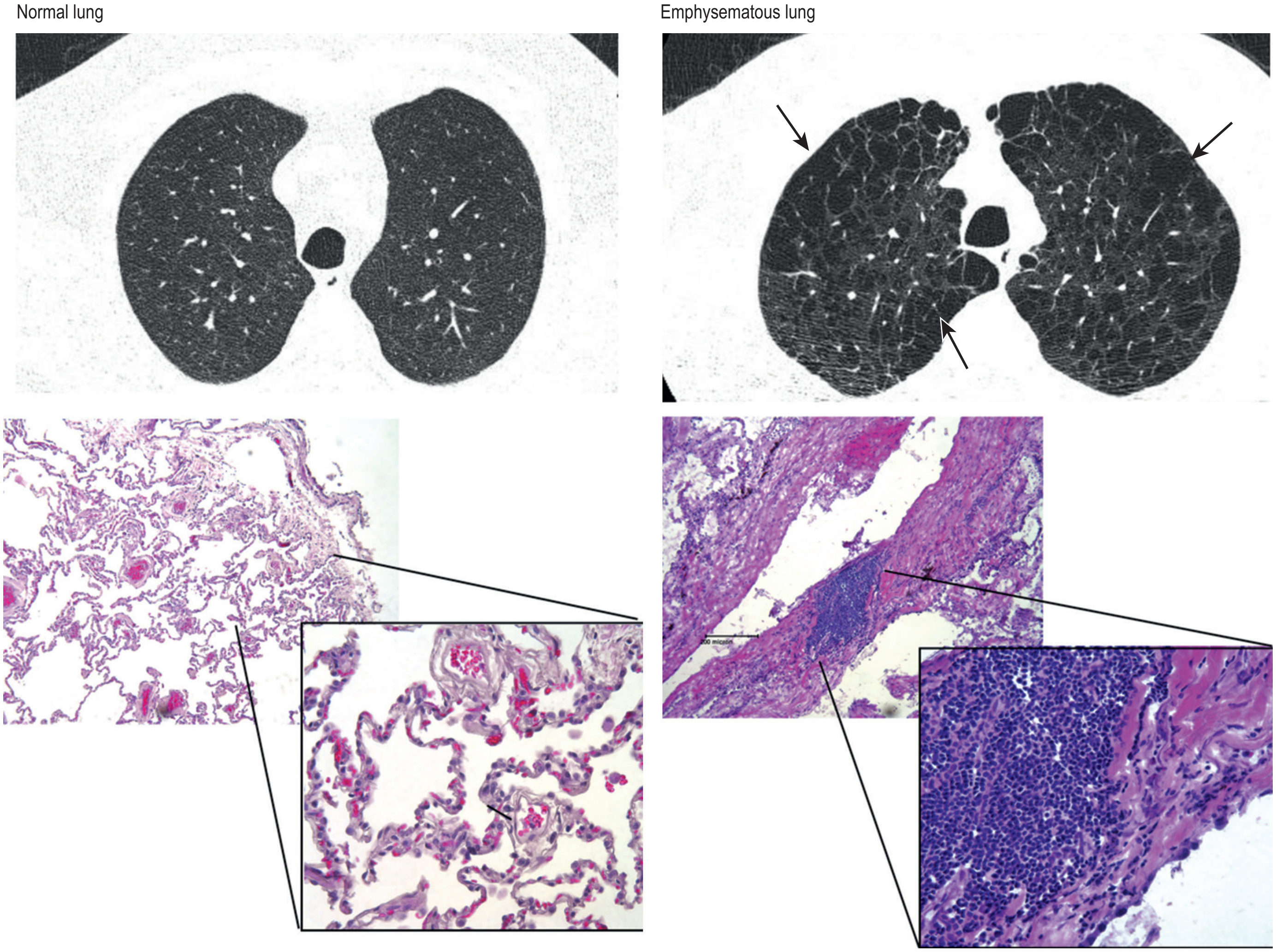


FIG 41.9 **Lung Parenchymal Destruction in Smokers With Chronic Obstructive Pulmonary Disease (COPD) and Emphysema.** Representative computed tomography (CT) images show lungs of a smoker without (A) and with (B) emphysema; arrows point to low attenuation dark black areas devoid of lung tissue (emphysematous regions) within the diseased lung. Lower panels show representative low- and high-magnification histology images from human lung in each case including the focal collection of inflammatory mononuclear cells in emphysematous lung.

NOVEL PATHWAYS TO THERAPY IN INFLAMMATORY AIRWAY DISEASE

Intensive research conducted over the past several decades has markedly improved our understanding of the immune and environmental basis of allergic and nonallergic airway diseases. Currently available therapies for these diseases have, however, failed to match this level of sophistication. It is now clear that allergic disease is immunologically complex and likely to involve multiple hypersensitivity mechanisms operating in parallel. Recent studies also show that factors other than allergens likely critically influence the airway immune response to inhaled antigens and particulate matter. These adjuvant factors include cell wall products of bacteria and fungi, such as lipopolysaccharide (LPS); secreted factors, such as proteases; and endogenous factors from cells damaged by exposure to allergens and smoke. There is increasing evidence that asthma and RS may be linked to respiratory tract infections involving viruses, especially HRV, and filamentous fungi. Future therapies are, therefore, likely to focus on both endogenous factors that coordinate allergic inflammation (*e.g.*, cytokines) and etiological environmental factors (*e.g.*, proteases, fungi, and nCB).

Once thought of as primarily a physical barrier, the respiratory epithelium is now recognized to actively coordinate immunity to allergens and other environmental challenges. Through release of IL-33, IL-25 and/or TSLP, respiratory epithelial cells can initiate and coordinate the innate and adaptive type 2 immune response (see Fig. 41.7). In addition, respiratory epithelial cells produce several chemokines in response to environmental triggers, recruiting DCs, ILC2s, basophils, eosinophils, and mast cell progenitors.[43] Several of the triggers listed above and also trauma to airway cells can activate the release of these type 2 cytokines and chemokines.

The prototypical epithelial cell cytokines, IL-25, IL-33, and TSLP, share many overlapping functions supporting a local type 2 inflammatory response despite having unique receptors. As expected, many of the type 2 cells, including eosinophils, DCs, ILCs, and mast cells, express receptors to all three cytokines. Despite functional overlap, these cytokines also possess unique functions. For example, IL-33 and IL-25 activate different types of ILCs, with IL-33 primarily activating ILC2 and IL-25 preferentially stimulating a novel population of type 2 innate cells. TSLP is associated more with activation of the adaptive type 2 inflammatory response via DCs. Also, TSLP uniquely supports the hematopoiesis of basophils. Together, these epithelial cell derived cytokines orchestrate a robust innate and adaptive type 2 inflammatory response as well as initiating necessary repairs to the epithelial cell barrier.[43]

ON THE HORIZON

- Increasingly sophisticated clinical trials focused on interrupting key immunological pathways (*i.e.*, interleukin [IL]-4, IL-13 signaling pathway) in allergic disease
- Focused research on the importance of infectious pathogens, microbial adjuvant moieties, and air pollution in allergic disease initiation and perpetuation
- Development of immunologically based diagnostic and prognostic strategies for asthma, chronic sinusitis, and possibly other allergic airway disorders
- Clinical studies that reassess the validity of the predominant use of general immunosuppression in comparison to targeted antiinflammatory approaches

Recent clinical studies have confirmed the importance of Th2- and ILC-derived cytokines, including IL-5 and IL-13, to the pathogenesis of human allergic disease. The anti–IL-5 antibody mepolizumab was recently approved by the US Food and Drug Administration (FDA) for use in eosinophilic asthma; anti–IL-13 (lebrikizumab) and anti–IL-4/IL-13 receptor (dupilumab) antibodies are likely to be approved for similar indications.

In addition to defining disease populations that are most likely to benefit from them, a major challenge with all biological agents is their prohibitive cost. Development of small, relatively easily manufactured molecules that antagonize disease-related pathways is, in many ways, a more attractive approach with lower pharmaceutical costs. Small molecule agonists of peroxisome proliferator activated receptor-γ (PPAR-γ) are currently approved for use in diabetes mellitus (pioglitazone, rosiglitazone) in the United States, but they are also outstanding antagonists of Th17 cells and can reverse experimental emphysema.[44] Thus PPAR-γ agonists are excellent candidates for future clinical trials in emphysema. Small molecule antagonists of IL-4/IL-13 signaling that target STAT6 have also been discovered and are potentially suitable for clinical trial and development.[45]

In future studies, it will be essential to clarify the role of chronic infections caused by viruses, bacteria, and fungi in asthma and RS. Although recent studies firmly implicate a central infectious role for fungi in CRSwNP and asthma, it is a major challenge to design clinical trials that unequivocally demonstrate whether antifungal antibiotics are useful in asthma and CRS.[9]

The central role played by proteases and the proteolytic activation of complement (*e.g.*, C3) and coagulation-related inflammatory pathways (*e.g.*, fibrinogen) in allergic diseases suggest that antiproteases might be effective in the management of these conditions. Broad-spectrum antiproteases are effective in attenuating human and experimental asthma and are, thus, suitable for further therapeutic development.[37] The major obstacles with these agents are discovering molecules with sufficiently broad spectrum to inhibit the proteases of the many fungi now linked to human allergic disease and overcoming potential safety issues related to their anticoagulant and immunosuppressive properties.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Bousquet J, Khaltaev N, Cruz AA, et al. Allergic Rhinitis and its Impact on Asthma (ARIA) 2008 update. Allergy 2008;63:8–160.
2. Fokkens W, Lund V, Mullol J. European position paper on rhinosinusitis and nasal polyps 2007. Rhinol Suppl 2007;20:1–136.
3. Calabria CW, Dietrich J, Hagan L. Comparison of serum-specific IgE (ImmunoCAP) and skin-prick test results for 53 inhalant allergens in patients with chronic rhinitis. Allergy Asthma Proc 2009;30:386–96.
4. Hsueh WD, Conley DB, Kim H, et al. Identifying clinical symptoms for improving the symptomatic diagnosis of chronic rhinosinusitis. Int Forum Allergy Rhinol 2013;3:307–14.
5. Bent JP 3rd, Kuhn FA. Diagnosis of allergic fungal sinusitis. Otolaryngol Head Neck Surg 1994;111:580–8.
6. Steinke JW, Wilson JM. Aspirin-exacerbated respiratory disease: pathophysiological insights and clinical advances. J Asthma Allergy 2016;9:37–43.
7. Meltzer EO. The role of nasal corticosteroids in the treatment of rhinitis. Immunol Allergy Clin North Am 2011;31:545–60.

8. Solelhac G, Charpin D. Management of allergic rhinitis. F1000Prime Rep 2014;6:94.
9. Porter PC, Lim DJ, Maskatia ZK, et al. Airway surface mycosis in chronic TH2-associated airway disease. J Allergy Clin Immunol 2014;134:325–31.
10. Rains BM 3rd, Mineck CW. Treatment of allergic fungal sinusitis with high-dose itraconazole. Am J Rhinol 2003;17:1–8.
11. Corry DB, Irvin CG. Promise and pitfalls in animal-based asthma research: building a better mousetrap. Immunol Res 2006;35:279–94.
12. Pavord ID, Korn S, Howarth P, et al. Mepolizumab for severe eosinophilic asthma (DREAM): a multicentre, double-blind, placebo-controlled trial. Lancet 2012;380:651–9.
13. Iyer VN, Lim KG. Bronchial thermoplasty: where there is smoke, there is fire. Allergy Asthma Proc 2015;36:251–5.
14. Lee JH, Park HK, Heo J, et al. Drug Rash with Eosinophilia and Systemic Symptoms (DRESS) syndrome induced by celecoxib and anti-tuberculosis drugs. J Korean Med Sci 2008;23:521–5.
15. Knutsen AP, Bush RK, Demain JG, et al. Fungi and allergic lower respiratory tract diseases. J Allergy Clin Immunol 2012;129:280–91, quiz 292–3.
16. Swartz J, Stoller JK. Acute eosinophilic pneumonia complicating *Coccidioides immitis* pneumonia: a case report and literature review. Respiration 2009;77:102–6.
17. Uchiyama H, Suda T, Nakamura Y, et al. Alterations in smoking habits are associated with acute eosinophilic pneumonia. Chest 2008;133:1174–80.
18. Yoshimi M, Nannya Y, Watanabe T, et al. Acute eosinophilic pneumonia is a non-infectious lung complication after allogeneic hematopoietic stem cell transplantation. Int J Hematol 2009;89:244–8.
19. Rose DM, Hrncir DE. Primary eosinophilic lung diseases. Allergy Asthma Proc 2013;34:19–25.
20. Curtis C, Ogbogu P. Hypereosinophilic syndrome. Clin Rev Allergy Immunol 2016;50:240–51.
21. Gioffredi A, Maritati F, Oliva E, et al. Eosinophilic granulomatosis with polyangiitis: an overview. Front Immunol 2014;5:549.
22. Gell PGH, Coombs RRA. Clinical aspects of immunology. Oxford, England: Blackwell; 1963.
23. Sica A, Mantovani A. Macrophage plasticity and polarization: in vivo veritas. J Clin Investig 2012;122:787–95.
24. Fireman P. Understanding asthma pathophysiology. Allergy Asthma Proc 2003;24:79–83.
25. Zhu J. T helper 2 (Th2) cell differentiation, type 2 innate lymphoid cell (ILC2) development and regulation of interleukin-4 (IL-4) and IL-13 production. Cytokine 2015;75:14–24.
26. Cajaiba MM, Borralho P, Reyes-Mugica M. The potentially lethal nature of bronchial casts: plastic bronchitis. Int J Surg Pathol 2008;16:230–2.
27. Molfino NA, Nannini LJ, Martelli AN, et al. Respiratory arrest in near-fatal asthma. N Engl J Med 1991;324:285–8.
28. Tomankova T, Kriegova E, Liu M. Chemokine receptors and their therapeutic opportunities in diseased lung: far beyond leukocyte trafficking. Am J Physiol Lung Cell Mol Physiol 2015;308:L603–18.
29. Davoine F, Lacy P. Eosinophil cytokines, chemokines, and growth factors: emerging roles in immunity. Front Immunol 2014;5:570.
30. Zhang X, Kohl J. A complex role for complement in allergic asthma. Exp Rev Clin Immunol 2010;6:269–77.
31. Kanaoka Y, Boyce JA. Cysteinyl leukotrienes and their receptors; emerging concepts. Allergy Asthma Immunol Res 2014;6:288–95.
32. Bonnans C, Levy BD. Lipid mediators as agonists for the resolution of acute lung inflammation and injury. Am J Respir Cell Mol Biol 2007;36:201–5.
33. Porter P, Polikepahad S, Qian Y, et al. Respiratory tract allergic disease and atopy: experimental evidence for a fungal infectious etiology. Med Mycol 2011;49:S158–63.
34. Newcomb DC, Peebles RS Jr. Bugs and asthma: a different disease? Proc Am Thorac Soc 2009;6:266–71.
35. Dozor AJ. The role of oxidative stress in the pathogenesis and treatment of asthma. Ann NY Acad Sci 2010;1203:133–7.
36. Porter PC, Yang T, Luong A, et al. Proteinases as molecular adjuvants in allergic airway disease. Biochim Biophys Acta 2011;11:1059–65.
37. Millien VO, Lu W, Mak G, et al. Airway fibrinogenolysis and the initiation of allergic inflammation. Ann Am Thorac Soc 2014;11:S277–83.
38. Richerson HB, Bernstein IL, Fink JN, et al. Guidelines for the clinical evaluation of hypersensitivity pneumonitis. Report of the Subcommittee on Hypersensitivity Pneumonitis. J Allergy Clin Immunol 1989;84:839–44.
39. Hobbs B, Parker M, Chen H, et al. Exome array analysis identifies a common variant in IL27 associated with chronic obstructive pulmonary disease. Am J Respir Crit Care Med 2016;194:48–57.
40. Vestbo J. COPD: definition and phenotypes. Clin Chest Med 2014; 35:1–6.
41. You R, Lu W, Shan M, et al. Nanoparticulate carbon black in cigarette smoke induces DNA cleavage and Th17-mediated emphysema. elife 2015;4:e09623.
42. Kheradmand F, Shan M, Xu C, et al. Autoimmunity in chronic obstructive pulmonary disease: clinical and experimental evidence. [Review]. Exp Rev Clin Immunol 2012;8:285–92.
43. Hammad H, Lambrecht BN. Barrier epithelial cells and the control of type 2 immunity. Immunity 2015;43:29–40.
44. Shan M, You R, Yuan X, et al. Agonistic induction of PPAR reverses cigarette smoke-induced emphysema. J Clin Invest 2014;124: 1371–81.
45. Mandal PK, Morlacchi P, Knight JM, et al. Targeting the Src homology 2 (SH2) domain of signal transducer and activator of transcription 6 (STAT6) with cell-permeable, phosphatase-stable phosphopeptide mimics potently inhibits Tyr641 phosphorylation and transcriptional activity. J Med Chem 2015;58:8970–84.

MULTIPLE-CHOICE QUESTIONS

1. What epithelial cell–derived cytokine contributes to the pathophysiology of chronic rhinosinusitis (CRS) by serving as the primary driver of mucus production?
   A. Thymic stromal lymphopoietin (TSLP) via activation of dendritic cells
   B. Chemokine ligand 11 (CXCL11) via its chemotactic role for activated T cells
   C. Interleukin-33 (IL-33) via activation of innate lymphoid cells
   D. IL-25 via increased expression of IL-4

2. Which of the following characteristics MOST clearly distinguishes chronic obstructive pulmonary disease/emphysema from asthma?
   A. Airway obstruction
   B. Lymphocytic inflammation
   C. Disease exacerbations leading to disabling dyspnea
   D. Autoimmune responses against elastin
   E. Disease is linked to environmental exposures

3. Which of the following factors is most clearly implicated as an etiological environmental factor in human emphysema?
   A. Molds (fungi)
   B. Nanoparticulate carbon black (nCB)
   C. *Haemophilus influenzae*
   D. Endotoxin (lipopolysaccharide; LPS)
   E. Radon gas

4. Future therapies for asthma could target which of the following?
   A. IL-13
   B. IL-4 receptor
   C. IL-5
   D. Fungi
   E. All of the above

42

Urticaria, Angioedema, and Anaphylaxis

Clive E.H. Grattan, Elena Borzova

Urticaria is a common skin disorder that can lead to severe impairment in quality of life. Recently, there has been substantial progress in our knowledge and understanding of the pathophysiology of the condition, offering new diagnostic and treatment approaches for many patients.

DEFINITION

Urticaria is characterized by itchy wheals, angioedema, or both.[1] Urticarial lesions resulting from localized edema of the upper dermis are called *wheals* (Fig. 42.1), whereas pronounced swelling of deeper dermal layers and subcutaneous and submucosal tissues is known as *angioedema* (Fig. 42.2).

KEY CONCEPTS

Definition of Urticaria

- Urticaria is an illness characterized by wheals, angioedema, or both.
- Wheals are superficial swellings of the dermis of skin: they are pale and itchy with surrounding redness when they appear and then become pink before fading.
- Angioedema is a deep swelling below the skin or mucosa that usually lasts longer than wheals and may be painful rather than itchy.
- Wheals may occur with or without angioedema in urticaria.
- Acute urticaria lasts <6 weeks. It is common and often caused by viral infections. Urticaria caused by drugs and foods falls in this category, but the diagnosis is usually clear from the history.
- Chronic urticaria lasts ≥6 weeks with continuous disease activity. A cause may not be found. Urticaria occurring intermittently over a period of >6 weeks can be defined as episodic.
- When angioedema occurs without wheals, C1 esterase inhibitor deficiency should be excluded.

EPIDEMIOLOGY

Urticaria is common, and up to 25% of the general population experience some form of urticaria at least once over their lifetime, whereas the lifetime prevalence of chronic urticaria (CU) is 0.1–3% of the general population.[2] Acute urticaria mainly affects young adults and has an obvious female preponderance,[2] whereas CU is more common in adults, affecting mainly middle-aged women, and is uncommon in children and adolescents.[3] By definition, acute urticaria resolves within 6 weeks, whereas CU may last for years. Rarely, CU may last >10 years. However, CU is almost always a self-limiting disorder, resolving spontaneously within 6 months of onset in 50% of patients.[3]

GENETICS

Although there is no indication of mendelian inheritance, the frequency of chronic spontaneous urticaria (CSU) is higher in first-degree relatives of patients.[4] A link between CU and human leukocyte antigen (HLA)-DRB1*04 (DR4) and its associated allele, -DQB1*0302 (DQ8), was found in British patients with CSU with evidence of functional autoantibodies compared with a control population.[5] Strong associations with HLA-DRB1*1302 and DQB1*0609 alleles were found in aspirin-induced urticaria in Korean patients.[6]

CLINICAL PATTERNS

Urticaria is classified by the duration of continuous activity into acute (<6 weeks), chronic (≥6 weeks), or episodic, where short episodes occur repeatedly over a period, with short or long intervals. It may be spontaneous, inducible, or both. Inducible urticarias may be triggered by mechanical, thermal, or other stimuli. The differential diagnosis of urticaria includes urticarial vasculitis, hereditary angioedema (HAE), and autoinflammatory syndromes presenting with urticarial rash.

ETIOPATHOGENESIS AND ETIOLOGICAL CLASSIFICATION

Although many aspects of the pathophysiology of urticaria remain unclear, our understanding has advanced considerably over the last two decades, allowing for an etiological classification (Table 42.1).

Mast Cell–Dependent Mechanisms

Skin mast cells are key players in the pathogenesis of urticaria. They are predominantly located around the small blood vessels and lymphatic vessels as well as around or within peripheral nerves. The population density of mast cells is greatest at distal body areas, including the face, hands, and feet.[7] There is conflicting evidence as to whether the number of cutaneous mast cells is increased in patients with urticaria, but it is accepted that these cells release mediators more readily compared with normal cells.

Human skin mast cells contain preformed mediators in their granules, including histamine, proteases (tryptase, chymase), and heparin. They express many membrane receptors, including high-affinity immunoglobulin E (IgE) and low-affinity IgG receptors. Unlike mast cells elsewhere (*e.g.*, lung and gastrointestinal [GI] tract), skin mast cells possess complement C5a

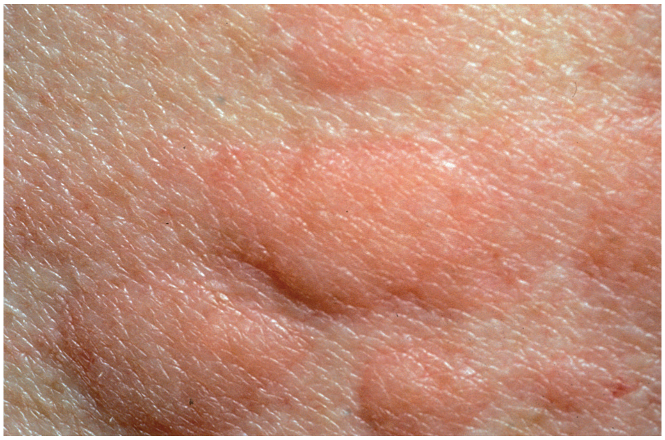


FIG 42.1 Spontaneous wheals in severe spontaneous urticaria showing superficial pink swellings with pale edematous centers.

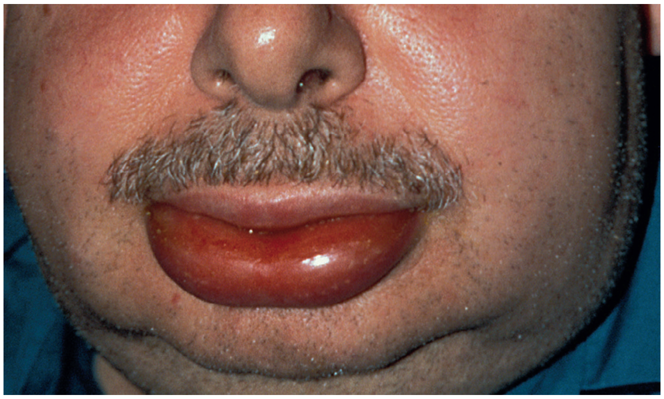


FIG 42.2 Angioedema of the mouth in acquired C1 esterase inhibitor deficiency.

TABLE 42.1 Etiopathogenesis of Urticaria

Acute
Idiopathic
Infection-related
Allergic (mediated by specific IgE)

Chronic
Idiopathic
Autoimmune (IgG against IgE or FcεRI)
Infection-related
Drug-induced
Diet-related

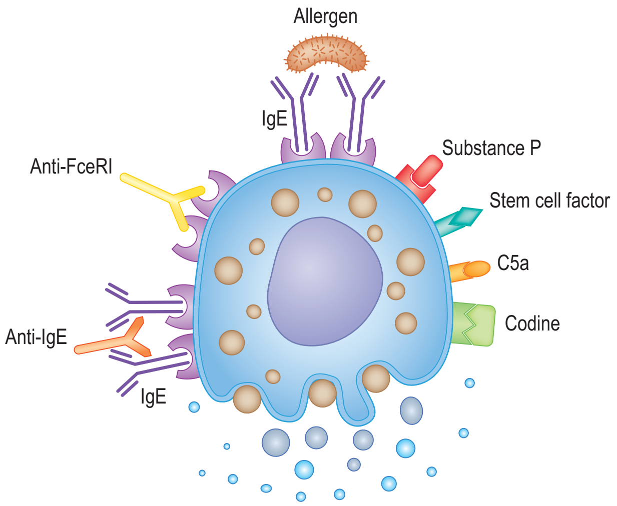


FIG 42.3 Schematic representation of a mast cell or basophil illustrating activation of the immunoglobulin E (IgE) receptor by cross-linking with immunological stimuli (allergen/specific IgE binding, anti-IgE, or anti-FcεRI autoantibodies) or independent activation by nonimmunological stimuli (substance P, stem cell factor, codeine, or C5a) leading to degranulation.

receptors and activation sites for neuropeptides and basic secretagogues.

Skin mast cell activation is central to the pathophysiology of CU. Mast cells can be activated by a variety of immunological and nonimmunological triggers (Fig. 42.3). Immunological and nonimmunological pathways of mast-cell activation are characterized by distinct patterns of mediator release. Immunological activation of mast cells is triggered by cross-linking of high-affinity IgE receptors (FcεRI) by antigen bound to antigen-specific IgE, by anti-FcεRIα, or by anti-IgE antibodies. Histamine release peaks at 5–10 minutes, followed by *de novo* synthesis of lipid-derived mediators (leukotriene C4 [LTC4] and prostaglandin D_2 [PGD_2]) and cytokines.[8] In contrast, nonimmunological stimulation of mast cells by neuropeptides, opiates, or C5a leads to rapid release of histamine within 15–20 seconds, without generation of eicosanoids and cytokines. Moreover, prolonged and subthreshold immunological stimulation may result in a state of receptor desensitization. For example, desensitization of FcεRI may lead to basophil hyporesponsiveness to anti-IgE in autoimmune CSU.[9] However, desensitization of receptors by immunological stimulation does not affect nonimmunological release.

Allergic Urticaria

The classic example of immunological mast-cell activation via high-affinity IgE receptors is IgE-mediated urticaria (often termed *allergic urticaria*). Cross-linking of receptor-bound IgE leads to the release of diverse preformed mediators and newly synthesized lipid mediators and cytokines, resulting in the early- and late-phase IgE-mediated allergic inflammatory responses.

IgE-mediated mast-cell activation can present as acute urticaria in those individuals previously sensitized to exogenous allergens. IgE to autoantibodies are implicated in chronic wheeling. Examples include some food- and drug-induced urticarias and latex-induced contact urticaria. Allergic urticaria to inhaled allergens (*e.g.*, latex, animal epithelia) is often accompanied by respiratory symptoms. Generalized allergic urticaria may progress to anaphylaxis. Allergic urticaria resolves rapidly on withdrawal of allergen exposure and recurs with each reexposure to the allergen or cross-reactive agents.

Autoimmune Urticaria

Functional autoantibodies directed against the extracellular α-chain of FcεRI on dermal mast cells or basophils or, less frequently, against receptor-bound IgE have been demonstrated in 25–50% of adult and pediatric patients with CSU.[9,10] The pathophysiology of autoimmune CU involves cross-linking of high-affinity IgE receptors by autoantibodies leading to degranulation of mast cells and basophils. There is good evidence that activation of

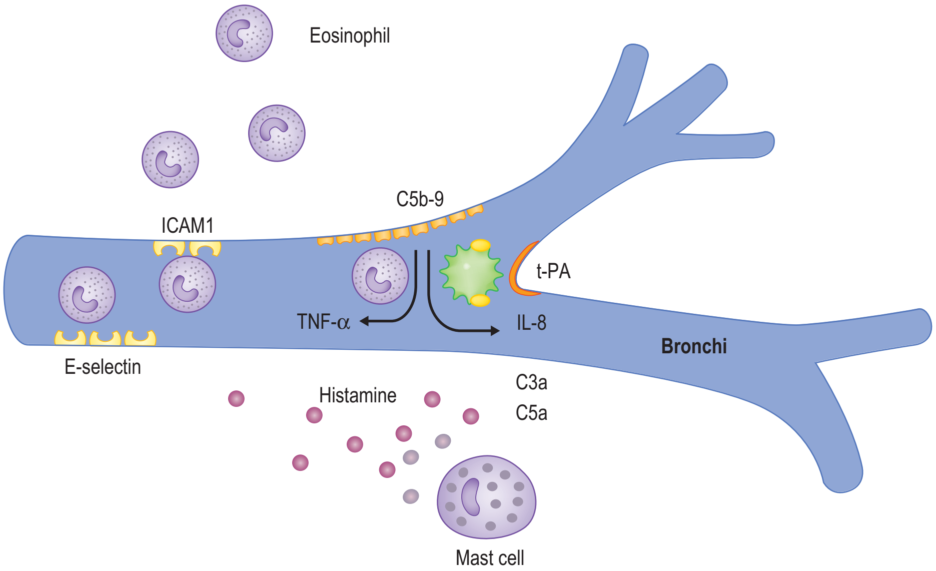


FIG 42.4 Immune Complex-Mediated Urticaria Initiated by Lodging of Antigenic Complexes in Small Blood Vessels Followed by C3a and C5a Generation. This results in mast cell degranulation and cytokine upregulation of adhesion molecules (intercellular adhesion molecule [ICAM]-1, E-selectin) by tumor necrosis factor-α (TNF-α) and interleukin-8 (IL-8), which leads to tissue recruitment of neutrophils and eosinophils and activation of tissue plasminogen activator (t-PA).

cutaneous mast cells by IgG derived from CSU sera *in vitro* is dependent on C5a, but not all histamine-releasing IgG derived from CSU sera are complement dependent on basophil assays. Most functional autoantibodies are IgG1 or IgG3, which are known to fix complement. By contrast, nonfunctional anti-FcεRI autoantibodies demonstrated by Western blot or enzyme-linked immunosorbent assay (ELISA) in other autoimmune diseases (including systemic lupus erythematosus [SLE] and bullous pemphigoid) are generally from the noncomplement-fixing subclasses IgG2 and IgG4.[10]

The low-affinity IgE receptor FcεRII/CD23 expressed on B lymphocytes and eosinophils may be another target antigen for an autoimmune response in CSU. Anti-CD23 antibodies have been detected in patients with CSU.[11] The interaction between anti-CD23 antibodies and CD23 on eosinophils may induce release of major basic protein 1 (MBP-1), a potent IgE-independent histamine-releasing agent, resulting in activation of mast cells and basophils.

Immune Complex–Mediated Urticarial Rash

Mast-cell activation can result from binding of circulating immune complexes to FcγRIII, expressed on mast cells. In addition, circulating immune complexes can activate complement, leading to C3a and C5a anaphylatoxin formation. Urticarial rash caused by immune complexes can occur acutely in serum sickness–like reactions, transfusion reactions, some drug-induced urticarias, and urticaria associated with infectious or autoimmune diseases. Immune complex–mediated urticarial rashes usually develop 1–3 weeks after initial exposure to the antigen and disappear several weeks after antigen discontinuation. Chronic wheeling mediated by immune-complex damage and associated with a histological evidence of leukocytoclastic vasculitis is known as *urticarial vasculitis* (see below). In this condition, urticaria can be associated with systemic symptoms, such as fever, arthritis, or nephritis, especially the uncommon hypocomplementemic variant (hypocomplementemic urticarial vasculitis syndrome [HUVS]). Damage to postcapillary venules results from the deposition of immune complexes in the vessel wall. Immune complexes are formed on exposure to external (drug or infections) or internal (collagen-like region of C1q) antigens. Complement activation via the classical pathway leads to neutrophil chemotaxis through cytokine expression and adhesion molecule activation (Fig. 42.4). Proteolytic enzymes released from neutrophils damage vessel walls, leading to wheal formation and red blood cell (RBC) extravasation.

Nonimmunological Mast-Cell Activation

Skin mast cells can be activated nonimmunologically by many agents, including neuropeptides (substance P, neuropeptide Y, vasoactive intestinal peptide, or somatostatin), calcitonin gene-related protein, compound 48/80, and natural polyamines. Drugs (*e.g.*, opiates, muscle relaxants, radiocontrast media, polymyxin B) can also cause dose-dependent nonimmunological urticaria by activating membrane-associated G-coupled proteins. A classic clinical example is codeine-induced acute urticaria.

Mast-Cell and Basophil Releasability in Urticaria

In patients with CSU, dermal mast cells show a decreased activation threshold. Although CSU basophils are often hyporesponsive to anti-IgE, they are paradoxically hyperresponsive to an as yet unidentified factor in normal human serum.[10]

Skin Response to Mast-Cell Activation in Chronic Urticaria

Despite the existence of different pathways for mast-cell activation, the end result is degranulation with release or secretion of mediators. Histamine is crucial for the development of cutaneous manifestations of urticaria and is found in high concentrations in tissue fluid of wheals. It causes localized redness resulting from vasodilatation, wheal formation from increased vascular permeability leading to plasma leakage, and a surrounding axon reflex-mediated flare-up. Histamine is also the main mediator of itch in urticaria. However, histamine-induced wheals are short-lived in contrast to urticarial lesions that may persist up to 24 hours, implying that other proinflammatory mediators and/or cellular infiltrates contribute to CSU pathogenesis.

Mast Cell–Independent Mechanisms of Urticaria

Pseudoallergy (Intolerance)

Pseudoallergy, or nonallergic hypersensitivity, mimics immediate-type allergic reactions clinically without evidence of underlying immunological mechanisms.[12] The most common triggers of pseudoallergic reactions are aspirin and other nonsteroidal antiinflammatory drugs (NSAIDs), as well as some food ingredients and additives, such as salicylates, benzoates, and tartrazine. These reactions do not involve IgE sensitization and can, therefore, occur on first exposure. Pseudoallergic reactions are dose-dependent and usually occur with chemically nonrelated substances. The diagnosis is difficult because skin tests and serology are uninformative. Diagnosis of nonallergic hypersensitivity is based on a distinctive clinical pattern, time course, clinical signs, and response to elimination of the cause. In the appropriate clinical context, pseudoallergy can be confirmed with oral challenge tests.

NSAIDs

Aspirin and other NSAIDs inhibit constitutive cyclooxygenase (COX-1) and inducible cyclooxygenase (COX-2), thereby diverting arachidonic acid metabolism toward the 5-lipoxygenase pathway in some cell types, notably eosinophils. This modulation of arachidonic acid metabolism results in overproduction of cysteinyl-leukotrienes LTC4, -D4, and -E4, leading to vasodilatation and edema. Furthermore, reduction of PGE_2 formation by COX inhibition has two further effects that promote urticaria: first, by reducing inhibition of cysteinyl leukotriene production and second, by reducing an inhibitory effect on immunologically-mediated mast-cell degranulation (Fig. 42.5). Cross-sensitivity occurs with other nonselective NSAIDs in susceptible individuals, depending on their pharmacologic potency for COX inhibition but not their chemical structure.

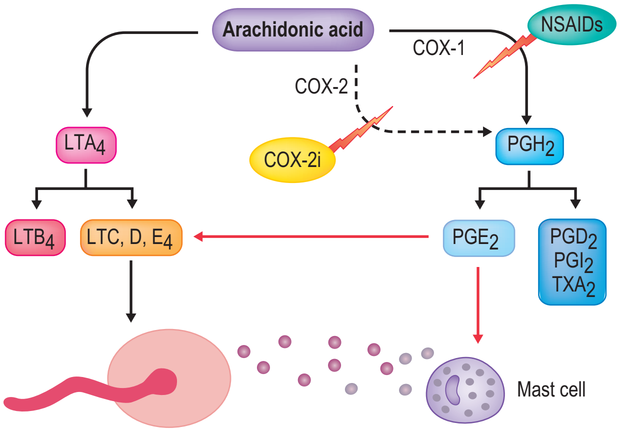


FIG 42.5 Arachidonic acid pathways illustrating potential diversion from prostaglandin synthesis to cysteinyl leukotrienes (LTC4, -D4, -E4) by blocking cyclooxygenase (COX), leading to increased vasopermeability. Reduction of prostaglandin E_2 also reduces its direct inhibitory effect on leukotriene production and on immunological mast cell degranulation.

Aspirin and NSAIDs can trigger acute urticaria as well as causing flare-ups of preexisting chronic spontaneous (but not physical) urticaria. Aspirin-induced exacerbations have been reported in 20–40% of patients with CSU.[13] In some patients, aspirin can act as a cofactor in food- or exercise-induced anaphylaxis.

Urticaria can develop from a few minutes to 24 hours after aspirin ingestion but usually within 1–2 hours. Angioedema of the lips and tongue, impaired swallowing, and laryngeal edema develop occasionally. Aspirin-induced skin symptoms usually subside 24–48 hours after discontinuing the drug. However, severe exacerbations of CSU caused by aspirin and other NSAIDs can last from several days to several weeks after aspirin intake.[13]

There are no skin tests or reliable *in vitro* diagnostic techniques for patients with aspirin-induced urticaria, and the diagnosis can only be established by challenge tests. Patients should avoid aspirin and other NSAIDs, but COX-2 inhibitors (coxibs) are usually well tolerated and can probably be used safely.

Food-Induced Pseudoallergic Reactions in CSU

Pseudoallergic food reactions appear to be important in some patients with CSU. Pseudoallergic food triggers include natural salicylates in fruit and vegetables and artificial food additives in processed foods, such as benzoates and tartrazine. Low-molecular-weight aromatic compounds in tomatoes, white wine, and herbs have also been implicated.[14] Clinically, exacerbations of CSU resulting from dietary pseudoallergens gradually subside within 10–14 days on an exclusion diet; in contrast, in 1–3 days in acute allergic urticaria. One study showed increased gastroduodenal and intestinal permeability in patients with CU. Responders to elimination of pseudoallergens in their diet demonstrated normalization of mucosal permeability and skin symptoms. Although the underlying mechanisms for pseudoallergic reactions in CU remain unproven, an impaired gastroduodenal barrier function may be a contributory factor.[14]

Kinin-Mediated Angioedema

The mediator of angioedema resulting from hereditary and acquired C1 esterase inhibitor (C1-INH) deficiency and angiotensin-converting enzyme inhibitor (ACEI)–induced angioedema is bradykinin rather than histamine or leukotrienes. There is currently no evidence that kinins are involved in other types of urticaria or angioedema. Bradykinergic angioedema is not considered a type of urticaria.

CLINICAL CLASSIFICATION

CU may be spontaneous, inducible, or both.

Spontaneous Urticaria

Spontaneous urticaria is the most common presentation. The term “spontaneous” makes no assumption about etiology, which may include autoimmunity, allergy, pseudoallergy, or infection. Those cases for which a cause cannot be ascertained with reasonable confidence are called *idiopathic*. In practice, this is the majority of patients. Spontaneous urticaria can be further classified by duration and frequency of attacks into acute, chronic, and episodic urticaria.

Acute Spontaneous Urticaria

Acute urticaria is defined as continuous disease lasting <6 weeks. Individuals with an atopic predisposition are at higher risk of acute urticaria; atopic diseases are found in about half the patients with acute urticaria.[3] Most acute urticaria resolves spontaneously within 3 weeks, but in 10% of patients, it may progress to CSU.

Infections are the most common identifiable cause of acute urticaria, with some cases resulting from foods or drugs. A specific cause of acute urticaria will not be found in about 50% of patients. Foods are the most common cause of acute urticaria in young children but are rarely responsible for acute urticaria in adults. In infancy, cow's milk allergy may cause acute allergic urticaria. Drug-induced urticaria is the most common presentation of drug hypersensitivity, accounting for a quarter of all adverse drug reactions, with penicillin and NSAIDs being the most common causes of allergic and nonallergic drug-induced urticaria, respectively. Drug-induced urticaria is more likely in older adults, perhaps reflecting polypharmacy and age-related pharmacokinetic changes, and in patients with human immunodeficiency virus (HIV) infection and those with renal or liver diseases.

Chronic Spontaneous Urticaria

CSU is characterized by daily or almost daily itchy wheals on skin, with or without angioedema, for ≥6 weeks. In 30–50% of patients, CSU is caused by functional autoantibodies against the α chain of the high-affinity IgE receptor or against IgE itself.[10] Functional abnormalities of basophils have also been described. However, in 50% of patients, the etiopathophysiology of the disease remains unknown and appears to be endogenous. Whatever the primary cause, CSU appears to be aggravated by a variety of exogenous exposures, including acute infections (most commonly viral upper respiratory tract infections), NSAIDs, dietary pseudoallergens, menses in women, tiredness, and stress. The presence of these factors probably explains why the clinical course is often erratic and unpredictable.

CSU affects women twice as often as men. There is a common link between CSU and thyroid autoimmunity (present in about 14–25% of CU cases) as well as other autoimmune diseases, including Graves disease, vitiligo, and insulin-dependent diabetes. CSU is often associated with inducible urticaria (*e.g.,* delayed-pressure urticaria [DPU]). Around 50% of patients present with both wheals and angioedema. CU can have a continuous or a relapsing course. CU of all types, including CSU, can cause serious disability in patients, including loss of sleep and energy, social isolation, altered emotional reactions, and difficulties in aspects of daily living similar in degree to patients with severe ischemic heart disease.

Episodic Spontaneous Urticaria

The cause of episodic urticaria often remains unknown, but the possibility of an allergic or pseudoallergic cause needs to be considered.

Inducible Urticarias

Inducible urticarias are common, accounting for 25% of all cases of CU. These include physical urticarias, such as symptomatic dermographism, cold and heat urticarias, DPU, solar urticaria, and vibratory angioedema, in which symptoms are triggered reproducibly by an external mechanical or thermal/ultraviolet stimulus. The inducible group now also includes cholinergic, adrenergic, aquagenic, and contact urticarias, in which the eliciting stimuli for mast-cell degranulation are defined by a nonphysical exposure. Inducible urticarias can coexist with CSU (*e.g.,* DPU or dermographism), and more than one inducible urticaria can occur in the same patient (*e.g.,* dermographic and cholinergic urticarias), which may lead to difficulty in diagnosis. With the exception of DPU, inducible urticarias develop rapidly after exposure to the relevant trigger and fade within an hour.[2]

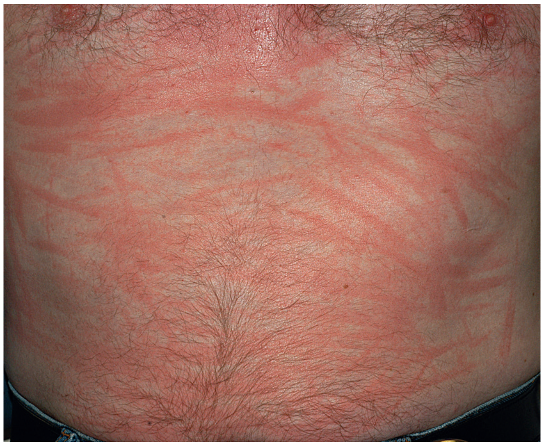


FIG 42.6 Extensive induced dermographic whealing on the chest of a patient as a result of scratching.

The pathogenesis of inducible urticarias is unclear except in the case of allergic contact urticaria, which is caused by mucocutaneous exposure to an allergen to which the individual has already developed IgE sensitization. A diagnosis of physical urticaria is confirmed if the symptoms can be reproduced by challenge testing with the suspected stimulus.[15] Challenge testing can be used for monitoring threshold changes during treatment. In general, treatment of inducible urticarias involves avoidance of known triggers and taking antihistamines. Sometimes tolerance can be induced for cold and solar urticarias.

Mechanical Urticaria

Symptomatic Dermographism

Dermographism is the most common physical urticaria, mainly affecting young people. Typical red, itchy, linear wheals are evoked within minutes of stroking, friction, rubbing, or scratching the skin (Fig. 42.6). Overheating, stress, and anxiety usually aggravate symptoms. The diagnosis of dermographism is confirmed by stroking the skin with a blunt firm object or a calibrated instrument, such as a dermographometer (HTZ, Croydon, U.K.) or FricTest (Moxie, Berlin, Germany), usually on the upper back.

Delayed Pressure Urticaria

Isolated DPU occurs in just 2% of all patients with urticaria, but DPU coexists with CSU in up to 40% of patients. DPU is the most debilitating of the physical urticarias and is triggered by sustained local pressure (*e.g.,* wearing tight shoes, carrying heavy bags, long walks, sitting or leaning against firm objects, climbing ladders, jogging, driving, or clapping hands). Deep and painful swellings, clinically resembling angioedema, develop 30 minutes to 12 hours after pressure, and may be associated with flu-like symptoms, fever, arthralgia, and fatigue. The most frequently affected sites are hands, soles, buttocks, shoulders, and areas under straps and belts. DPU lesions last 12–48 hours and are usually painful rather than itchy, especially on the hands and feet. Laboratory investigation reveals transitory leukocytosis and elevated erythrocyte sedimentation rate (ESR). Hanging a heavy weight suspended on a narrow band over the forearm or thigh for 15 minutes may be used as a challenge test, but more reliable results can be obtained with a dermographometer applied at 100 g/mm^2 for 70 seconds. The reaction should be assessed after 2–6 hours. DPU is difficult to treat because it responds poorly to antihistamines.

Vibratory Angioedema

Vibratory angioedema is rare. Familial cases have been described. Local swelling develops several minutes to 6 hours after using vibrating machinery, lawn mowing, applauding, and jogging, for instance. Systemic symptoms may occur (headache, chest tightness, diffuse flare). Placing the elbow or hand on a laboratory vortex for 5–15 minutes is a useful challenge test. Avoidance of the trigger is the only helpful treatment strategy.

Thermal or Ultraviolet-Induced Urticaria

Cold Urticaria

Cold urticaria accounts for about 3% of physical urticarias.[16] It occurs in both children and adults and is more common in cold climates, in women, and in atopic patients. The majority of cases are primary with no identifiable cause, but some cases are secondary to internal disease. Clinical manifestations can be local or generalized. Mucosal involvement may develop after drinking cold beverages. Systemic symptoms may be respiratory (laryngeal angioedema, tongue or pharyngeal swelling, wheezing), vascular (hypotension, tachycardia), GI (hyperacidity, nausea, diarrhea), or neurological (disorientation, headache). Cold-induced urticaria can be evoked by low ambient temperature, contact with cold objects, food or beverages, or immersion in cold water. Wheals develop during the cold exposure or, more commonly, on rewarming. The severity of cold urticaria depends on the intensity and duration of the cold stimulus. Cold urticaria is potentially life threatening, with a risk of anaphylaxis and death on exposure of large skin areas to cold, for example, jumping into cold water and hypothermia in neurosurgical and cardiothoracic operations. Familial cold urticaria is caused by mutations in the cold-induced autoinflammatory syndrome (*NLRP-3*) gene and is no longer classified as an inducible urticaria.

In 1–5% of patients, cold urticaria is secondary to cryoproteins (mainly cryoglobulins). These can be associated with infections (hepatitis C, infectious mononucleosis, syphilis, *Mycoplasma* infection), autoimmune diseases, and lymphoreticular malignancy (Waldenström macroglobulinemia, myeloma), but these are rare. Cold urticaria can precede these diseases by several years. Secondary cold urticaria can also be drug-related (penicillin, oral contraceptives, ACEIs).

The diagnosis of cold urticaria is confirmed by an ice cube challenge or TempTest (Moxie, Berlin, Germany). Some atypical cold-induced urticarias have negative results in the ice cube test.

The clinical workup in cold urticaria includes measurement of cryoproteins. Patients should be cautious about bathing or swimming in cold water and consuming cold food or drinks. Antihistamine treatment is often helpful but does not prevent anaphylaxis caused by swimming in cold water. In severe cold urticaria, tolerance induction may be attempted: this involves depletion of mast-cell histamine by repeated cold exposure.

Heat Urticaria

Heat-induced urticaria is very rare. It is induced by local heating of the skin at 38–44°C. Challenge test is done by application of hot water in a tube or beaker at up to 44°C for 4–5 minutes or TempTest (Moxie, Berlin, Germany). Symptoms develop several minutes after exposure. Management of heat-induced urticaria presents some difficulty, as antihistamines are only of limited value.

Solar Urticaria

Solar urticaria affects about 1% of all patients with urticaria and has a slight female predominance. It can be associated with erythropoietic porphyria. Wheals are caused by electromagnetic wavelengths ranging from 290 to 760 nm (ultraviolet B [UVB], UVA, and visible spectrum). It develops within minutes or hours after sun exposure and fades within 24 hours. Lesions are usually confined to sun-exposed skin, although they can also develop under clothing. The severity of solar urticaria depends on the wavelength, intensity, and duration of irradiation. Short exposures induce flare and pruritus, whereas longer exposures cause whealing. In patients sensitive to the visible spectrum, reactions may occur through window glass.

Solar urticaria is diagnosed by phototesting. Patients are advised to use creams with a high sun protection factor (SPF), protective clothing, and protective window shields and to limit the time spent outdoors.

Other Patterns of Inducible Urticaria

Cholinergic Urticaria

Cholinergic urticaria is the second most common physical urticaria and occurs mainly in adolescents, young adults, and patients with atopy. Cholinergic urticaria usually follows a rise in core temperature resulting from physical exercise, fever, or external passive heat (hot bath, shower, sauna) but may also be provoked by emotional stress and spicy food. The characteristic lesions are highly pruritic pinpoint pale wheals of 1–3 mm surrounded by a red flare. The wheals may occur anywhere except the soles and palms. Lesions usually begin on the trunk and the neck, extending outward to the face and limbs. As lesions progress, confluent areas of whealing and redness may develop. Severely affected patients may develop angioedema and even anaphylaxis. Most patients with mild disease do not seek medical help. The rash is triggered by activation of cholinergic sympathetic innervation of sweat glands, but the mechanism of activation remains unclear. Decreased blood protease inhibitor levels have been reported, and this is the rationale for using anabolic steroids to treat occasional severely affected individuals who are unresponsive to other measures. The prognosis is reasonably favorable, with spontaneous resolution within 8 years in most patients. However, 30% of patients are affected for over 10 years.

Cholinergic urticaria can be confirmed by reproducing the rash with exercise or passive heating in a hot bath at up to 42°C. Treatment is primarily with antihistamines, but beta-blockers, danazol, ketotifen, and montelukast have also been used. The condition may be refractory for up to 24 hours, and this may enable patients to prevent attacks by taking daily exercise.

Aquagenic Urticaria

Aquagenic urticaria is very rare. It occurs in women more often than in men and is triggered by water contact but not after drinking water. Scattered small papular wheals, similar to cholinergic urticaria but with a larger flare, appear within 10–20 minutes of water contact and resolve in 30–60 minutes. Diagnosis is made by using a challenge test in which a wet compress at body temperature is applied for up to 10 minutes on whichever part of the body is usually affected. Associations with HIV and hepatitis B infection have been described.

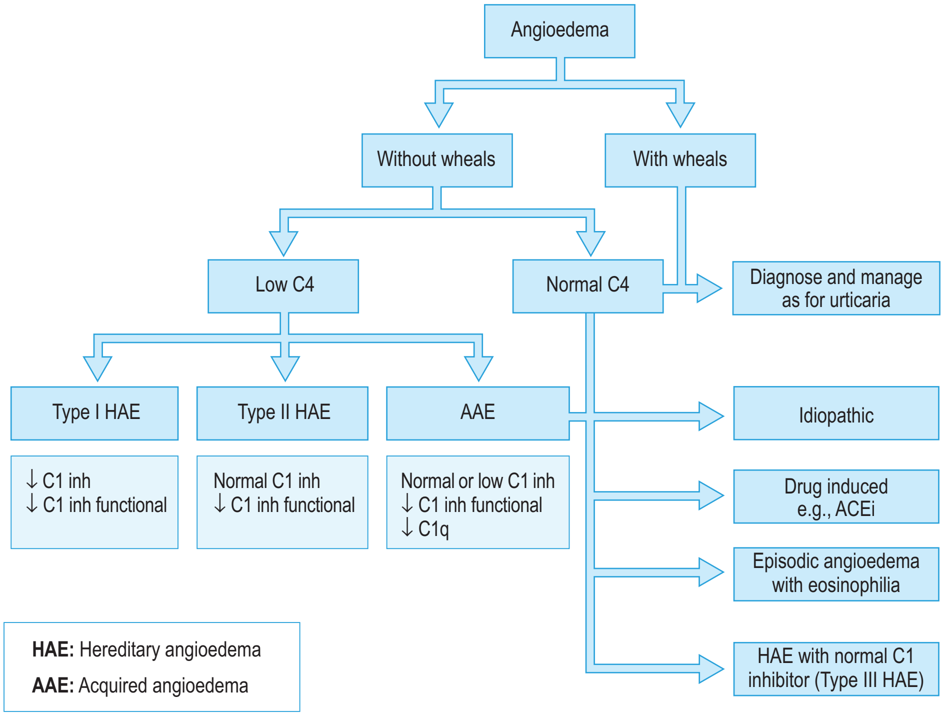


FIG 42.7 Diagnostic pathway for the differential diagnosis of angioedema.

Contact Urticaria

Contact urticaria occurs locally after skin or mucosal contact with the eliciting agent.[2,3] Reactions usually develop within a few minutes and resolve over 2 hours, although delayed-contact urticaria can occur with a latent period up to 48 hours. Contact urticaria can be caused by many organic and inorganic stimuli, such as latex, animal danders and secretions, foods, plants, topical drugs, and cosmetics.

Allergic contact urticaria occurs mainly in atopic subjects. IgE sensitization can occur through skin, mucous membranes, and the respiratory or GI tract. Foods are the most common cause. In ≈15% of patients, contact with the allergen may induce anaphylaxis. The severity of contact reactions depends on many factors (area of exposure, duration of contact, amount and concentration of substance, patient reactivity, comorbidity, and concomitant treatment).

Nonallergic contact urticaria is commoner than allergic contact urticaria and is usually caused by low-molecular-weight chemicals. Contact urticaria to benzoic acid in food is thought to involve generation of eicosanoids because it can be blocked by NSAIDs, but for other substances (*e.g.,* polyvinylchloride, polyethylene glycol), the mechanism remains unknown. Nonallergic contact urticaria may occur on first exposure and may depend on dose and concentration of the chemical.

DIFFERENTIAL DIAGNOSIS OF URTICARIA

Urticarial Vasculitis

Urticarial vasculitis is characterized clinically by CU and histologically by leukocytoclastic vasculitis on lesional skin biopsy.[17] It is important to differentiate urticarial vasculitis from CSU in terms of prognosis, approach to diagnostic evaluation, and therapy.

Angioedema Without Wheals

Angioedema occurs in nearly half the patients with CSU; in these patients, the disease tends to be more severe and more difficult to treat. Angioedema without wheals occurs in about 10% of patients with CSU but needs specific investigation because it may be caused by C1 inhibitor deficiency or drugs, although many cases remain unexplained (Fig. 42.7).

Angioedema Caused by C1 Inhibitor Deficiency

The diagnosis and clinical presentation of hereditary and acquired C1-INH deficiency are covered in Chapter 21.

Angioedema With Normal C1 Inhibitor

Histaminergic angioedema (without wheals)

Idiopathic. Idiopathic angioedema is diagnosed when no cause can be established. It is generally self-limiting but may follow a prolonged course. Treatment of idiopathic angioedema is as for spontaneous urticaria. Antihistamines are the mainstay of therapy; corticosteroids (prednisolone 30–40 mg/day) are useful for up to 3 days to cover more severe episodes, and epinephrine is lifesaving in laryngeal angioedema. Tranexamic acid can help in idiopathic angioedema without wheals, but it is usually ineffective when angioedema occurs with urticarial wheals.

Drug-induced. Angioedema without wheals occasionally occurs with NSAIDs, usually within several hours of intake.

Bradykininergic angioedema

Hereditary. Some kindreds have been described with a dominant pattern of angioedema inheritance, mainly in women, in whom a gain-of-function (GOF) mutation of the contact pathway of coagulation in *F12* has been found in about 20% of affected families. This used to be called type III HAE, but the term *hereditary angioedema with normal C1 inhibitor* is now preferred.[18]

Drug-induced. Angioedema develops in 0.1–2% of patients on ACEIs. Of these, up to 40% of patients present with life-threatening angioedema of the upper airway. It is thought to result from inhibition of kininase II, which breaks down bradykinin, as well as converting angiotensin I to angiotensin II in the renin–aldosterone pathway. It usually presents with episodic and unpredictable swellings of the head and neck, especially of the tongue and oropharynx. Although, in most cases, angioedema develops within the first week of treatment with ACEIs, symptom onset may occur after several years on treatment. Management involves discontinuation of ACEI therapy. Angioedema often recurs on reexposure to ACEIs. ACEIs may also precipitate angioedema in patients with angioedema from other causes, including C1-INH deficiency. Rare instances of angioedema have been reported with angiotensin II receptor antagonists.

Autoinflammatory Syndromes Presenting With Urticarial Rash

Acquired

Schnitzler syndrome. Schnitzler syndrome is a rare form of CU with intermittent fever, bone pain, high ESR, and monoclonal IgM, or, very rarely, IgG gammopathy.[19] Clinically, patients present with nonpruritic or mildly pruritic CU, mainly affecting the trunk and limbs. The wheals are resistant to antihistamines, and angioedema is rare. Fever bouts may exceed 40°C, sometimes with chills and nocturnal sweating. Patients often suffer from bone pains, mainly in the pelvis or tibias, arthralgia, and sometimes full-blown arthritis. Lymphadenopathy, hepatomegaly, and splenomegaly may be present.

Kappa light-chain monoclonal IgM and, less commonly, IgG paraproteins are found on serum electrophoresis. The ESR is persistently elevated at 60–100 mm/hour, with leukocytosis, elevated platelet count, and anemia. Skin histology shows neutrophilic urticaria with a tendency to localize around appendages in most cases; monoclonal IgM is deposited in the epidermis around the keratinocytes and along basement membranes on direct immunofluorescence. Bone examination may demonstrate hyperostosis on radiography and hyperfixation on bone technetium scanning. Bone marrow examination shows normal results in most patients, but nonspecific lymphocytic, plasmocytic, or polyclonal infiltration is present in about 20%.

The pathophysiology of Schnitzler syndrome is still unclear, and the severity of urticarial rash does not depend on the paraprotein level. Evidence of activation of interleukin-1 (IL-1), increased IL-6, and granulocyte macrophage–colony-stimulating factor (GM-CSF) and anecdotal reports of complete clinical responses to the IL-1 receptor antagonist anakinra suggest that cytokines play a leading role in its pathogenesis. The prognosis is generally good. However, long-term follow-up is recommended because patients may develop B-cell lymphomas 10–20 years after its onset.

Hereditary (Cryopyrin-Associated) Periodic Syndromes

Several hereditary autoinflammatory urticarial syndromes show mutations of the *NLRP-3* gene on chromosome 1q44. *NLRP-3* encodes a protein called cryopyrin, which is involved in apoptosis and inflammation. These rare autosomal dominant disorders include familial cold autoinflammatory syndrome (FCAS), Muckle-Wells syndrome (MWS), and chronic infantile neurological, cutaneous and articular (CINCA) syndrome, now grouped under the inclusive term *cryopyrin-associated autoinflammatory syndrome* (CAPS).[20]

DIFFERENTIAL DIAGNOSIS

Several dermatoses may present with urticarial lesions, including erythema multiforme minor, bullous pemphigoid, and dermatitis herpetiformis. These dermatoses can nearly always be distinguished clinically from urticaria on the basis of their polymorphic pattern, prolonged duration of individual lesions, lack of daily fluctuation, development of vesicles or blisters, and resistance to conventional therapy for urticaria. Very occasionally, skin biopsy, with or without indirect immunofluorescence, may be required to make the distinction.

Papular urticaria is an urticarial reaction to insect bites in sensitized individuals. The lesions are fixed rather than fluctuating, may take days or weeks to resolve fully, and may leave pigmentation or scars. Bites often occur asymmetrically in groups or lines. Although histamine is involved in the initial pruritic lesions, oral antihistamines are usually unhelpful, and potent topical steroids may be required to speed up natural resolution.

WORKUP IN PATIENTS WITH URTICARIA

Evaluation of patients with urticaria requires a detailed history and a physical examination.[21] The history is particularly important in patients with urticaria and should include a thorough inquiry for all potential causes of the disorders, possible precipitating and aggravating factors, the timing of onset and duration of individual wheals, associated symptoms as well as travel history, recent infection, occupational exposure, food and drug intake, and comorbidity. Patients may be asked to keep a diary of attacks and complete weekly urticarial activity scores for CSU.[1] The duration of individual lesions can be very helpful in distinguishing the different clinical patterns of urticaria.

Physical examination should focus on skin lesion morphology and careful systemic evaluation for underlying disease or comorbidities. If the patient is symptom-free at the time of evaluation, photographs taken when the symptoms occur are helpful. The approximate duration of individual lesions can be assessed by outlining a particular lesion with a pen and observing it for a day. The appearance and distribution of skin lesions may suggest a diagnosis (*e.g.,* the pinpoint lesions with a large flare in cholinergic urticaria or lesions on sun-exposed areas in solar urticaria).

Further evaluation of patients with urticaria is guided by the patient history and clinical pattern of disease. However, it must be remembered that there can be more than one cause for urticaria and that different clinical subtypes of urticaria can coexist in one patient.

CLINICAL PEARLS

Diagnosis of Clinical Patterns of Urticaria

- The duration of individual wheals can help define the pattern of urticaria.
- Wheals lasting ≤1 hour are usually triggered by a physical stimulus.
- Localized wheals lasting up to 2 hours may be caused by skin or mucosal contact with an allergen or a nonimmunological exposure.
- Wheals that take 1–24 hours to fade are usually a presentation of chronic spontaneous urticaria.
- Wheals lasting >24 hours may be caused by delayed pressure urticaria or urticarial vasculitis.

Workup in Acute Urticaria

When allergens are suspected as the cause there should be a close temporal relationship to the time of exposure, usually starting within minutes, a history of previous exposure causing sensitization and prompt resolution on allergy withdrawal. IgE sensitization can be assessed by skin prick testing or fluoroenzymatic immunoassay (FEIA) (also known as serum specific IgE testing).

Workup in Physical Urticarias

When physical urticaria is suspected, appropriate challenge testing should be performed to confirm the diagnosis.[22] Generally, there is no need for further investigation, except for cold urticaria, where blood cryoglobulins should be assayed.

Workup in Chronic and Episodic Spontaneous Urticarias

No laboratory workup is recommended for patients with mild CSU that is easily controlled by antihistamines, unless the history points to an underlying disease. Studies have shown that random laboratory testing very rarely yields evidence of unsuspected internal diseases as a cause of CSU, and this should therefore be discouraged.[21]

Screening laboratory evaluation can be considered in patients with poor response to first-line antihistamine treatment. A complete blood count (CBC) with differential, ESR, thyroid-stimulating hormone (TSH), thyroid antibodies, liver function tests, and urinalysis will exclude most diseases associated with urticaria.

Other evaluations should be guided by abnormal findings in the history and physical examination in patients with CSU. Additional tests may include stool examination for ova and parasites and *Helicobacter pylori*, antinuclear antibody titer, hepatitis viral screening, and skin prick or blood tests for specific IgE in episodic urticaria if there is a story of intermittent allergen exposure. Tests for immediate hypersensitivities should not be undertaken in chronic continuous urticaria unless there are compelling reasons for doing so. Rarely, CSU can be caused by a specific food additive, which should be confirmed by dietary exclusion and double-blind, placebo-controlled oral challenge.

The Diagnosis of Autoimmune Chronic Urticaria

The diagnosis of autoimmune CU is not straightforward and involves *in vivo* and *in vitro* approaches. Autologous serum skin testing (ASST) is a simple and useful screening method for autoreactivity in patients with CU. For skin testing, 0.05 mL of the patient's own serum should be injected intradermally into clinically uninvolved forearm skin, together with an equal volume of saline as a negative control. Histamine 10 mg/mL can be used as a positive control. The reaction is considered positive if the serum skin test forms a pink wheal at least 1.5 mm greater than the negative control at 30 minutes. The test is 80% specific and 70% sensitive for autoimmune CU as defined by a positive basophil histamine release assay.[23]

The current diagnostic gold standard in autoimmune CSU is a functional release assay using whole sera on healthy donor basophils or mast cells, but these only give indirect evidence of functional autoantibodies. Being technically difficult, these assays are mainly confined to research centers. Nonfunctional immunoassays (Western blot, ELISA) based on binding of autoantibodies to relevant antigens (FcεRIα or IgE) do not correlate well with the results of functional assays.[10]

MANAGEMENT OF URTICARIA

Finding effective treatment for urticaria can be challenging. Treatment should be tailored to the clinical pattern, duration, and severity of the urticaria. Management should include nonpharmacological measures and drug therapy with a stepwise approach.[1]

THERAPEUTIC PRINCIPLES

Management of Urticaria

- Eliminate infectious, drug, or food causes.
- Minimize nonspecific aggravators, including heat, stress, alcohol, nonsteroidal antiinflammatory drugs, and pressure.
- Regular oral H1 antihistamines are the first line of therapy for all spontaneous and inducible urticarias.
- Second-line treatments, including short courses of oral corticosteroids, may be necessary for specific clinical situations.
- Immunosuppressive therapies should be reserved for patients with severe autoimmune urticaria or steroid-dependent urticaria that has not responded to other first- and second-line measures.
- Omalizumab (monoclonal anti-IgE) is effective in chronic spontaneous urticaria refractory to H1 antihistamines

General Measures

Causes, triggers, and aggravating factors should be avoided or minimized, whenever possible. Patients with CSU should minimize exposure to nonspecific aggravating factors, such overheating, wearing tight clothes and shoes, stress, alcohol, dietary pseudoallergens, and some drugs. NSAIDs aggravate up to 30% of patients with CSU and are usually avoided. This does not apply to the physical urticarias, in particular DPU, where NSAIDs may be used as treatment. ACEIs are contraindicated in angioedema without wheals that may be mediated by kinins but not in other patterns of urticaria. Although it is often recommended that patients with CSU avoid codeine and penicillin, clinical experience suggests that this is not necessary. Cooling lotions and creams, such as 1% menthol in aqueous cream, may help relieve pruritus. Some patients with spontaneous but not inducible CU appear to respond to a low-pseudoallergen diet.[12] However, controlled clinical trials are lacking.

First-Line Therapy

Antihistamines are the cornerstone of treatment for all types of histaminergic urticaria. Second-generation antihistamines offer several advantages over classic H1 antihistamines, including lack of sedation and impairment of performance, longer duration of action, and absence of anticholinergic side effects. Meta-analysis indicates that antihistamines are effective in 40–90% of patients with CU. Second-generation antihistamines are inverse agonists of H1 receptors, which stabilize H1 receptors in the inactive conformation and are, therefore, most effective in CU when taken regularly for prophylaxis. The timing of antihistamine intake should be adjusted to suit the diurnal pattern of urticaria for each individual. Although the evidence base for combining H1 and H2 antihistamines is poor, this may be helpful. H2 antihistamines also suppress the dyspepsia that often accompanies severe urticaria.

Second-Line Therapy

It has become common practice to increase second-generation antihistamines above their licensed doses up to fourfold when CU does not respond because clinical experience and updosing studies show that this achieves better control in some patients. When urticaria does not respond to first-line measures, systemic corticosteroids are commonly used as short-term rescue therapy for acute urticaria or severe exacerbations of CU. Long-term treatment with oral corticosteroids is not recommended because of safety concerns. In corticosteroid-dependent patients, an alternate-day dosing schedule may be used, and steroid-sparing drugs should also be considered. A wide variety of medications have been reported to be of benefit for antihistamine-unresponsive urticaria when oral corticosteroids might otherwise have to be considered, but they are invariably used off-label, and controlled trials are usually lacking. Where possible, these should be targeted at specific subgroups of urticaria patients (*e.g.,* the LT receptor antagonist montelukast has been shown to be effective in aspirin-sensitive patients with CSU, and there is limited evidence for a trial of thyroxine in biochemically euthyroid patients with CSU who have thyroid autoimmunity).

Third-Line Therapy

The recent approval of omalizumab (anti-IgE) as add-on therapy for CSU not responding to antihistamines has transformed the management of severe disease in many patients. The recommended dose is 300 mg by monthly subcutaneous injection, although clinical experience indicates that lower (and occasionally higher) doses may be effective and the interval between treatments can be extended in good responders. Clinical experience suggests that when used off-label, it may also be effective for many patients with inducible urticarias. The mechanism of action is still not certain but probably involves reduction of high-affinity IgE receptors on mast cells and basophils as a consequence of a rapid fall in circulating free IgE and a later reduction in IgE bound to mast cells, resulting in a stabilizing effect. Experience to date suggests that omalizumab is often highly effective for control of urticaria symptoms (including itch) but probably does not modify the natural history of the illness.

Immunosuppressive therapy is mainly considered for autoimmune CSU. However, it may also benefit therapy-resistant CSU without evidence of circulating antibodies.

The best-studied immunosuppressive therapy is cyclosporine, shown to be effective initially at 4 mg/kg daily in patients with CSU. Patients must be monitored carefully for renal impairment and hypertension; treatment should normally be limited to 4 months. Cyclosporine is contraindicated in patients with previous malignant disease except nonmelanoma skin cancer. There is some evidence for efficacy of plasmapheresis and immunoglobulins in chronic autoimmune urticaria, although these are expensive options, and controlled clinical trials are needed. Methotrexate, mycophenolate mofetil, and azathioprine have also been used alone or with corticosteroids.[14]

MANAGEMENT OF HEREDITARY ANGIOEDEMA

The swellings of HAE are mediated by kinins rather than histamine, so the management of HAE is completely different from mast cell–dependent urticaria. The primary aim of therapy is to replace the missing functional C1 esterase inhibitor or stabilize the coagulation, fibrinolysis, complement, and kallikrein–kinin pathways. ACEIs should be avoided because ACE is a key enzyme involved in the breakdown of bradykinin. Angiotensin II receptor agonist (ARA 2) are probably safe alternatives. Exogenous estrogen (oral contraceptives and hormone replacement therapies) should be avoided in women, since it appears to activate kallikrein through activated factor XII. Lifestyle events that exacerbate HAE vary in their importance among individuals but may include local trauma (*e.g.,* dental extraction), stress, tiredness, and intercurrent infections, possibly including *Helicobacter pylori* infection. Attacks involving the extremities or abdomen are the most common. It is estimated that over their lives, up to 50% of patients will experience an attack of oropharyngeal swelling, with risk of asphyxiation, so treatment must be rapidly available for emergencies in all patients as well as as a preventive measure. The intention of treatment is to reduce or curtail the severity of a swelling and to reduce "downtime" if one occurs. Some patients with frequent peripheral swellings prefer to treat themselves symptomatically with analgesics and wait for spontaneous resolution over 3–4 days, rather than have specific treatment, since no treatment is completely without risk of adverse effects, and there are potential problems with cost and/or availability. The management of types I and II HAE is better defined than type III but follows similar principles.[24]

Treatment of the Acute Attack

The "gold standard" treatment for oropharyngeal or GI attacks in most countries (except the United States until recently) has been plasma-derived nanofiltered C1 esterase inhibitor (pd-C1-INH) concentrate given by intravenous infusion. It has proven to be effective and safe. The recommended dose is 20 units/kg body weight in adults and children, although smaller doses have been used in the past and may be effective. Up to 3 units of fresh frozen plasma (containing C1 inhibitor and its substrate, complement) may be given as an alternative in an emergency when pd-C1-INH is not available, but it will not have had the same stringent screening and production measures to eliminate the possibility of transmissible infection. Initial improvement in the swelling may be seen within 30–60 minutes, and time to clearance is usually in the order of 24 hours after pd-C1-INH. Self-administration can reduce the severity of attacks by allowing earliest treatment and should be encouraged. A recombinant C1-INH is licensed in Europe. Icatibant (a bradykinin 2 receptor antagonist) is also highly effective for acute episodes of HAE. Ecallantide (a kallikrein antagonist) is currently licensed in the United States but not in Europe. Both are given by subcutaneous injection, which provides an easier route for administration than intravenous infusion. Ecallantide should be given under medical supervision, since there is a small risk of allergic reactions. Icatibant is licensed for self-administration, which provides a potential advantage. There have been no head-to-head studies of pd-C1-INH, ecallantide, and icatibant, but experience suggests that the effectiveness of these products is comparable. It should be noted that antihistamines and steroids have no place in the management of HAE. Any improvement resulting from epinephrine's effects on vasopermeability will be transient, and there is no evidence that it changes the overall course of an attack.

Short-Term Prophylaxis

It has become common practice to give prophylactic treatment with pd-C1-INH before procedures involving local trauma to the oropharynx, including dental extraction and intubation for general anesthesia. The dose for prophylaxis may be less than that needed for treatment of an acute attack. Common practice is to give at least 500 units pd-C1-INH 1 hour before the procedure, with a further dose available in case a swelling develops. Other strategies include increasing the dose of prophylactic treatment with anabolic steroids (danazol or stanozolol) or plasmin inhibitors (tranexamic acid or ε-aminocaproic acid) for at least 48 hours before and after the procedure. Specific guidance on dosing for adults and children can be found elsewhere.

Long-Term Prophylaxis

Anabolic steroids are the mainstay of long-term prophylaxis in countries where they are licensed for use in the treatment of HAE. They increase the production of C1-INH by the liver in heterozygotes with a functioning allele. The dose should be titrated against the clinical response rather than blood levels of C1-INH to the lowest level that prevents or ameliorates the condition. Virilizing side effects can be problematic for women, and anabolic steroids are avoided in children because of concerns about growth retardation. Monitoring of liver function and lipid profiles should be undertaken periodically. Performing a liver ultrasound examination every 3 years to screen for development of hepatoma is usually recommended in patients on long-term prophylaxis. Plasmin inhibitors are generally less effective for prophylaxis compared with anabolic steroids but are preferred in children. pd-C1-INH infusions twice a week may be given as prophylaxis during pregnancy and in rare situations when alternative therapies are not appropriate.

Anaphylaxis

Anaphylaxis is a severe, life-threatening, systemic reaction of sudden onset and involves respiratory compromise, cardiovascular collapse, or both. Its clinical features and management have been summarized in international guidelines.[25-28]

Epidemiology of Anaphylaxis

Estimates of the incidence of anaphylaxis in the general population in Europe and the United States vary from 50 to 103 cases per 100 000 per annum.[29] Lifetime personal risk estimate for anaphylaxis in the general population is believed to be 0.05–2%.[30] The mortality rate in anaphylaxis is estimated to be below 0.001%.[31] In the United Kingdom, anaphylaxis accounts for 20 deaths per year, which represents 1 death in 3 million people.[32] According to the European Anaphylaxis Registry, a life-threatening or fatal anaphylaxis occurs in about 1 in 100 children experiencing severe anaphylaxis.[33]

There was a sevenfold increase in anaphylaxis admissions in England and Wales in 1992–2002.[34] Severe anaphylaxis was diagnosed in 1–9 per 10 000 people attending emergency departments in the United Kingdom, Australia, and the United States.[35] It is estimated that in a 12-month period, 1 in 12 patients with previous anaphylaxis will have a recurrence, and 1 in 50 will require hospitalization or treatment with epinephrine.[36]

Pathophysiology of Anaphylaxis

Although anaphylaxis is often subdivided into immunological and nonimmunological types, the clinical presentation is similar in both, and most authorities no longer make a distinction. Immunological anaphylaxis is further classified as IgE-mediated and non-IgE-mediated anaphylaxis. In IgE-mediated anaphylaxis, allergen cross-links allergen-specific IgE on the surface of mast cells and basophils, leading to their degranulation. Release of mediators causes bronchoconstriction, mucus secretion, diminished cardiac contractility, increased vascular permeability, vasoconstriction of coronary and peripheral arteries, and vasodilation of venules, thereby producing clinical symptoms of anaphylaxis. IgE-mediated reactions occur in presensitized patients (*e.g.,* those experiencing penicillin-, insulin-, latex-, or peanut-induced anaphylaxis). IgG- or IgM-related transfusion reactions should be classified as immunological, non–IgE-mediated anaphylaxis. In contrast, some substances, such as opioids, radiocontrast media, vancomycin, and some muscle relaxants, are capable of direct release of mediator (histamine) from basophils and mast cells without involvement of IgE. Although reactions to NSAIDs are considered pharmacological rather than immunological (because of the downstream effects of COX inhibition), an IgE-mediated mechanism has been suspected in a few patients but is difficult to prove. Apart from IgE, other antibodies may be involved: In murine models IgG-mediated FcγRIII-dependent anaphylaxis elicited by a high dose of allergen has been described. The key participating cells in this type of anaphylaxis are macrophages, with platelet-activating factor being the main mediator.[37]

Etiology of Anaphylaxis

Anaphylaxis is most commonly caused by foods, drugs, general anesthetic agents, insect stings, and latex. Rare causes include vaccines, semen, and aeroallergen inhalation. Exercise can occasionally cause anaphylaxis either on its own (exercise-induced anaphylaxis) or after ingestion of a food to which the individual is presensitized (food- and exercise-induced anaphylaxis). Up to 20% of patients with systemic mastocytosis present with anaphylaxis during their lifetime.[38] Anaphylaxis mostly occurs in reaction to hymenoptera stings, NSAIDs, and opioids and in the perioperative setting.[39] Idiopathic anaphylaxis accounts for up to 60% of anaphylaxis cases in ambulatory adults and for 10% of cases in children.[40] It is increasingly recognized that some anaphylactic cases are multifactorial. Cofactors are thought to lower the threshold for the induction of anaphylaxis and are implicated in about 30% of anaphylaxis cases in adults.[41] The risk of severe anaphylaxis is increased in patients taking beta-blockers or ACEIs, or both.[28]

The most common routes of allergen exposure are oral and parenteral, although inhalation of allergens (*e.g.,* fish or legume allergens after cooking, latex particles in health care settings) or percutaneous penetration after skin contact can induce anaphylaxis in highly sensitized patients.

Food-Induced Anaphylaxis

According to a meta-analysis of 34 studies, food-induced anaphylaxis was reported to occur with an incidence of 0.14 cases per 100 000 person-years at all ages and up to 7 per 100 person-years in children aged 0–4 years.[42] According to the European Anaphylaxis Registry, foods are a predominant cause of anaphylaxis in the first decade of life. Teenagers and young adults in their second and third decades of life are known to be at the highest risk of fatal food-induced anaphylaxis.[34]

Peanuts, tree nuts, fish, and shellfish are the most frequent culprits in food-induced anaphylaxis, but almost any food can be implicated. Many cases of severe anaphylaxis are caused by unintended exposure to hidden food allergens. In addition, cofactors, such as alcohol, NSAIDs, and exercise, may increase the severity of a food-induced allergic reaction.[27]

Drug-Induced Anaphylaxis

Drug-induced anaphylaxis is more common in hospitalized patients than in the community. Any drug can cause anaphylaxis, but the prevalence varies: one study found 5–15 cases per 100 000 hospitalized patients for most analgesics and antibiotics, whereas for dextran, penicillin, pentoxifylline, and streptokinase, the figure was over 30 cases per 100 000.[43] Anaphylaxis has been reported after treatment with monoclonal antibodies (mAbs; basiliximab, rituximab, infliximab, omalizumab, etc.).[44] IgE antibodies to galactose-α-1,3-galactose (α-gal) are thought to cause immediate anaphylaxis on first exposure to cetuximab.[45]

Drugs are known to be the leading cause of fatal anaphylaxis, with the mortality rate in drug-induced anaphylaxis about one death per 50 000–100 000 treatment courses.[46]

All routes of administration can potentially be fatal, including oral, intravenous, subcutaneous, intraarticular, intrauterine, inhalational, rectal, or topical, but the risk is greatest after parenteral administration.

Perioperative Anaphylaxis

Perioperative anaphylaxis occurs in up to 1 of 13 000 anesthesia administrations.[28] The common causes of perioperative anaphylaxis include neuromuscular-blocking agents, antibiotics, blood and blood products, dyes, chlorhexidine, or natural rubber latex.[28] Reactions to neuromuscular-blocking agents mostly occur on first exposure and have been associated with a 70% rate of cross-reactivity in this group; risk factors for fatal anaphylaxis include male gender, history of cardiovascular disease, emergency setting, and use of beta-blockers.[47] Latex remains an important cause of intraoperative anaphylactic reactions. The antiseptic chlorhexidine is increasingly recognized as a cause of IgE-mediated perioperative anaphylaxis.[48]

Insect Sting–Induced Anaphylaxis

Following insect stings, life-threatening systemic reactions have been reported in 0.4–0.8% of children and up to 3% of adults.[49] Severe anaphylaxis from insect stings causes approximately 40 deaths annually in the United States and nearly 100 in Europe.[49] Occupational anaphylaxis caused by venom allergy can occur in bee keepers, gardeners, forestry or greenhouse workers, farmers, truck drivers, and masons.[50] The severity of reaction depends on the type of insect, amount of venom, location of sting, the patient's sensitivity, older age, preexisting diseases, previous less severe systemic reactions, concomitant treatment, mast-cell disorder, and elevated baseline tryptase.[51] Allergen-specific immunotherapy with venom extracts has been shown to be safe and effective in patients with *Hymenoptera* venom allergy, providing some clinical protection within the first 8 weeks of treatment and a long-lasting effect after 3–5 years of maintenance treatment (Chapter 43). It is noteworthy that patients with systemic mastocytosis are at risk of potentially fatal anaphylaxis to insect stings even if they are not presensitized to venom: this may be attributed to venom components, such as phospholipase A2, acting as mast-cell liberators.[51]

Latex-Induced Anaphylaxis

The prevalence of latex allergy has been estimated to be as high as 1–6% in the general population, 8–17% in health care workers, and 67% in patients with spina bifida.[31] Latex-induced anaphylaxis has been reported in surgery and dentistry and can be fatal. More than half the patients with latex allergy report allergic reactions to fruits, such as banana, avocado, kiwi fruit, chestnut, pear, pineapple, grape, and papaya.

Other Rare Causes of Anaphylaxis

Anaphylaxis occurs during 1 in 20 000–47 000 transfusions of blood or blood products, especially in patients with IgA deficiency.[52] IgA deficiency affects 1 in 500–700 Caucasians. One-third of these patients have circulating anti-IgA antibodies, which are associated with serious life-threatening anaphylactic reactions to blood products containing IgA.[52] Seminal fluid allergy is extremely rare, mostly affecting young women with atopy, with 20% of cases developing life-threatening anaphylaxis. These reactions can be prevented with condom use or intravaginal desensitization with seminal fluid. In cystic echinococcosis, anaphylaxis can follow the rupture of a hydatid cyst during surgery or as a result of trauma.[53]

Anaphylaxis in Clonal Mast-Cell Disorders

There is a link between unexplained anaphylaxis and clonal mast-cell disease (systemic mastocytosis or monoclonal mast cell activation syndrome).[54] In systemic mastocytosis, anaphylaxis may occur to *Hymenoptera* stings and drugs (NSAIDs, opioids, and drugs used in the perioperative setting). Patients with unexplained anaphylaxis should be evaluated for mast-cell clonality to exclude systemic mastocytosis or mast-cell activation syndrome.[54]

Clinical Diversity of Anaphylaxis

In anaphylaxis, there is a remarkable range of clinical symptoms. Anaphylaxis can be preceded by prodromal symptoms, such as tingling and redness of the palms and soles, anxiety, sense of impending doom, and disorientation. Anaphylaxis most commonly begins in skin and mucous membranes, is followed by involvement of the respiratory and GI tracts and the cardiovascular system, and may finally proceed to cardiac and/or respiratory arrest. Generalized urticaria and angioedema are the most common manifestations of anaphylaxis, observed in over 90% of cases, but may be absent. Respiratory symptoms may vary from rhinitis to laryngeal edema and airway obstruction, which are potentially life threatening. Cardiovascular manifestations in anaphylaxis include hypotension and/or cardiac arrhythmias. In adults, reduced blood pressure is regarded as systolic blood pressure of <90 mm Hg or >30% decrease from that person's baseline values.[27] Some patients present with only cardiovascular collapse in the absence of other signs of anaphylaxis, especially during general anesthesia. Anaphylaxis is usually associated with tachycardia caused by increased cardiac sympathetic drive in response to a decreased effective vascular volume, but bradycardia may also occur. Anaphylaxis may result in up to 35% of intravascular fluid leaking into the extracellular space. A two-phase reaction to the hypovolemia may present with tachycardia as the first phase, followed by bradycardia when effective blood volume falls by 20–30%.[55]

CLINICAL PEARLS

Diagnosis of Anaphylaxis

- Anaphylaxis is characterized by extreme difficulty with breathing as a result of airway obstruction from angioedema/bronchoconstriction, circulatory collapse, or both.
- It is nearly always accompanied by tachycardia, usually by flushing, urticaria, and panic, and sometimes by vomiting and diarrhea.
- Panic attacks do not involve airways obstruction, hypotension, or urticaria but may be accompanied by faintness or tetany of the hands as a result of hyperventilation.
- Vasovagal attacks present with fainting, nausea, slow pulse, and pallor without respiratory difficulty, diarrhea, or urticaria.

Four clinical patterns of anaphylaxis have been described: immediate, biphasic, protracted, and delayed. Anaphylaxis can occur within seconds after allergen exposure: The more rapid the onset of anaphylaxis after allergen exposure, the more severe and life-threatening is the reaction. Food-induced anaphylaxis takes slightly longer to develop than drug- or insect-induced anaphylaxis. Some patients may have biphasic anaphylaxis, with recurrence after 1–72 hours, but usually within 8–10 hours of the initial attack.[28] According to a meta-analysis of 27 studies, patients with anaphylaxis presenting with hypotension or having an unknown trigger may be at increased risk of biphasic anaphylaxis.[56] Some patients may develop protracted anaphylaxis, which sometimes lasts longer than 24 hours, may be extremely severe, and is often resistant to treatment. Delayed onset of anaphylaxis has been reported anecdotally but is very rarely encountered in clinical practice.

The diversity and severity of symptoms in anaphylaxis depend on the dose of allergen, the route of allergen exposure, the extent of allergen absorption, the degree of sensitization, individual allergen threshold for a reaction, target tissue sensitivity, cofactor involvement, coexisting atopic diseases and their severity, and concomitant treatment. “Summation anaphylaxis” has been recognized as occurring after simultaneous exposure to various stimuli (physical exercise, infection, stress, or concomitant exposure to other allergens or treatment with NSAIDs, ACEIs, or beta-blockers).

Fatal reactions to foods are usually characterized by respiratory symptoms (bronchospasm and hypoxia). In contrast, anaphylaxis induced by insect stings is more likely to lead to cardiovascular collapse. Asthma sufferers are at higher risk of fatal anaphylaxis.[27] The risk of relapse in anaphylaxis depends on the type of allergen, individual allergen threshold, success of allergen avoidance, and the availability of immunotherapy.

The most dangerous symptoms are laryngeal edema, respiratory failure, and circulatory collapse, which may lead to death. Deaths from acute asthma in anaphylaxis occur predominantly in patients with preexisting unstable asthma. Rapid fatal shock often occurs without other symptoms, and laryngeal angioedema is the least common cause of fatality. Fatal anaphylaxis occurs mostly within an hour of the onset of anaphylaxis.[32] According to the UK Fatal Anaphylaxis Registry, the earliest arrest in fatal food anaphylaxis occurred within 25–35 minutes of exposure, slightly slower than with insect stings (10–15 minutes) or drugs (≤5 minutes in hospital and 10–20 minutes outside hospital).[32]

Diagnosis of Anaphylaxis

Measurement of blood tryptase is now widely used as a marker of mast cell degranulation for *in vitro* confirmation of anaphylaxis. Beta-tryptase is released from mast cells, but not from basophils, and diffuses more slowly compared with histamine. The concentration of tryptase peaks 1–2 hours after the onset of reaction, with a half-life of 1.5–2.5 hours. Samples for tryptase testing should be collected as soon as possible after emergency treatment of the patients and within 1–2 hours (but not later than 4 hours) of anaphylaxis onset, and again after 24 hours (baseline sample) to check that the value has returned to normal.[57] Tryptase may also be detected in postmortem specimens after death from suspected anaphylaxis.[58]

Normally, mature tryptase is below detection limits in the serum of healthy subjects, whereas it is elevated in most cases of anaphylaxis with vascular compromise, especially if it is parenterally induced. However, a normal tryptase result does not exclude anaphylaxis. Tryptase within the normal range during anaphylaxis is often observed in food-induced anaphylaxis. In infants, tryptase may not be elevated after anaphylaxis, although the baseline levels may be increased.[59] The diagnostic value of other mast-cell products, such as histamine, platelet-activating factor, and proteases (carboxypeptidase A3), in anaphylaxis is under investigation.

In anaphylaxis, component-resolved diagnosis can be useful to stratify risk in certain clinical scenarios.[60] For example, patients with wheat-dependent exercise-induced anaphylaxis should be tested for omega-5-gliadin sensitization,[61] whereas patients with anaphylaxis to vegetables, fruits, nuts, and cereals may have IgE to nonspecific lipid transfer proteins (mostly Pru p 3 and Tri a 14).[62] In delayed anaphylaxis to mammalian meat or anaphylaxis to cetuximab, tests for IgE against galactose-α-1,3-galactose (α-gal) should be considered.[45] Certain components are regarded as risk factors for anaphylaxis; for example, Mal d 3 sensitivity carries a sevenfold greater risk of anaphylaxis to apple compared with Bet v 1.[63]

Management of Anaphylaxis

Early recognition of anaphylaxis facilitates removal of the cause and prompt institution of treatment. The patient with anaphylaxis should lie down with the legs elevated to increase venous blood return and maintain cardiac output, and intravenous fluids should be given. In drug-induced or insect-induced anaphylaxis a tourniquet may be placed proximal to the site of the injection or insect sting to slow absorption of injected antigens. The tourniquet should be released for 3 minutes at 5-minute intervals, with the total duration of application not exceeding 30 minutes. Epinephrine should be administered by an intramuscular injection in the mid-outer thigh at the first sign of respiratory failure or cardiovascular collapse and repeated after 5–15 minutes if the response to the first injection is suboptimal.[25] Epinephrine autoinjectors for self-administration are available, but a single pen may be insufficient to reverse severe reactions. Use of these pens in anaphylaxis outside hospital can be lifesaving. Overall, prompt diagnosis of anaphylaxis, early administration of epinephrine, and fast transport to emergency rooms are crucial factors for successful management of anaphylaxis.

Epinephrine is both an α and β adrenergic agonist with cyclic adenosine monophosphate (cAMP)–mediated pharmacological effects on target organs. In patients with anaphylaxis, stimulation of α_1 adrenergic receptors increases peripheral vascular resistance, thereby improving blood pressure and coronary perfusion, reversing peripheral vasodilation, and decreasing angioedema. Activation of β_1 adrenergic receptors increases myocardial contractility (inotropy, chronotropy) while stimulation of β_2 adrenoreceptors causes bronchodilation and decreases the release of inflammatory mediators from mast cells and basophils.[27]

Current guidelines recommend the intramuscular route for epinephrine administration because of faster absorption and higher plasma level of epinephrine after intramuscular injection compared with subcutaneous injection.[25] The appropriate dosage of epinephrine is 0.01 mg/kg of a 1:1000 (1 mg/mL) solution, to a maximum of 0.5 mg in adults.[27] Epinephrine has a rapid but short action, so the dose may need to be repeated every 5–15 minutes until symptoms improve.[64] More than one dose is required in a third of patients. Intravenous administration of epinephrine should be reserved for severe anaphylaxis with profound life-threatening hypotension that is refractory to other treatment because of a risk of potentially fatal cardiac arrhythmias and myocardial infarction.[26] The recommended dose is a 1 : 100 000 solution (*i.e.,* 1 mg in 100 mL saline at initial infusion rate of 2–10 μg/min, titrated up or down, depending on the clinical response or epinephrine side effects).

Common pharmacological adverse effects of epinephrine include anxiety, fear, headache, pallor, tremor, dizziness, and palpitation. In the event of an overdose, unwanted effects may include increased Q–T interval on electrocardiography, ventricular arrhythmias, angina, myocardial infarction, increased blood pressure, pulmonary edema, and intracranial hemorrhage. Patients with cardiovascular diseases and thyrotoxicosis and cocaine users are particularly susceptible to the adverse effects of epinephrine.

The efficacy of epinephrine can be decreased by concomitant therapy with beta-blockers, which is associated with unopposed stimulation of α adrenoreceptors and reflex vagotonic effects, leading to bradycardia, hypertension, coronary artery constriction, bronchoconstriction, and augmented mediator release. Anaphylaxis in patients on beta-blockers can be severe, protracted, and unresponsive to treatment. Patients treated with beta-blockers may require fluid replacement and treatment with glucagon, which increases intracellular cAMP independently of β adrenergic receptors. Glucagon can be administered intravenously 1–5 mg (20–30 mg/kg in children, maximum 1 mg) over 5 minutes, followed by infusion at 5–15 mg/min titrated to clinical response.[26] Glucagon may improve hypotension in 1–5 minutes with maximal effect at 5–15 minutes. Side effects of glucagon include nausea and vomiting.

Corticosteroids are often administered in anaphylaxis to minimize the risk of recurrent or protracted anaphylaxis. The beneficial effects of corticosteroids develop 6–12 hours after administration. Therefore their main role in anaphylaxis is likely to be the prevention of relapse, but it is still unclear how they work.

If there is no response to epinephrine, life support measures should be instituted. The treatment choice depends on the clinical presentation. In hypotension, large volumes of fluids should be given rapidly using 1–2 L of 0.9% normal saline, infused rapidly (5–10 mL/kg within the first 5 minutes for an adult and up to 30 mL/kg in the first hour for children), to compensate for peripheral vasodilatation and for fluid loss into the extravascular space.[26] According to a Cochrane review, crystalloid normal saline should be preferred to colloids for intravenous fluid resuscitation.[65] Other vasopressors (dopamine, glucagon) may be needed to reverse severe hypotension. Oxygen should also be administered in circulatory or respiratory failure. Bronchospasm should be treated with nebulized or inhaled β_2 agonists. If there is severe laryngeal edema, endotracheal intubation and even emergency tracheostomy may be needed to maintain the airway. Methylene blue, a selective nitric oxide cyclic guanosine monophosphate (cGMP) inhibitor, can prevent vasodilatation and has been reported to be effective in refractory anaphylaxis.[66] Patients presented with respiratory compromise should be observed in a hospital setting for at least 6–8 hours.[27] In severe cases of hypotension, an observation period of 12–24 hours is advisable.[27]

Prevention of Anaphylaxis

The first step in prevention is to identify those at risk of anaphylaxis. Therefore all patients with a history of anaphylaxis should be referred for assessment and undergo allergy evaluation. Patients should be instructed how to avoid culprit allergens and cross-reactive agents and should be advised on safe alternatives. The education of patients, their families, and, in the case of children, caregivers and school staff about anaphylaxis, and availability of first-aid measures is of primary importance. Written personalized emergency action plans should be provided to patients at special risk, such as school children. Emergency medications, such as epinephrine autoinjectors, should be dispensed, and patients should receive training on their correct use. Patients should be advised to carry an epinephrine autoinjector with them at all times. Immunotherapy is very effective for prophylaxis of bee- and wasp venom–induced anaphylaxis in sensitized patients and can be lifesaving. Drug-induced anaphylaxis can be prevented by avoidance of culprit drugs and cross-reacting agents, by premedication (for radiocontrast media), and, in some cases, by drug desensitization for antibiotics, chemotherapeutic agents, insulin, vaccines, biological agents, and so on. For food-induced anaphylaxis, avoidance of the culprit food is essential; oral immunotherapy is available in some allergy centers. In idiopathic anaphylaxis, patients with frequent episodes (>6 episodes per year or ≥2 episodes within 2 months) can be treated with steroids to prevent further episodes. Omalizumab[67] and rituximab[68] have been reported to be effective in preventing idiopathic anaphylaxis. Prevention strategies for anaphylaxis should also involve promoting public awareness and public health measures, such as appropriate food labeling, disclosure of food ingredients in restaurants, withdrawal of peanuts from in-flight refreshments, first-aid training in anaphylaxis for school staff, and establishment of national anaphylaxis registries.

TRANSLATIONAL RESEARCH OPPORTUNITIES

ON THE HORIZON

- Defining the role of functional autoantibodies and understanding the mast-cell activation signals in chronic spontaneous urticaria should improve clinical assessment and management of patient subgroups.
- Development of new bradykinin and kallikrein inhibitors for patients with hereditary angioedema should further improve the acute management of this rare but very important condition.
- Understanding the full clinical spectrum of patients with cryopyrin-associated periodic syndrome with *NLRP-3* mutations should allow earlier identification and treatment of these individuals with interleukin-1 blockers to improve quality of life and prevent later complications.

Ongoing research into the importance of functional autoantibodies in chronic urticaria and the cause or causes of mast-cell activation in patients with “idiopathic” disease should help to refine clinical assessment and management pathways.

The effectiveness of a bradykinin receptor antagonist and a kallikrein inhibitor for emergency treatment of HAE has illuminated the key role of the kallikrein–kininogen–kinin pathway in disease pathogenesis. Development of new inhibitors of this pathway may offer additional benefits to patients in the future.

Clarification of the phenotypic spectrum of patients with cryopyrin-associated periodic syndrome with *NLRP-3* mutations will facilitate early detection of affected individuals who present in childhood with persistent urticaria and would benefit from treatment with IL-1 blockers. Early treatment should improve quality of life and may prevent the development of systemic amyloidosis, nephropathy, and deafness in adulthood.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Zuberbier T, Aberer W, Asero R, et al. The EAACI/GA(2)LEN/EDF/WAO Guideline for the definition, classification, diagnosis, and management of urticaria: the 2013 revision and update. Allergy 2014;69:868–87.
2. Greaves MW, Kaplan A, editors. Urticaria and angioedema. New York: Marcel Dekker; 2004.
3. Beltrani VS. An overview of chronic urticaria. Clin Rev Allergy Immunol 2002;23:147–69.
4. Asero R. Chronic idiopathic urticaria: a family study. Ann Allergy Asthma Immunol 2002;89:195–6.
5. O'Donnell BF, O'Neill CM, Francis DM, et al. Human leucocyte antigen class II associations in chronic idiopathic urticaria. Br J Dermatol 1999;140:853–8.
6. Kim SH, Choi JH, Lee KW, et al. The human leucocyte antigen-DRB 1*1302-DQB1*0609-DPB1*0201 haplotype may be a strong genetic marker for aspirin-induced urticaria. Clin Exp Allergy 2005;35: 339–44.
7. Maurer M, Metz M. The status quo and quo vadis of mast cells. Exp Dermatol 2005;14:923–9.
8. Church MK, el-Lati S, Caulfield JP. Neuropeptide-induced secretion from human skin mast cells. Int Arch Allergy Appl Immunol 1991;94:310–18.
9. Sabroe RA, Greaves MW. Chronic idiopathic urticaria with functional autoantibodies: 12 years on. Br J Dermatol 2006;154:813–19.
10. Kaplan AP. Chronic urticaria: pathogenesis and treatment. J Allergy Clin Immunol 2004;114:465–74.
11. Puccetti A, Bason C, Simeoni S, et al. In chronic idiopathic urticaria autoantibodies against FcεRII/CD23 induce histamine release via eosinophil activation. Clin Exp Allergy 2005;35:1599–607.
12. Zuberbier T. The role of allergens and pseudoallergens in urticaria. J Investig Dermatol Symp Proc 2001;6:132–4.
13. Grattan CE. Aspirin sensitivity and urticaria. Clin Exp Dermatol 2003;28:123–7.
14. Kozel MM, Sabroe RA. Chronic urticaria: aetiology, management and current and future treatment options. Drugs 2004;64:2515–36.
15. Buhner S, Reese I, Kuehl F, et al. Pseudoallergic reactions in chronic urticaria are associated with altered gastroduodenal permeability. Allergy 2004;59:1118–23.
16. Wanderer AA, Hoffman HM. The spectrum of acquired and familial cold-induced urticaria/urticaria-like syndromes. Immunol Allergy Clin North Am 2004;24:259–86.
17. Davis MD, Brewer JD. Urticarial vasculitis and hypocomplementemic urticarial vasculitis syndrome. Immunol Allergy Clin North Am 2004;24:183–213.
18. Cicardi M, Aberer W, Banerji A, et al. Classification, diagnosis, and approach to treatment for angioedema: consensus report from the Hereditary Angioedema International Working Group. Allergy 2014;69:602–16.
19. Simon A, Asli B, Braun-Falco M, et al. Schnitzler's syndrome: diagnosis, treatment, and follow-up. Allergy 2013;68:562–8.
20. Shinkai K, McCalmont TH, Leslie KS. Cryopyrin-associated periodic syndromes and autoinflammation. Clin Exp Dermatol 2008;33:1–9.
21. Kozel MM, Bossuyt PM, Mekkes JR, et al. Laboratory tests and identified diagnoses in patients with physical and chronic urticaria and angioedema: a systematic review. J Am Acad Dermatol 2003;48:409–16.
22. Magerl M, Altrichter S, Borzova E, et al. The definition, diagnostic testing and management of chronic inducible urticarias - update and revision of the EAACI/GA(2)LEN/EDF/UNEV 2009 consensus panel recommendations. Allergy 2016;71:780–802.
23. Sabroe RA, Grattan CE, Francis DM, et al. The autologous serum skin test: a screening test for autoantibodies in chronic idiopathic urticaria. Br J Dermatol 1999;140:446–52.
24. Zuraw BL. Hereditary angioedema. N Engl J Med 2008;359:1027–36.
25. Simons FE, Ardusso LR, Bilò MB, et al. International consensus on (ICON) anaphylaxis. World Allergy Organ J 2014;7(1):9.
26. Lieberman P, Nicklas RA, Randolph C, et al. Anaphylaxis – a practice parameter update 2015. Ann Allergy Asthma Immunol 2015;115:341–84.
27. Muraro A, Roberts G, Worm M, et al. EAACI Food Allergy and Anaphylaxis Guidelines Group. Anaphylaxis: guidelines from the European Academy of Allergy and Clinical Immunology. Allergy 2014;69:1026–45.
28. Simons FE, Ebisawa M, Sanchez-Borges M, et al. 2015 update of the evidence base: World Allergy Organization anaphylaxis guidelines. World Allergy Organ J. 2015;8:32.
29. Tejedor Alonso MA, Moro Moro M, Múgica García MV. Epidemiology of anaphylaxis. Clin Exp Allergy 2015;45:1027–39.
30. Lieberman P, Nicklas RA, Oppenheimer J, et al. The diagnosis and management of anaphylaxis practice parameter: 2010 update. J Allergy Clin Immunol 2010;126(3):477–80.
31. Panesar SS, Javad S, de Silva D, et al. EAACI Food Allergy and Anaphylaxis Group. The epidemiology of anaphylaxis in Europe: a systematic review. Allergy 2013;68:1353–61.
32. Pumphrey RS. Fatal anaphylaxis in the UK, 1992-2001. Novartis Found Symp 2004;257:116–28.
33. Grabenhenrich LB, Dölle S, Moneret-Vautrin A, et al. Anaphylaxis in children and adolescents: The European Anaphylaxis Registry. J Allergy Clin Immunol 2016;137:1128–37.
34. Turner PJ, Gowland MH, Sharma V, et al. Increase in anaphylaxis-related hospitalizations but no increase in fatalities: an analysis of United Kingdom national anaphylaxis data, 1992-2012. J Allergy Clin Immunol 2015;135:956–63.
35. Moneret-Vautrin DA, Morisset M, Flabbee J, et al. Epidemiology of life-threatening and lethal anaphylaxis: a review. Allergy 2005;60(4):443–51.
36. Mullins RJ. Anaphylaxis: risk factors for recurrence. Clin Exp Allergy 2003;33(8):1033–40.
37. Vadas P, Gold M, Perelman B, et al. Platelet-activating factor, PAF acetylhydrolase, and severe anaphylaxis. N Engl J Med 2008;358:28–35.
38. Valent P. Risk factors and management of severe life-threatening anaphylaxis in patients with clonal mast cell disorders. Clin Exp Allergy 2014;44:914–20.
39. Bonadonna P, Pagani M, Aberer W, et al. Drug hypersensitivity in clonal mast cell disorders: ENDA/EAACI position paper. Allergy 2015;70:755–63.
40. Greenberger PA, Lieberman P. Idiopathic anaphylaxis. J Allergy Clin Immunol Pract 2014;2(3):243–50.
41. Wolbing F, Fisher J, Koberle M, et al. About the role and underlying mechanisms of cofactors in anaphylaxis. Allergy 2013;68:1085–92.
42. Umasunthar T, Leonardi-Bee J, Turner PJ, et al. Incidence of food anaphylaxis in people with food allergy: a systematic review and meta-analysis. Clin Exp Allergy 2015;45:1621–36.
43. International Collaborative Study of Severe Anaphylaxis. Risk of anaphylaxis in a hospital population in relation to the use of various drugs: an international study. Pharmacoepidemiol Drug Saf 2003;12(3):195–202.
44. Galvão VR, Castells MC. Hypersensitivity to biological agents-updated diagnosis, management and treatment. J Allergy Clin Immunol Pract 2015;3(2):175–85.
45. Steinke JW, Platts-Mills TA, Commins SP. The alpha-gal story: lessons learned from connecting the dots. J Allergy Clin Immunol 2015;135(3):589–96.
46. Weiss WE, Adkinson NF. Immediate hypersensitivity reactions to penicillin and related antibiotics. Clin Allergy 1988;18:515–40.
47. Reitter M, Petitpain N, Latarche C, et al. French network of regional pharmacovigilance centres. Fatal anaphylaxis with neuromuscular blocking agents: a risk factor and management analysis. Allergy 2014;69:954–9.
48. Rutkowski K, Wagner A. Chlorhexidine: a new latex? Eur Urol 2015;68:345–7.
49. Golden DB, Moffitt J, Nicklas RA, et al, Joint Task Force on Practice Parameters; American Academy of Allergy, Asthma & Immunology

(AAAAI); American College of Allergy, Asthma & Immunology (ACAAI); Joint Council of Allergy, Asthma and Immunology. Stinging insect hypersensitivity: a practice parameter update 2011. J Allergy Clin Immunol 2011;127:852–4.

50. Siracusa A, Folletti I, Gerth van Wijk R. Et al. Occupational anaphylaxis – an EAACI task force consensus statement. Allergy 2015;70:141–52.
51. Bilo MB, Cichocka-Jarosz E, Pumphrey R, et al. Self-medication of anaphylactic reactions due to Hymenoptera stings – An EAACI Task Force Consensus Statement. Allergy 2016;71:931–43.
52. Salama A, Kardashi R, Arbach O. Long-term treatment and transfusion of normal blood components following tolerance induction in patients with anti-IgA anaphylactic reactions. Transfus Med Hemother 2014;41:381–7.
53. Vuitton DA. Echinococcosis and allergy. Clin Rev Allergy Immunol 2004;26(2):93–104.
54. Gülen T, Hägglund H, Sander B, et al. Presence of mast cell clonality in patients with unexplained anaphylaxis. Clin Exp Allergy 2014;44:1179–87.
55. Brown SG. Cardiovascular aspects of anaphylaxis: implications for treatment and diagnosis. Curr Opin Allergy Clin Immunol 2005;5(4):359–64.
56. Lee S, Bellolio F, Hess EP, et al. Time of onset and predictors of biphasic anaphylactic reactions: a systematic review and meta-analysis. J Allergy Clin Immunol Pract 2015;3:408–16.
57. Anaphylaxis. NICE guideline 14, December 2011; http://guidance.nice.org.uk/CG134.
58. McLean-Tooke A, Goulding M, Bundell C, et al. Postmortem serum tryptase levels in anaphylactic and non-anaphylactic deaths. J Clin Pathol 2014;67:134–8.
59. Simons FE, Sampson HA. Anaphylaxis: unique aspects of clinical diagnosis and management in infants (birth to age 2 years). J Allergy Clin Immunol 2015;135:1125–31.
60. Wölbing F, Biedermann T. Anaphylaxis: opportunities of stratified medicine for diagnosis and risk assessment. Allergy 2013;68:1499–508.
61. Morita E, Chinuki Y, Takahashi H. Recent advances of in vitro tests for the diagnosis of food-dependent exercise-induced anaphylaxis. J Dermatol Sci 2013;71:155–9.
62. Van Winkle RC, Chang C. The biochemical basis and clinical evidence of food allergy due to lipid transfer proteins: a comprehensive review. Clin Rev Allergy Immunol 2014;46:211–24.
63. Fernández-Rivas M, Bolhaar S, González-Mancebo E, et al. Apple allergy across Europe: How allergen sensitization profiles determine the clinical expression of allergies to plant foods. J Allergy Clin Immunol 2006;118:481–8.
64. Simons FE, Ardusso LR, Bilò MB, et al, World Allergy Organization. WAO guidelines for the assessment and management of anaphylaxis. World Allergy Organ J 2011;4:13–37.
65. Perel P, Roberts I, Ker K. Colloids versus crystalloids for fluid resuscitation in critically ill patients. Cochrane Database Syst Rev 2013;(2):CD000567.
66. Bauer CS, Vadas P, Kelly KJ. Methylene blue for the treatment of refractory anaphylaxis without hypotension. Am J Emerg Med 2013;31:264.
67. Warrier P, Casale TB. Omalizumab in idiopathic anaphylaxis. Ann Allergy Asthma Immunol 2009;102:257–8.
68. Borzutzky A, Morales PS, Mezzano V, et al. Induction of remission of idiopathic anaphylaxis with rituximab. J Allergy Clin Immunol 2014;134:981–3.

MULTIPLE-CHOICE QUESTIONS

1. Urticaria is defined by:
   A. Itchy wheals, angioedema, or both
   B. Urticarial rash for >6 weeks
   C. An inducible elicitor
   D. Spontaneous wheals
   E. Recurrent angioedema without wheals

2. The main mediator of chronic spontaneous urticaria (CSU) is:
   A. Interleukin-1
   B. Bradykinin
   C. Tryptase
   D. Histamine
   E. Cysteinyl leukotrienes

3. The first line of management of urticaria is:
   A. Oral corticosteroids
   B. Diet
   C. Nonsedating H1 antihistamines
   D. Omalizumab
   E. Cyclosporin

4. The first-line treatment of anaphylaxis is:
   A. Subcutaneous epinephrine
   B. Intramuscular epinephrine
   C. Nonsedating H1 antihistamines
   D. Oral corticosteroids
   E. Parenteral corticosteroids

5. What is the treatment of choice for refractory anaphylaxis in a patient who is taking beta-blockers?
   A. Prednisolone
   B. Glucagon
   C. Hydrocortisone
   D. Loratadine
   E. Salbutamol

43

Allergic Reactions to Stinging and Biting Insects

Arthur Helbling, Ulrich R. Müller

Insects can induce allergic reactions by stinging and biting. They sting humans to defend themselves; they bite to feed. With regard to allergies, stings by insects of the order Hymenoptera (bees, wasps, or ants) are much more important than bites, which can be inflicted by various insects, such as mosquitoes, horseflies, midges, fleas, and bed bugs. Hymenoptera stings can induce acute systemic allergic reactions that kill several hundred patients in Europe and America every year, whereas systemic allergic reactions to bites are very rare. Both stings and bites can induce local allergic or toxic reactions. Infections are only transmitted by bites.

ENTOMOLOGICAL ASPECTS[1-4]

Stinging Hymenoptera all belong to the suborder Aculeatae (Fig. 43.1) with the families Apideae, Vespidae, Formicidae, and Myrmicinae. Latin and popular names in English, which partly differ in Europe and the United States, are given in Table 43.1.

Apidae

In this family, the honeybee (*Apis mellifera*) (Fig. 43.2A) is clinically the most important cause of allergies. Stings occur either in the vicinity of beehives or when the insect feels threatened during human activities, such as cutting flowers or walking barefoot in the grass. Because the whole beehive survives over winter, stings can occur not only in spring and summer but occasionally also on warm winter days. Honeybees are brown and moderately hairy. In contrast to other Hymenoptera, they usually lose their barbed sting when stinging.

Bumblebees (*e.g., Bombus terrestris*) are increasingly used as pollinators in greenhouses and occasionally cause allergic sting reactions in greenhouse workers.[2] Bumblebees are distinctly larger and more hairy compared with honeybees, and most species have distinct yellow or white bands on their abdomen (Fig. 43.2B).

Vespidae

The vespids are divided into the subfamilies Vespinae and Polistinae, which differ morphologically in the junction between thorax and abdomen (Figs. 43.2C, D). Vespids are almost hairless, and in most species, the abdomen is striped black and yellow. Vespids do not usually lose their sting when stinging and may therefore sting several times, even the same victim. Because only the queen survives over winter, larger populations develop only in summer, and most stings occur in summer and fall. The subfamily Vespinae contains the three genera *Vespula, Dolichovespula,* and *Vespa.*

Vespula (*V. germanica* and *V. vulgaris* in Europe; *V. maculifrons* and *V. germanica* in the United States) are called *wasps* in Europe and *yellow jackets* in the US. They breed in the ground, in attics, or in shelters. The genus *Vespula* is, by far, the most aggressive. Stings occur not only near the nests but also more often while victims are eating outdoors.

Most species of *Dolichovespula* (*D. media, D. sylvestris, D. saxonica* in Europe; *D. maculata, D. arenaria* in the United States) look very similar to *Vespula,* with black and yellow stripes on the abdomen and only a slightly larger size. They can be distinguished from *Vespula* by the larger distance between the eyes and the mandibles. Only *D. maculata,* the bald-faced hornet in the United States, is easy to distinguish from other vespids by its mostly black abdomen. *Dolichovespula* build their nests in tree branches or under the roofs of houses. They sting almost exclusively in the vicinity of their nests.

The genus *Vespa* (*V. crabro,* the European hornet[Fig. 43.2E]; *V. orientalis*) is easy to distinguish from other vespids by its much larger size. The European hornet was brought to America in the late nineteenth century, whereas the Asian hornet (*V. velutina*) has recently (2005) been introduced to the South of France. The oriental hornet is present in southeastern Europe, Asia, and Africa.

Nevertheless, hornet stings are rare and occur almost exclusively in the vicinity of nests, which are usually in hollow tree trunks or bird nesting-boxes.

Polistinae (*P. annularis, P. exclamans, P. fuscatus,* "paper wasp" in the United States; *P. dominulus, P. gallicus, P. nympha,* "field wasp" in Europe) live mostly in the southern United States and in the Mediterranean area of Europe, but small colonies have been observed all over Europe except in the British Isles. Their small nests consist only of one womb and are built in trees or under roofs.

Ants (Myrmicinae, Formicinae)

In South and Central America, and in the southern states of the United States, fire ants (*Solenopsis invicta*[see Fig. 43.3F]; *S. richteri*) are responsible for many systemic allergic sting reactions.[3] Fire ants build their mounds in yards, playgrounds, and fields. Occasional allergic sting reactions to *Pogonomyrmex,* the North American harvester ant, have been described and extremely rarely to the European red ant, *Formica rufa.* In contrast, species of Myrmicinae, especially *Myrmecia pilosula,* the "jack-jumper ant," are an important cause of allergic sting reactions in southern Australia.[4] Another group of aggressive ants are the ponerinae of the genus *Pachycondyla,* including *P. chinesis* in the Far East

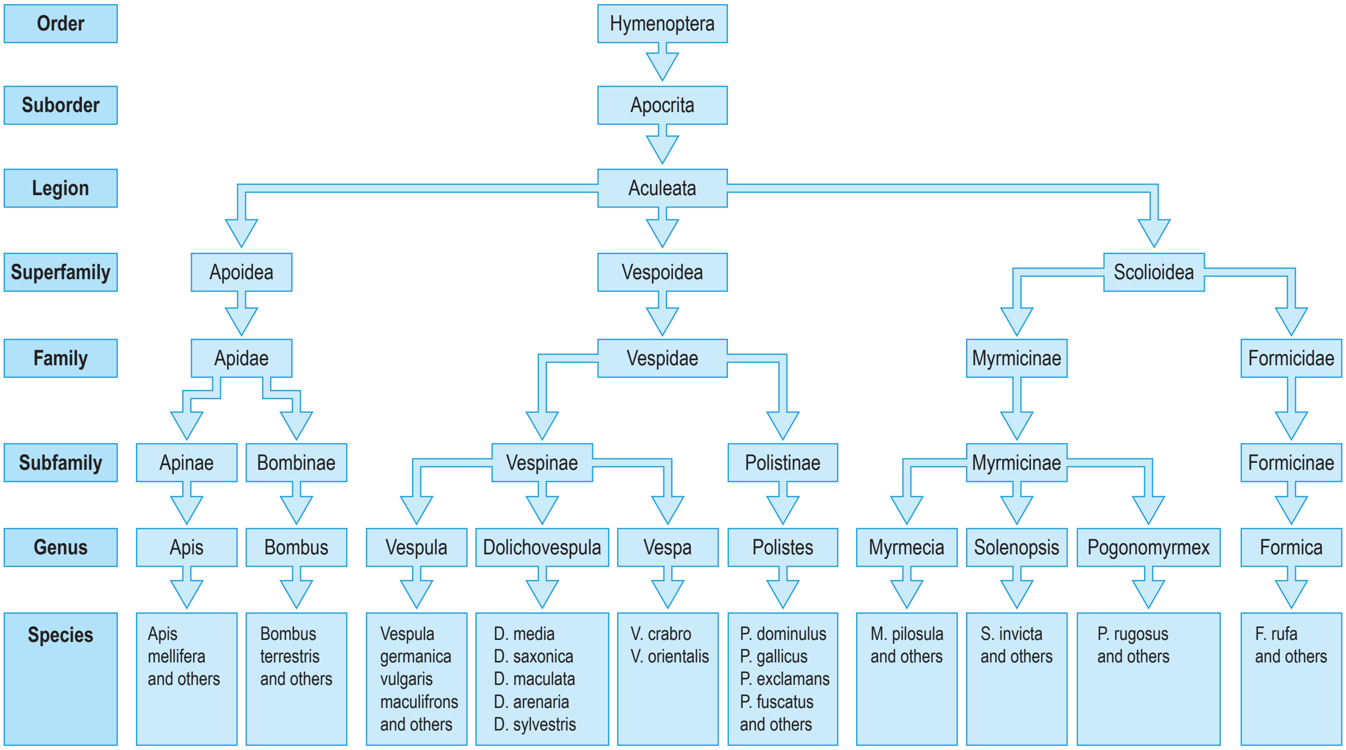

Order
Hymenoptera
Suborder
Apocrita
Legion
Aculeata
Superfamily
Apoidea
Vespoidea
Scolioidea
Family
Apidae
Vespidae
Myrmicinae
Formicidae
Subfamily
Apinae
Bombinae
Vespinae
Polistinae
Myrmicinae
Formicinae
Genus
Apis
Bombus
Vespula
Dolichovespula
Vespa
Polistes
Myrmecia
Solenopsis
Pogonomyrmex
Formica
Species
Apis mellifera and others
Bombus terrestris and others
Vespula germanica vulgaris maculifrons and others
D. media D. saxonica D. maculata D. arenaria D. sylvestris
V. crabro V. orientalis
P. dominulus P. gallicus P. exclamans P. fuscatus and others
M. pilosula and others
S. invicta and others
P. rugosus and others
F. rufa and others

FIG 43.1 Taxonomy of Hymenoptera.

TABLE 43.1 **Popular Names of the Most Frequent Hymenoptera in Europe and in the United States**

Latin Name	Popular Name in Europe	Popular Name in the United States
Apis mellifera	Honeybee	Honey bee
Bombus	Bumblebee	Bumble bee
Vespula ssp.	Wasp	Yellow jacket
Dolichovespula maculata	—	Bald-faced hornet
arenaria	—	Yellow hornet
media	Median wasp	—
Vespa crabro	Hornet	European hornet
Polistes spp.	Field wasp	Paper wasp
Solenopsis invicta	—	Fire ant

and *P. senna arensis* in the Middle East, which may also cause systemic allergic reactions.[3]

ALLERGENS IN HYMENOPTERA VENOMS

All Hymenoptera venoms contain low-molecular-weight substances, such as biogenic amines, phospholipids, amino acids, and carbohydrates, as well as peptides, such as melittin, apamin, and kinins, which contribute to the toxic effect but which—except melittin—are probably irrelevant with regard to allergies. The allergens of the important stinging Hymenoptera are shown in Table 43.2. Most of them are glycoproteins of 10–50 kilodaltons (kDa) and some, for example, dipeptidylpeptidases, are glycoproteins of up to 102 kDa.[5] The major allergens in bee venom are phospholipase A2, hyaluronidase, and acid phosphatase; in vespid venoms, antigen 5 and phospholipase A1. Ant venom from *S. invicta* contains a 37-kDa allergen and a 24-kDa allergen, with some sequence homology to phospholipase A1 and to antigen 5 from vespid venom.[3] In ant venom from *M. pilosula*, only pilosulin 1 has been identified as a major allergen so far. The genes of many major venom allergens have been cloned, and many of them have also been expressed as recombinant proteins comparable in allergenic activity and enzymatic function to their natural counterparts (see Table 43.2).[6-9] The amount of venom injected during a sting varies between and within species, especially in vespids. Bees release 50–140 μg venom per sting and vespids much less: between 2 and 17 μg.[2]

CLINICAL PICTURE

CLINICAL RELEVANCE

- Not every swelling after an insect sting is caused by allergy.
- Hymenoptera venom allergy is generally caused by systemic IgE.
- Although Hymenoptera venom allergy is not as prevalent as respiratory allergies, severe systemic reactions—even fatal—occur regularly worldwide.

Symptoms of venom hypersensitivity are most often caused by immunoglobulin E (IgE)-mediated immunological mechanisms, but occasionally by non–IgE-mediated immunological mechanisms. Rarely, nonimmunological mechanisms of mediator release play a role. The clinical presentation is classified into normal, large local, systemic allergic, systemic toxic, and unusual reactions.[1,2]

Normal Local Reactions

The normal local reaction of a nonallergic subject to a Hymenoptera sting consists of a painful, sometimes itchy, local wheal

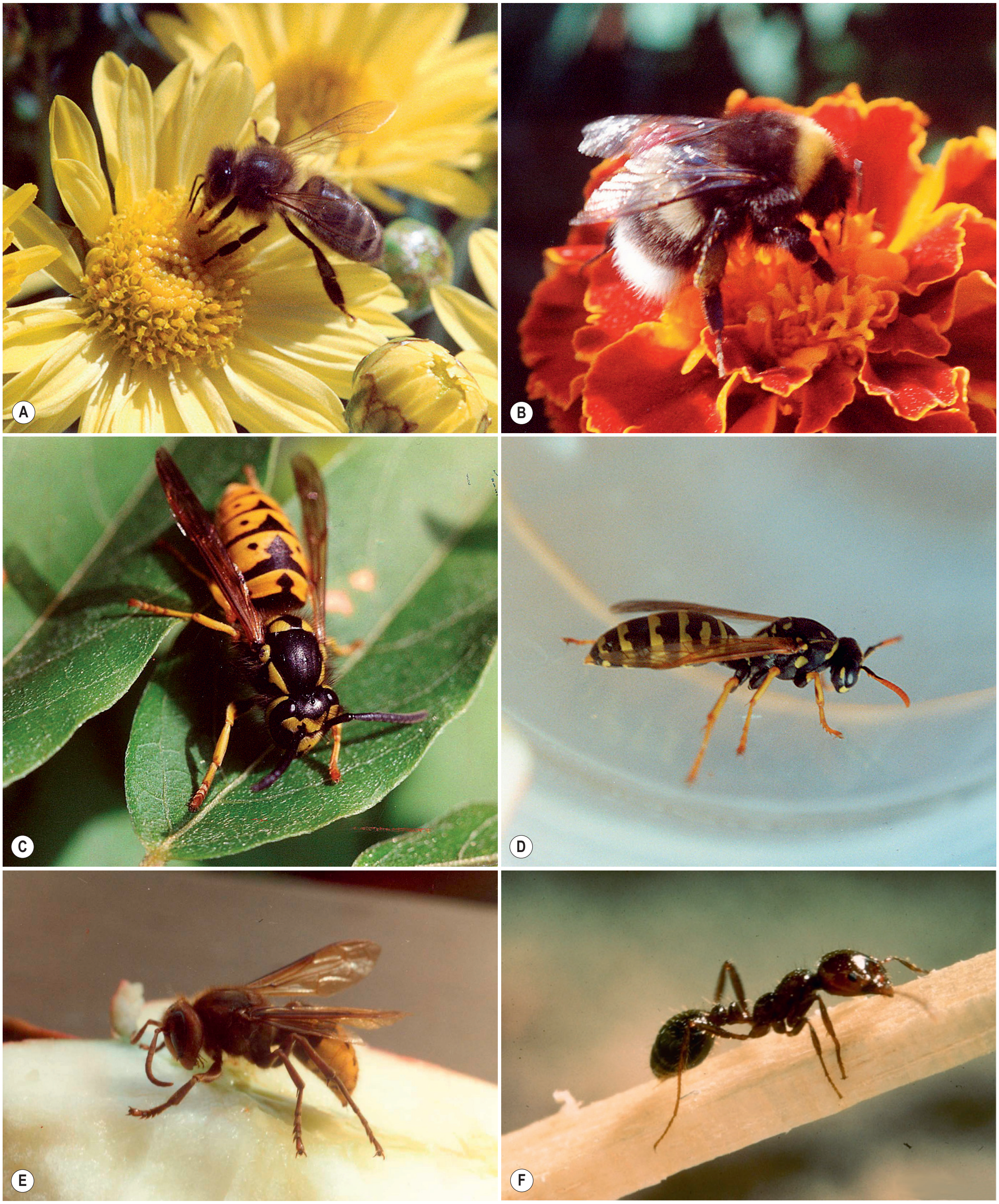


FIG 43.2 Common Hymenoptera. (A) Honeybee (*Apis mellifera*), (B) bumblebee (*Bombus terrestris*), (C) wasp/yellow jacket (*Vespula* spp.), (D) field wasp (*Polistes gallicus*), (E) European hornet (*Vespa crabro*), (F) fire ant (*Solenopsis invicta*). (Courtesy of US Department of Agriculture.)

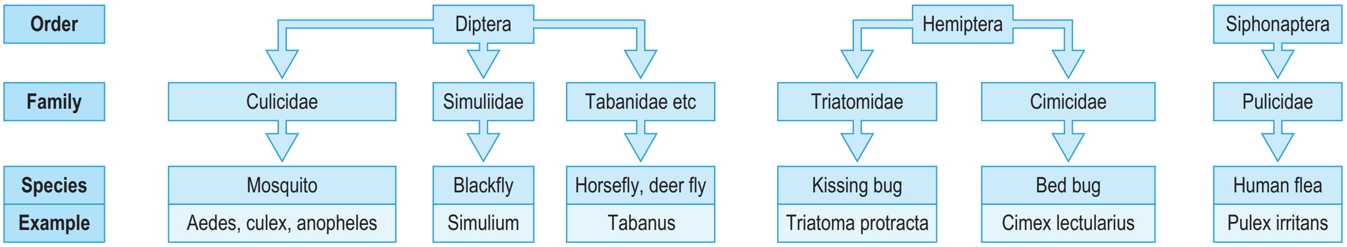

Order
Family
Species
Example
Diptera
Hemiptera
Siphonaptera
Culicidae
Simuliidae
Tabanidae etc
Triatomidae
Cimicidae
Pulicidae
Mosquito
Blackfly
Horsefly, deer fly
Kissing bug
Bed bug
Human flea
Aedes, culex, anopheles
Simulium
Tabanus
Triatoma protracta
Cimex lectularius
Pulex irritans

FIG 43.3 Biting insects that may cause allergic reactions.

TABLE 43.2 Allergens of Hymenoptera Venoms

Allergen	Molecular Weight (kDa)	Glycosylation	Major/Minor Allergen[a]	Expressed as Recombinant
Allergens in Bee Venom (*Apis mellifera*)				
Phospholipase A2 Api m1	16	Yes	Major	Yes
Hyaluronidase Api m2	45	Yes	Major	Yes
Acid phosphatase Api m3	49	Yes	Major	Yes
Melittin Api m4	2.8	No	Minor	Yes
Dipeptidylpeptidase Api m5	102	Yes	Major	Yes
Trypsin inhibitor Api m6	8	No	Minor	Yes
CUB serine protease Api m7	39	Yes	Major?	Yes
Carboxylesterase Api m8	70	Yes	Minor	
Carboxypeptidase Api m9	60	Yes	Minor	
Icarapin Api m10	29	Yes	Major	Yes
MRJP8[b]	65	Yes	?	
Api m11.0101 MRJP9 11.0201	60	Yes		
Vitellogenin Api m12	200	Yes	?	
Allergens of Wasp/Yellow Jacket (*Vespula* spp., Ves v) and Ant (*Solenosis invicta,* Sol i), *Myrmecia pilosula,* (Myr p) Venoms				
Vespula Vulgaris				
Phospholipase A1 Ves v1	34	No	Major	Yes
Hyaluronidase Ves v2	45	Yes	Minor	Yes
Dipeptidylpeptidase Ves v3	100	Yes	?	Yes
Antigen 5 Ves v5	23	No	Major	Yes
Vitellogenin Ves v6	200	Yes		
Solenopsis Invicta				
Phospholipase A1 Sol i1	37	Yes	Major	Yes
Sol i2	14	No	Minor	
Antigen 5 Sol i3	26	Yes	Major	Yes
Sol i4	12	No	Minor	
Myr p1	6.1		Minor	
Myr p2	5.6		Major	
Myr p3	8.2		Minor	

[a] Major allergen: ≥50% of allergic patients have specific immunoglobulin E (IgE) against. Minor allergen: <50% of allergic patients have specific IgE against.
[b] MRJP, major royal jelly protein.

and flare reaction, followed by a swelling of up to 5–10 cm in diameter. Usually, local symptoms resolve within a few hours. The fire ant (*S. invicta*) attaches to the skin by means of its powerful mandibles and then stings, releasing venom that produces a characteristic firelike pain. If not removed, it will continue to rotate in a pivotal fashion, repeatedly injecting further small amounts of venom.[10] After the ant stings, a vesicle is left behind, which later develops into a pustule that only heals after 1–2 weeks.[3]

Large Local Reactions

Large local reactions (LLRs) are defined as swellings around the sting site exceeding 10 cm in diameter, developing minutes to hours after the sting and lasting more than 24 hours.[1] LLRs may cause significant discomfort, especially when they last for days or even weeks and involve a whole limb, eyelids, or lips. Sometimes they are accompanied by lymphadenopathy or lymphangitis. They may also be associated with nonspecific systemic inflammatory symptoms, such as malaise, fever, shivering, or headache. However, the development of local infection, abscesses, or a phlegmon at the sting site is inhibited by the bacteriostatic effect of Hymenoptera venoms. In contrast, scratching after stings by the American fire ant or bites from blood-sucking insects, such as midges, can lead to skin infection.

The pathogenesis of LLRs is thought to be based on IgE-mediated and/or cell-mediated immune mechanisms, most likely a combination of both.[1,2]

Systemic Reactions

Systemic reactions (SRs), including anaphylactic reactions, are usually mediated by IgE (Chapter 42). Affected organs include skin (pruritus, urticaria, flush, angioedema), the gastrointestinal (GI) tract (cramps, vomiting or diarrhea, dysphagia), the respiratory tract (laryngeal edema, bronchial obstruction, pulmonary edema), and the cardiovascular system (arterial hypotension, shock, arrhythmias, loss of consciousness with incontinence).

The most commonly used classification of SRs was developed by Müeller.[11] Symptoms appear most often within a few minutes to 1 hour after the sting. The patient recovers usually within a few hours. A protracted course over more than a day or a biphasic course has been described but is rare.

Lasting morbidity, for example, myocardial or cerebrovascular infarction, as a consequence of SRs or even fatal reactions can occur, but these are rare (see "Epidemiological aspects").

Systemic Toxic Reactions

Toxic reactions are dose dependent, and a clinically significant toxic effect only needs to be considered after multiple stings—usually 50 to several hundred.[1,2] The main toxic effects develop within hours to days and comprise rhabdomyolysis and intravascular hemolysis leading to acute renal failure with tubular necrosis. Myocardial damage, hepatic dysfunction, coagulation disorders, and brain edema and/or necrosis may occur. The number of stings needed to cause a fatal reaction in adults varies between 200 and 1000. In small children, however, fewer than 50 stings may be lethal. In most cases, death is not immediate but occurs after several days.

Unusual Reactions

Various unusual sting reactions can occur: these are rare and appear after hours to days. More than half of them follow immediate local or SRs.[1,12] Non–IgE-mediated immunological mechanisms are likely involved. The causal relation to the sting event often remains uncertain. Serum sickness–like syndromes with fever, arthralgias, exanthema, and lymphadenopathy are well documented. Less frequently reported complications of Hymenoptera stings include diseases of the nervous system (peripheral neuropathy, polyradiculomyelitis, extrapyramidal syndromes, acute disseminated encephalomyelitis), kidneys (glomerulonephritis, interstitial nephritis), or blood and blood vessels (hemolytic anemia, thrombocytopenia, Henoch-Schönlein syndrome and other forms of vasculitis). In these situations, a causal relation to the sting is less well documented.

EPIDEMIOLOGICAL ASPECTS

Prevalence of Allergy to Stings by Flying Hymenoptera

KEY CONCEPTS

- Awareness of the most relevant insects, specifically Hymenoptera and their entomological subdivisions, causing immunoglobulin E (IgE)–mediated allergies
- Assessment of clinical symptoms and classification of Hymenoptera venom allergy
- Evaluation of persons and circumstances at risk for Hymenoptera venom allergy
- Knowledge of appropriate diagnosis of Hymenoptera venom allergy, potential diagnostic and therapeutic control test tools

Cumulative lifetime sting rates of 61–95% have been reported in people aged 16–65 years. Of course, this can vary considerably in different regions of the world. Asymptomatic sensitization with the development of specific IgE to bee and wasp venoms is a frequent finding and has been reported in up to 40%.[13] Hymenoptera sting allergy can occur at any age. In general, because of their higher level of outdoor activities, men are more frequently stung compared with women, and children more often compared with adults.[14]

The reported cumulative lifetime prevalence of LLRs ranges from 2% to 26%, that of SRs from 0.3% to 5%.[2] Among beekeepers, it varies between 14% and 43%.[15] LLRs are not predictive for SRs in the event of a subsequent Hymenoptera sting.[13] In Europe and the United States, SRs are more frequently caused by vespids than by honeybees. In the southern states of the United States and in Australia, ants are important causes of SR.

Risk Factors for Hymenoptera Allergy

The risk of developing a sting allergy increases with the number of stings, especially if two stings occur within a short period (up to 2 months).[16] However, beekeepers stung less than 10 times a year have a much higher risk for SRs compared with those stung more than 200 times a year.[15] Even after successful venom immunotherapy, the risk of recurrent SRs after Hymenoptera stings persists throughout life.[17] Cardiovascular diseases and their treatment with β-blocking drugs and angiotensin-converting enzyme (ACE) inhibitors are associated with more severe sting reactions, sometimes also with lasting morbidity caused by cardiac or cerebrovascular infarction as a result of anaphylaxis.[1,18,19] However, β-blockers do not increase the overall risk of systemic sting reactions.[18] Systemic mastocytosis (indicated by elevated baseline serum tryptase) is a risk factor for severe or even fatal systemic sting reactions.[19] Atopy is not more frequent in patients with Hymenoptera sting allergy than in the general population. However, in patients with atopy and Hymenoptera venom allergy,

TABLE 43.3 Natural History of Hymenoptera Venom Allergy Based on Prospective Studies With Sting Reexposure in Patients Without Venom Immunotherapy

			% WITH SYSTEMIC REACTION TO REEXPOSURE BY			
Previous Reaction	**Author, Year**	**Reexpo-sure by**	**Bee or Vespid**	**Bee**	**Vespid**	**Ant**
Large local reactions	Müller, 1990	FS	6			
Systemic reactions (SRs)	van der Linden, 1994	CH		51	25	
Mild SRs in children	Schuberth, 1983	FS	16			
Mild SRs adults	Blaauw, 1985	CH		31	10	
Severe SRs in adults	Blaauw, 1985	CH	44	60	33	
SR in controls of controlled studies	Hunt, 1978	CH	58			
	Müller, 1979	FS		75		
	Brown, 2003	CH				72

Summarized in Rueff, Przybilla, Müller et al Allergy 1996;51:21-25. FS, field sting; CH, sting challenge.

SRs may be more severe and more often affect the respiratory tract. About 50% of beekeepers with venom allergy have atopy, and they may become sensitized by inhaling dust containing venom when working with beehives.[15]

Mortality Caused by Hymenoptera Stings

Mortality resulting from Hymenoptera stings varies from 0.09 to 0.48 deaths per million persons per year.[20] Over the past 40 years, the mean annual incidence of fatal Hymenoptera sting reactions in Switzerland has been 3.1; when extrapolated to the whole of Europe, this would be about 200. The vast majority of fatal sting reactions occur in adults over 45 years of age.[2,20] Additional risk factors include a positive history of sting allergy, male gender, and stings on the head or neck. Autopsies have documented preexisting cardiovascular disease in most cases.[20] Approximately half of the deaths occur in subjects with no known prior history of allergic sting reactions.

Natural History of Hymenoptera Sting Allergy (Table 43.3)

In prospective clinical studies, after an LLR, the risk of developing an SR to the next sting is between 5% and 10%; after a mild SR, the risk is between 15% and 30%; and after a severe SR, the risk is between 50% and 75%.[13,21] The severity of an index sting reaction is an important factor determining the risk at reexposure. Children are at lower risk of re-sting SR compared with adults.[1,22] The re-sting SR risk is definitely lower in patients with *Vespula* allergy than in those with bee venom allergy, probably because of the smaller and more variable amount of venom injected.

Epidemiological Aspects of Allergic Reactions to Ant Stings

Nearly 50% of inhabitants are stung each year in areas of the United States that are fire ant endemic.[3,10] Many report LLRs, but up to 1% of patients who are stung by imported fire ants develop anaphylaxis, and some deaths have been reported. In Australia, more than 90% of ant venom anaphylaxis is caused by *M. pilosula.* Age >35 years, annual sting rate, and allergy to bee venom were predictive factors for more severe reactions.[23]

DIAGNOSIS

History

As in all forms of allergy, the clinical history is key to diagnosing Hymenoptera sting allergy. This includes the date, number, and circumstances of stings (*e.g.,* environment, activities); kind and severity of symptom; sting site; retained or removed stinger; interval to onset of symptoms; emergency treatment; risk factors for a particularly severe reaction (*e.g.,* comorbidity, drugs, elevated baseline serum tryptase); tolerated stings after the first SR; reduction of the quality of life[24]; and other allergies.[1,2] In individuals with only LLRs, no further diagnostic tests are recommended. For those with SRs, diagnostic tests include skin tests, venom-specific serum IgE antibodies, and, in severe SRs, baseline serum tryptase.

Skin Tests

Skin tests should not be performed until at least 3 weeks after an SR, as false-negative results can occur (the refractory period). They are performed by intradermal or skin prick endpoint titration.[2] Injection of 0.02 mL of venom solution is given intradermally in increasing concentrations—0.00001–1 µg/mL—into the volar surface of the forearm. For skin prick tests, concentrations of 0.01–300 µg/mL are used. However, even at 300 µg/mL, the sensitivity of skin prick is clearly lower than that of the intradermal test. We therefore prefer the intradermal test.

Venom-Specific Serum IgE Antibodies

Several different *in vitro* immunoassays for the detection of venom-specific serum IgE antibodies (sIgE) are commercially available (Chapter 95). Immediately after a sting, sIgE may be low or undetectable, but it usually increases within days or weeks after an SR. If no sIgE is found in a patient with a clear history of venom allergy, the test should be repeated after 2–4 weeks.[2]

Elevated sIgE levels after Hymenoptera stings that cause a normal sting reaction are not predictive for future development of an SR. [13] sIgE can be determined against the whole venom or components of the venoms. Measuring sIgE to recombinant venom allergens has improved the precision of allergic diagnostic tests. Tests are currently available for the species-specific, nonglycosylated major allergens Api m 1, Api m 10, Ves v 5, and Ves v 1. Tests for Api m 2, Api 3, and Api m 5 will be available soon.[6,25]

Sensitivity and Specificity of Skin Tests and sIgE

The sensitivity of these tests is >90% in patients with a history of SR within the past year and decreases steadily thereafter with or without venom immunotherapy but may stay positive for many years. However, their specificity is limited, since up to 20% of an unselected population have positive results, whereas only 0.3–5% have a history of allergic sting reactions.[1,2] As a rule, the intradermal test result remains positive longer than that of sIgE. However, there is currently no reliable test that predicts the risk of future SRs in untreated or treated patients. Despite a history of typical SRs to stings, a few patients have no detectable sIgE and have negative skin test results.[26] This may be attributed to insufficient sensitivity of the available tests, a long interval between the SR and testing with spontaneous decrease of sensitization, or non–IgE-mediated pathogenesis.

Specificity may cause problems: more than 20% of people without a history of SR have a positive diagnostic test result.[13] Although it is difficult to exclude sensitization after a previous sting, this positivity may reflect cross-reactivity (see below).

For skin testing and venom immunotherapy for fire ant (*S. invicta*) stings, only whole-body extracts are currently available. These extracts have good sensitivity but low specificity. Therefore they should only be used in patients with a history of systemic sting reactions, at least 30 days after the SR.[10] In contrast, venoms of *M. pilosula* have been shown to have excellent sensitivity and specificity.[4]

Cross-Reactivity

Cross-reactivity between venom allergens is strong between species within a family, for example, among *Vespula, Dolichovespula,* and *Vespa,* but only limited between *Vespinae* and *Polistinae* or between honeybees and bumblebees.[6] There is some homology at the protein level between bee and vespid venoms, with ~50% sequence identity between the hyaluronidases and dipeptidylpeptidases of the two families, but double-positivity with diagnostic tests to both venoms is frequently observed. This may reflect true double-sensitization or cross-reactivity. Cross-reacting carbohydrate determinants (CCDs) are present in many major Hymenoptera venom allergens, such as hyaluronidase, acid phosphatase, and phospholipase A2, but also in many plant proteins (*e.g.,* rapeseed pollen or bromelain). CCDs are certainly responsible for a significant proportion of the double-positivity of diagnostic tests to bee and vespid venoms. They may also explain why individuals with no history of SRs have positive test results. The CCDs are probably of no clinical relevance.[27] The radioallergosorbent test (RAST)for venoms and CCDs can help distinguish between true double-sensitization and cross-reactivity but is not always conclusive.

Assessing IgE antibodies to species-specific nonglycosylated, recombinant major allergens (Api m1 [phospholipase A2] of bee venom and Ves v5 [antigen 5] of *Vespula* venom) reduces the rate of double-positivity very significantly and is helpful in choosing venoms for immunotherapy.[6,25,28]

Some cross-reactivity between allergens of vespid venoms and those of *S. invicta* has also been documented.[3,6]

Cellular Tests

If routine tests in patients with a history of SRs are negative, cellular tests may be helpful to demonstrate sensitization.[2]

In the *basophil histamine release test* peripheral blood leukocytes are incubated with venom allergens. The reaction with cell-bound IgE antibodies leads to histamine release from basophils. In the cellular antigen stimulation test (CAST), leukocytes of patients are prestimulated with interleukin-3 (IL-3) and exposed to venom allergens. The released sulfidoleukotrienes are determined by using enzyme-linked immunosorbent assay (ELISA).

The basophil activation test (BAT) is based on flow-cytometric demonstration of an altered membrane phenotype of basophils stimulated by IL-3 and allergen exposure. At present, the most commonly used expression marker is CD63. The sensitivity and specificity of BAT seems to be superior to skin tests and venom-specific sIgE. It may also have better predictive value. The test must, however, be performed with fresh blood; it is expensive and not yet well standardized: there are few data on specificity and predictive value in relation to a sting reexposure during or after venom immunotherapy.

Allergen-Specific IgG

The presence of sIgG and IgG4 primarily reflects exposure to the respective venom. sIgG titers increase after a sting, irrespective of the presence or absence of an allergic sting reaction. Venom immunotherapy induces a rise in sIgG. However, there is no close correlation between the concentration of sIgG or the sIgE/sIgG ratio and the clinical response to a resting during or after venom immunotherapy.[1] Routine assessment of sIgG is therefore not recommended but may be helpful if the responsible insect is unidentified.

Baseline Serum Tryptase

Because elevated baseline serum tryptase levels (>11.4 μg/L) are associated with severe, sometimes IgE-negative, systemic sting reactions and with cutaneous or systemic mastocytosis, this enzyme should be determined in all patients with a history of SRs.[19] The commercially available fluorescence immunoassay measures total tryptase. α-Tryptase is secreted continuously and reflects whole-body mast cell load. Elevated values are seen in cutaneous and systemic mastocytosis. β-Tryptase is released during mast cell activation and is a marker of anaphylaxis.

Sting Challenge Tests

Sting challenge with a live insect is not recommended as a diagnostic tool in untreated patients, but a sting challenge under well-supervised clinical conditions may be helpful in evaluating the efficacy of venom immunotherapy.[21,29] However, a tolerated sting challenge does not definitely exclude a reaction to future stings after immunotherapy, especially if these are repeated.[17]

PREVENTION AND TREATMENT

Prevention

All patients with a history of SRs should receive detailed instruction on the avoidance of future stings and measures to take if stung.[1] Bee stings occur most often when walking barefoot on grass, and wasp stings occur when eating outdoors, in orchards with fallen fruits, and near open waste-bins. The risk of a sting is especially high near beehives or vespid nests. While gardening, wearing long trousers, shirts with long sleeves, and gloves are recommended. Strongly scented perfumes, sun creams, or shampoos, as well as brightly colored garments, should be avoided.

Treatment of Large Local Reactions

Oral antihistamines and cooling of the sting site (*e.g.,* with ice cubes) reduces local swelling, pain, and itching. Antiinflammatory ointments or topical corticosteroids may diminish the local inflammatory process. In cases of severe swellings, oral corticosteroids together with antihistamines over several days are recommended.[10]

Systemic Allergic Reactions

Sympathomimetics, antihistamines, and corticosteroids are the most effective drugs for symptomatic treatment of SRs. All patients experiencing SRs should seek medical advice and should be medically observed until the symptoms resolve and blood pressure is stable.

Mild reactions confined to the skin may be treated with rapid-acting oral antihistamines alone. If respiratory or cardiovascular symptoms occur, intramuscular epinephrine must be given immediately, intravenous (IV) access should be established, and antihistamines and corticosteroids given intravenously (Chapter 42). All patients with severe SRs should be hospitalized and supervised until they recover completely. Patients with cardiovascular symptoms must be treated and transported in the supine position, and IV volume replacement is indicated.

Subsequently, every patient who has experienced a systemic allergic sting reaction should be investigated with a view to prevention measures and immunotherapy.[10]

Emergency Medication Kit

All patients with a history of SRs should carry an emergency kit for self-administration.[30] After a sting, patients should immediately take both antihistamines and corticosteroids, whether or not symptoms occur. In case of systemic symptoms, such as urticaria, dyspnea, generalized weakness, or dizziness, epinephrine should be administered intramuscularly via an autoinjector, such as the Epipen (dosage 0.3–0.5 mg of epinephrine).[30] In children with body weight <25 kg, Epipen Junior (0.15 mg epinephrine) or an equivalent product should be administered, along with half the adult dose of antihistamines and corticosteroids. If any SR occurs, medical care must be sought immediately.

Venom Immunotherapy[1,10]

Indications

Venom immunotherapy (VIT) is indicated in children and adults with a history of severe SRs (grade III/IV), provided sensitization to the relevant venom has been demonstrated by skin and/or blood test. LLRs or unusual reactions do not require VIT. VIT is also recommended for patients who experience repeated mild, non–life-threatening reactions and are at high risk for reexposure, such as beekeepers or their family members. Concomitant cardiovascular disease, mastocytosis, and strongly impaired quality of life because of the venom allergy are relative indications for VIT in patients with non–life-threatening sting reactions.[19,24]

Contraindications for VIT are the same as for immunotherapy with other allergens (Chapter 91).

Dosage and Treatment Regimens

The recommended maintenance dose for both children and adults is 100 µg of the venom. This maintenance dose is equivalent to approximately two bee stings or several vespid stings. A higher dose (*e.g.,* 200 µg) is recommended when SRs occur after reexposure to a field sting or a sting challenge.[29] In highly exposed subjects, such as beekeepers or professional gardeners, a maintenance dose of 200 µg is advised.

VIT may be initiated by a conventional or an ultra-rush protocol. During the first year, maintenance VIT is given every 4 weeks. Subsequently, the intervals may be extended to 6 weeks if VIT is well tolerated.

Adverse Reactions to VIT

The overall incidence of systemic adverse reactions to VIT varies between 5% and 40%.[1,10] VIT with bee venom causes side effects more often than with *Vespula* venom.[29] Ultra-rush protocols are associated with a higher rate of side effects compared with conventional protocols. Most systemic side effects are mild; only one-third of patients require medical treatment.

Premedication with antihistamines reduces large local and other cutaneous reactions, such as urticaria, but severe SRs may not be suppressed. Some units routinely give antihistamines 2 hours before up-dosing injections until the maintenance dose has been repeatedly well tolerated. In patients with recurrent SRs during up-dosing, off-label use of anti-IgE (omalizumab) has made it possible to complete up-dosing and to continue maintenance VIT.[31]

Efficacy of VIT

In addition to three prospective controlled trials, the efficacy of VIT has been confirmed by well-tolerated sting challenges during VIT in several uncontrolled prospective studies.[1,10] Full protection is achieved in 80–85% of patients receiving bee venom and in 95–100% of patients receiving *Vespula* venom. The higher risk of treatment failure in bee venom allergy may relate to differences in venom compositions. Important allergens, such as Api m10 or Api m 3, may be missing or present in low concentrations in some VIT products.[32] It has therefore been suggested that patients with a predominant sensitization to these allergens may not be properly protected by VIT.[33]

The efficacy of immunotherapy with commercially available whole-body extract of the ant *Solenopsis* has not been documented in controlled studies. However, excellent results, comparable with those of VIT with *Vespula* venom, were obtained in a double-blind placebo-controlled study using *M. pilosula* venom.[4]

Duration of VIT[17]

Lifelong treatment may be the safest recommendation, but in most allergy centers, VIT is given for up to 5 years. If VIT is given for at least 3 years, >80% of both adults and children are still protected, when reassessed 1–7 years after discontinuation. Longer courses of treatment should be considered in high-risk patients, such as those with very severe systemic sting reactions, coexisting cardiovascular or pulmonary disease, and systemic allergic reactions to VIT or stings during VIT and in subjects with elevated basal serum tryptase levels. Lifelong VIT is advised for patients with cutaneous or systemic mastocytosis.[10]

Risk Factors for Recurrence of SRs After Stopping VIT

A number of risk factors have been identified for relapse of Hymenoptera venom allergy after discontinuation of VIT[17]: in general, adults, especially older people, because of concomitant diseases, have a less favorable prognosis compared with children. Patients with bee venom allergy have a higher relapse risk compared with those allergic to *Vespula* venom. Patients with more severe pretreatment reactions and those with more SRs during VIT (to injections or field stings) have a higher risk for recurrent SRs to Hymenoptera stings. The risk of relapse is lower after 5 years of VIT, compared with only 3 years of VIT.

ALLERGIC REACTIONS TO BITING INSECTS

Biting insects may cause local allergic reactions as a result of sensitization to their salivary proteins introduced during the process of blood sucking. Systemic reactions are very rare. The responsible insects belong to the orders Diptera, Hemiptera, and Siphonaptera (Fig. 43.3).[34]

Clinical Symptoms

Local reactions to insect bites may be characterized by immediate wheal and flare, or may be delayed, with pruritic erythema and papules developing after 12–24 hours and lasting for days to weeks; combined reactions can also occur. In delayed and combined reactions, vesicular, bullous, or even necrotic lesions may develop. SRs are much rarer than those to Hymenoptera stings. They have, however, occasionally been described, especially after bites by horseflies (*Tabanus* spp.) and the kissing bug (*Triatoma protracta*).

Allergens

The salivary proteins involved have either digestive (amylases, esterases) or hemostatic function (*e.g.*, factor Xa inhibition). Numerous IgE-binding proteins with a molecular weight of 15–81 kDa have been described, especially in the saliva of mosquitoes, and in horseflies and kissing bugs as well; some of these proteins have been cloned.[35]

Prevention and Treatment

The application of insect repellents and prophylactic intake of antihistamines can prevent or mitigate outdoor exposure. Screens on windows and doors and mosquito nets over the beds are effective in homes. Bedbug infestation should be eliminated by using appropriate pesticides; flea infestation of pet animals should be managed by a veterinarian. Local reactions may be treated with topical steroids or oral antihistamines.

ON THE HORIZON

The availability of recombinant venom allergens offers several promising perspectives for diagnosis and immunotherapy.[5-9]

- The specificity of these tests could be considerably improved by using recombinant cocktails with species-specific major allergens without cross-reacting carbohydrate determinants (CCDs) for diagnosis instead of the whole venom.
- In patients with double-positive diagnostic tests to honeybee and *Vespula* venom, recombinant species-specific nonglycosylated major allergens from honeybee and *Vespula* venom make it possible to distinguish reliably between true double sensitization and cross-reactivity, an important issue for the choice of venoms for immunotherapy.
- Addition of such allergens to therapy solutions can increase the efficacy of venom immunotherapy (VIT) with bee venom, as these are lacking or insufficiently represented in currently available preparations.
- Allergic side effects of VIT might be reduced by using modified major allergens with reduced immunoglobulin E (IgE) binding, but retained specific T-cell interaction.[36]

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Müller U. Insect sting allergy. Stuttgart, Germany: Gustav Fischer Verlag; 1990.
2. Bilò MB, Ruëff F, Mosbech H, et al. Diagnosis of Hymenoptera venom allergy. Allergy 2005;60:1339–49.
3. Hoffmann DR. Ant venoms. Curr Opin Allergy Clin Immunol 2010;10:342–6.
4. Brown SGA, Franks RW, Baldo BA, et al. Prevalence, severity, and natural history of jack jumper ant venom allergy in Tasmania. J Allergy Clin Immunol 2003;111:187–92.
5. Ollert M, Blank S. Anaphylaxis to insect venom allergens: role of molecular diagnostics. Curr Allergy Asthma Rep 2015;15:26.
6. Müller U. Insect venoms. In: Ring J, editor. Anaphylaxis, chemical immunol allergy. Basel, Switzerland: Karger; 2010. p. 141–56.
7. King TP, Guralnick M. Hymenoptera allergens. In: Lockey R, Ledford D, editors. Allergens and allergen immunotherapy, vol. 21. 4th ed. (Clinical Allergy and Immunology). Boca Raton, FL: Taylor & Francis; 2008. p. 237–49.
8. Grunwald T, Bockisch B, Spillner E, et al. Molecular cloning and expression in insect cells of honeybee venom allergen acid phosphatase (Api m 3). J Allergy Clin Immunol 2006;117:848–54.
9. Seismann H, Blank S, Braren I, et al. Dissecting cross-reactivity in Hymenoptera venom allergy by circumvention of a-1,3-core fucosylation. Mol Immunol 2010;47:799–808.
10. Müller U, Golden DBK, Lockey RF. Byol S. Immunotherapy for Hymenoptera venom hypersensitivity. In: Lockey R, Ledford D, editors. Allergens and allergen immunotherapy, vol. 21. 4th ed. (Clinical Allergy and Immunology). Boca Raton, FL: Taylor & Francis; 2008. p. 377–92.
11. Mueller HL. Diagnosis and treatment of insect sensitivity. J Asthma Res 1966;3:331–3.
12. Reisman RE. Unusual reactions to insect stings. Curr Opin Allergy Clin Immunol 2005;5:355–8.
13. Sturm GJ, Kranzelbinder B, Schuster C, et al. Sensitization to Hymenoptera venoms is common, but systemic sting reactions are rare. J Allergy Clin Immunol 2014;133:1635–43.
14. Graif Y, Confino-Cohen R, Goldberg A. Allergic reactions to insect stings: Results from a national survey of 10 000 junior high school children in Israel. J Allergy Clin Immunol 2006;117:1435–9.
15. Müller U. Bee venom allergy in beekeepers and their family members. Curr Opin Allergy Clin Immunol 2005;5:343–7.
16. Antonicelli A, Bilò MB, Bonifazi F. Epidemiology of Hymenoptera allergy. Curr Opin Allergy Clin Immunol 2002;2:341–6.
17. Müller UR, Ring J. When can immunotherapy for insect sting allergy be stopped? J Allergy Clin Immunol Pract 2015;3:324–8.
18. Müller U, Haeberli G. Use of beta-blockers during immunotherapy for Hymenoptera venom allergy. J Allergy Clin Immunol 2005;115:606–10.
19. Ruëff F, Przybilla B, Bilò M, et al. Predictors of severe anaphylactic reactions in patients with hymenoptera venom allergy: importance of baseline serum tryptase. J Allergy Clin Immunol 2009;124:1047–55.
20. Sasvary T, Müller U. Fatalities from insect stings in Switzerland 1978 to 1987. Schweiz Med Wschr 1994;124:1887–94.
21. Ruëff F, Przybilla B, Müller U, et al. The sting challenge test in Hymenoptera venom allergy. Allergy 1996;51:216–25.
22. Golden DBK, Kagey-Sobotka A, Norman PS, et al. Outcomes of allergy to insect stings in children, with and without venom immunotherapy. N Engl J Med 2004;351:668–74.
23. Brown SG. Cardiovascular aspects of anaphylaxis: implications for treatment and diagnosis. Curr Opin Allergy Clin Immunol 2005;5:359–64.
24. Oude Elberink JN, Dubois AE. Quality of life in insect venom allergic patients. Curr Opin Allergy Clin Immunol 2003;3:287–93.
25. Frick M, Müller S, Bantleon F, et al. rApi m 3 and rApi m 10 improve detection of honey bee sensitization in Hymenoptera venom-allergic patients with double sensitization to honey bee and yellow jacket venom. Allergy 2015;70:1665–8.
26. Golden DBK, Kagey-Sobotka A, Norman PS, et al. Insect sting allergy with negative venom skin test responses. J Allergy Clin Immunol 2001;107:897–901.
27. Hemmer W, Focke M, Kolarich D, et al. Antibody binding to venom carbohydrates is a frequent cause for double positivity to honey bee and yellow jacket venom in patients with stinging insect allergy. J Allergy Clin Immunol 2001;108:1045–52.
28. Müller U, Johansen N, Petersen A, et al. Hymenoptera venom allergy: analysis of double positivity to honey bee and *Vespula* venom by estimation of species-specific major allergens Api m1 and Ves v5. Allergy 2009;64:543–8.
29. Ruëff F, Vos B, Elberink JO, et al. Predictors of clinical effectiveness of Hymenoptera venom immunotherapy. Clin Exp Allergy 2014;44:736–46.

30. Bilò MB, Cichocka-Jarisz E, Pumphrey R, et al. Self-medication of anaphylactic reactions due to Hymenoptera stings – An EAACI Task Force Consensus Statement. Allergy 2016;71:931–43.
31. Sokol KC, Ghazi A, Kelly BC, et al. Omalizumab as a desensitizing agent and treatment in mastocytosis: a review of the literature and case report. J Allergy Clin Immunol Pract 2014;2:266–70.
32. Blank S, Seismann H, Michel Y, et al. Api m 10, a genuine A. mellifera venom allergen, is clinically relevant but underrepresented in therapeutic extracts. Allergy 2011;66:1322–9.
33. Frick M, Pfützner W, Helbling A, et al. Predominant Api m 10 sensitization as risk factor for treatment failure in honey bee venom immunotherapy. J Allergy Clin Immunol 2016;138:1663–71.
34. Hoffman DR. Allergic reactions to biting insects. In: Levine MI, Lockey RF, editors. Monograph on insect allergy. Pittsburgh, PA: Dave Lambert Associates; 2003. p. 161–74.
35. Simons FER, Peng Z. Mosquito allergy. In: Levine MI, Lockey RF, editors. Monograph on insect allergy. Pittsburgh, PA: Dave Lambert Associates; 2003. p. 175–203.
36. Karamloo R, Schmid-Grendelmeier P, Kussebi F, et al. Prevention of allergy by a recombinant multi-allergen vaccine with reduced IgE binding and preserved T cell epitopes. Eur J Immunol 2005;35:3268–76.

MULTIPLE-CHOICE QUESTIONS

1. Which factor has NO influence on the severity of an allergic reaction as a result of a Hymenoptera sting?
   A. Elevated basal tryptase
   B. Treatment with β-blockers
   C. Two yellow jacket stings within 3 weeks
   D. Atopy
   E. Previous repeated severe allergic sting reactions

2. The average amount of venom released by a single bee sting is approximately:
   A. 0.5 μg
   B. 5 μg
   C. 50 μg
   D. 0.5 mg
   E. 0.5 g

3. Which important bee allergens are missing or found only in small concentrations in some therapeutic products used for venom immunotherapy (VIT)?
   A. Api m 1
   B. Api m 2
   C. Api m 4
   D. Api m 6
   E. Api m 10

4. Concerning the duration of VIT with Hymenoptera venoms, which statement is correct?
   A. Once the maintenance dose of 100 μg can be given without any side effect, VIT can be discontinued.
   B. The recommended duration of VIT is 3–5 years.
   C. Most patients will need a lifelong VIT.
   D. Once the patient on VIT has tolerated a sting of the relevant insect, VIT can be stopped.
   E. The duration of VIT—analogue to pollen immunotherapy—is maximally 3 years.

44

Atopic and Contact Dermatitis

Mark Boguniewicz, Luz Fonacier, Donald Y.M. Leung

Atopic dermatitis (AD) and contact dermatitis (CD) are common inflammatory skin diseases.[1,2] The complex pathophysiology of AD involves both underlying skin barrier abnormalities and immune dysregulation. Its course is augmented by environmental influences, including stress, allergen exposure, and microbial infection. Patients with AD experiencing a relapsing course benefit from a proactive approach with skin emollients and antiinflammatory therapy (*e.g.,* intermittent use of a topical steroid or calcineurin inhibitor). The identification of new immune pathways associated with AD has resulted in the development of novel biologicals targeting patients with refractory AD.

CD is a skin disorder caused by contact with an exogenous substance that elicits an allergic and/or irritant response. Allergic contact dermatitis (ACD) accounts for 20%, whereas irritant contact dermatitis (ICD) accounts for 80% of all cases of CD. ACD is a delayed-type, T cell–mediated immune response consisting of an afferent limb and an efferent limb. Irritants cause direct activation of the innate immune system through hyperproduction of cytokines and chemokines. The management of ACD includes identification of the allergen, avoidance, pharmacological intervention, and prevention. This chapter reviews the clinical and mechanistic aspects of both AD and CD.

CLINICAL ASPECTS OF ATOPIC DERMATITIS

Epidemiology

Prevalence of AD has been documented in up to 24% of children, with no apparent difference seen between urban and rural districts or between the two sexes. Children from families with atopy have a significantly higher risk of developing AD. Data from an international study (ISAAC Phase 3) showed that prevalence of current eczema in children varies widely among countries, ranging from 0.9% to 22.5%, and in adolescents, the prevalence ranges from 0.2% to 24.6%.[3] These data point to the significance of AD being a global health problem in both developed and developing countries and suggest that the ultimate presentation of an atopic disease may depend on a complex interaction between environmental exposures and end-organ response in genetically predisposed individuals. Prevalence of adult eczema from a nationally representative sample found that the 1-year prevalence of eczema was 10.2% and that the 1-year prevalence of eczema with asthma and/or hay fever was 3.2%.[4]

Natural History

AD typically manifests in early childhood, with onset in the first year occurring in >50% of patients and before 5 years of age in approximately 90% of patients. Although AD can present at any age, other diseases need to be considered in adults with new-onset dermatitis, especially in those without a history of childhood eczema, asthma, or allergic rhinitis (see Table 44.1, Differential Diagnosis). Although in some children, AD may progress to milder forms of eczema, only in approximately one-third will the disease be eliminated altogether. Severity and atopic sensitization are major determinants of prognosis. A US registry of patients with mild-moderate AD showed that it was not until age 20 years that 50% had experienced at least one 6-month period free of symptoms and treatments.[5]

Clinical Features

AD has no pathognomonic skin lesions or unique laboratory parameters; the diagnosis is, therefore, based on the presence of major and associated clinical features (Table 44.2). The principal features include pruritus, a chronically relapsing course, typical morphology and distribution of skin lesions, and a history of atopic disease. Patients usually have dry xerotic skin, and those with mutations of the gene encoding the epidermal barrier protein filaggrin (*FLG*) typically have prominent scaling and hyperlinear palms. Acute AD is characterized by pruritic, erythematous papules associated with excoriations, vesiculations, and serous exudate. Subacute AD is characterized by erythematous, excoriated, scaling papules, whereas chronic AD is characterized by lichenification and fibrotic papules. During infancy, AD involves primarily the face, scalp, and the extensor aspects of the extremities; the diaper region is typically spared, although some infants manifest the flexural involvement typical of older patients (Fig. 44.1).

Complicating Features

Ocular Problems

Atopic keratoconjunctivitis is bilateral, with intense pruritus, burning, tearing, and copious mucoid discharge. It is frequently associated with eyelid dermatitis and chronic blepharitis and may result in visual impairment from corneal scarring. Patients may also develop keratoconus from persistent rubbing of the eyes or anterior subcapsular cataracts.

Hand Dermatitis

Patients with AD may have a nonspecific irritant hand dermatitis aggravated by repeated wetting, especially in an occupational setting.

Infections

Patients with AD have an increased susceptibility to infection or colonization with a variety of organisms.[2] *Staphylococcus aureus* can be cultured from the skin of the majority of patients with

TABLE 44.1 Differential Diagnosis of Atopic Dermatitis

Congenital Disorders

- Netherton syndrome
- Familial keratosis pilaris

Chronic Dermatoses

- Seborrheic dermatitis
- Contact dermatitis (allergic or irritant)
- Nummular eczema
- Psoriasis
- Ichthyoses

Infections and Infestations

- Scabies
- Human immunodeficiency virus (HIV)–associated dermatitis
- Dermatophytosis

Malignancies

- Cutaneous T-cell lymphoma (mycosis fungoides/Sézary syndrome)
- Letterer-Siwe disease

Autoimmune Disorders

- Dermatitis herpetiformis
- Pemphigus foliaceus
- Graft-versus-host disease (GvHD)
- Dermatomyositis

Immunodeficiencies

- Wiskott-Aldrich syndrome (WAS)
- Severe combined immunodeficiency syndrome (SCID)
- Hyperimmunoglobulin E (IgE) syndrome (HIES)
- Dedicator of cytokinesis 8 (DOCK8)–associated immunodeficiency
- Tyrosine kinase 2 (TYK2) deficiency
- Immune dysregulation, polyendocrinopathy, enteropathy X-linked (IPEX) syndrome

Metabolic Disorders

- Zinc deficiency
- Pyridoxine (vitamin B6) and niacin deficiency
- Multiple carboxylase deficiency
- Phenylketonuria

TABLE 44.2 Clinical Features of Atopic Dermatitis

Major Features

- Pruritus
- Facial and extensor involvement in infants and children
- Flexural lichenification in adults
- Chronic or relapsing dermatitis
- Personal or family history of atopic disease

Minor Features

- Xerosis
- Cutaneous infections
- Nonspecific dermatitis of the hands or feet
- Ichthyosis, palmar hyperlinearity, keratosis pilaris
- Pityriasis alba
- Nipple eczema
- White dermatographism and delayed blanch response
- Anterior subcapsular cataracts
- Elevated serum immunoglobulin E (IgE) levels
- Positive immediate-type allergy skin tests

Modified from Hanifin JM, Rajka G. Diagnostic features of atopic dermatitis. Acta Derm Venereol (Stockh) 1980;92:44–7.

AD. Preferential adherence of this organism in AD may be related to expression of adhesins, such as fibronectin and fibrinogen, in inflamed skin. Recurrent pustulosis has become a significant problem for a number of patients, especially with the emergence of methicillin-resistant *S. aureus* (MRSA) as an important pathogen in AD.

Viral infections in AD include herpes simplex, molluscum contagiosum, and human papillomavirus infections. Patients who have atopic dermatitis with eczema herpeticum (ADEH) have more severe disease, increased body surface area affected, and an increase in biomarkers, including circulating eosinophil counts, serum immunoglobulin E (IgE), thymus and activation-regulated chemokine, and cutaneous T cell–attracting chemokine, compared with patients with AD who do not have a history of EH. Patients with ADEH also have more cutaneous infections caused by *S. aureus* or molluscum contagiosum virus and are more likely to have a history of asthma, as well as food and inhalant allergies. Patients with ADEH have reduced interferon-γ (IFN-γ) production. IFNγ and IFNγR1 SNPs are significantly associated with ADEH and may contribute to an impaired immune response to herpes simplex virus. In addition, genetic variants in IFN-γ regulatory factor 2 (IFN-γR2) have been shown to be associated with ADEH and may contribute to abnormal immune responses to herpes simplex virus.

Superimposed dermatophytosis may contribute to cutaneous inflammation in AD. The opportunistic yeast *Malassezia sympodialis* (previously *Pityrosporum ovale*) has been associated with a predominantly head and neck distribution of AD.

Systemic Complications

A significant number of patients with AD do not outgrow their disease and can have systemic inflammation with increased risk for systemic diseases, including cardiovascular disease,[6] rheumatoid arthritis, and/or inflammatory bowel disease.[7]

Psychosocial Implications

Patients with AD frequently respond to stress or frustration with itching and scratching. Stimulation of the central nervous system (CNS) may intensify cutaneous vasomotor and sweat responses and contribute to the itch–scratch cycle. Scratching can be associated with significant secondary gain or with a strong habitual component. Severe disease can lead to problems with self-esteem and social interactions. Sleep abnormalities are common and can contribute to impaired quality of life of patients and family members, even when the skin disease appears to be in remission.

Differential Diagnosis

A number of diseases can be confused with AD (see Table 44.1). Immunodeficiency with eczematous rash includes immune dysregulation, polyendocrinopathy, enteropathy X-linked (IPEX) syndrome. IPEX results from mutations of *FOXP3,* a gene located on the X chromosome that encodes a DNA-binding protein required for development of regulatory T cells (Tregs). Patients may present with enteropathy, type 1 diabetes, thyroiditis, hemolytic anemia, and/or thrombocytopenia. Wiskott-Aldrich syndrome (WAS) is an X-linked recessive disorder characterized by an eczematous rash, associated with thrombocytopenia along with variable abnormalities in humoral and cellular immunity and severe bacterial infections. Hyper-IgE syndrome (HIES) with mutations in the signal transducer and activator of transcription 3 gene (*STAT3*) is an autosomal dominant multisystem disorder characterized by recurrent deep-seated bacterial infections, including cutaneous cold abscesses and pneumonias caused by *S. aureus.* Patients with mutations in the gene encoding dedicator of cytokinesis 8 protein (*DOCK8*) have an immunodeficiency with eczema, recurrent viral infections, including some with CNS involvement, and many have associated food allergies. Patients with tyrosine kinase 2 (TYK2) deficiency can also present with an eczematous rash with high serum IgE and recurrent cutaneous staphylococcal infections. Other diseases to consider in the differential diagnosis of AD include cutaneous T-cell lymphoma, especially in adults with no history of childhood eczema and without other atopic features, human immunodeficiency virus (HIV) infection, and CD (see section on ACD). Contact allergy can also complicate AD, especially in patients whose AD does not respond to or worsens with therapy.

PATHOGENESIS OF ATOPIC DERMATITIS

Genetics

The genetics of AD are complex with key roles played by skin barrier or epidermal differentiation genes and immune response or host defense genes.

KEY CONCEPTS

Pathogenesis of Atopic Dermatitis (AD)

- Patients with AD have abnormalities of skin barrier and immune dysregulation.
- Some patients with AD have decreased filaggrin in the skin based on mutations in the gene encoding filaggrin (*FLG*), copy number variation of *FLG,* or secondary to type 2–mediated cytokine (*e.g.,* IL-4, IL-13) suppression of *FLG.*
- *FLG* mutations are associated with early-onset, severe, persistent AD with increased risk for asthma and allergic sensitization.
- Normal-appearing skin in patients with AD is associated with immune activation and epidermal terminal differentiation abnormalities.
- Patients with AD are predisposed to *Staphylococcus aureus* colonization or infection, and a small subset are prone to eczema herpeticum.
- Patients with an “Asian phenotype” of AD have features of psoriasis with increased T-helper 17 (Th17) and Th22 along with increased Th2 skewing.

Loss-of-function mutations of *FLG* are a major predisposing factor for AD. Patients with *FLG* gene mutations have early-onset, severe, and persistent AD and are at increased risk for developing asthma, as well as food and inhalant allergies.[8] Other skin barrier proteins implicated in AD include loricrin and involucrin, which are both significantly decreased in lesional and nonlesional skin of patients with AD. Candidate gene approaches have also implicated variants in the *SPINK5* gene, and haplotype tagging of SNPs for the claudin-1 gene (*CLDN1*) has revealed associations with AD. These observations have established a key role for impaired skin barrier function in the pathogenesis of AD that allows increased transepidermal water loss and penetration of allergens, antigens, and chemicals from the environment, resulting in cutaneous inflammation. Additional genome-wide significant susceptibility loci identified through the genome-wide association study (GWAS) and immunochip analysis continue to be reported and suggest a role for epidermal barrier function, innate-adaptive immunity, interleukin-1 (IL-1) family signaling, Tregs, the vitamin D pathway, and the nerve growth factor pathway in the pathogenesis of AD.[9] Thus a combination of varied genetic factors may influence a wide range of phenotypes of AD among individuals.

Immune Abnormalities in Atopic Dermatitis

Immunohistology

On the basis of clinical appearance and duration of illness, AD skin can be characterized as nonlesional AD, acute AD lesions (≤3 days after onset), and chronic skin lesions. Nonlesional AD skin is not normal and is characterized by a sparse perivascular T-cell infiltrate, suggesting the presence of minimal inflammation at baseline. Langerhans cells (LCs) also exhibit surface-bound IgE molecules, which have enhanced capacity to present allergen to T cells. In acute lesions, there is epidermal spongiosis with an increased infiltration of activated memory T cells bearing the skin-homing cutaneous lymphocyte-associated antigen (CLA$^+$). Eosinophils, basophils, and neutrophils are rare. Mast cells are seen in various stages of degranulation.

Chronic lichenified AD lesions are characterized by epidermal hyperplasia, prominent hyperkeratosis, and an increased number of dendritic cells (DCs) bearing surface IgE. Macrophages dominate the dermal mononuclear cell infiltrate, but lymphocytes remain prominent. Although intact eosinophils are rarely seen, eosinophil product deposition can be readily identified, suggesting they contribute to allergic skin inflammation.

Immune Pathways in Atopic Dermatitis

AD is associated with a combination of defective innate responses to microbes and altered adaptive responses to environmental allergens (reviewed in references[1,2,10]). A key difference between epidermal keratinocytes found in AD skin, compared with normal skin, is the presence of thymic stromal lymphopoietin (TSLP) and IL-33 in the AD epidermis. TSLP and IL-33 are key cytokines secreted by epithelial cells that induce DCs to drive T-helper 0 (Th0) cells into the Th2-cell differentiation pathway. Nonlesional AD and acute AD skin lesions are predominantly associated with expression of IL-4, IL-5, IL-13, IL-25, and IL-33. These type 2 cytokines are present in all stages of AD and can be secreted by multiple cell types, including innate lymphoid type 2 cells (ILCs), mast cells, and basophils, which are present in AD skin lesions and contribute to substantial redundancy in allergic inflammation. As such, cytokine targeting, as opposed to cell targeting, is considered a more effective approach in the treatment of AD. It is noteworthy that experimental studies demonstrate pretreatment with IL-4 and IL-13 dampens responses to IFNs and IL-17, suggesting that once the early AD lesion is exposed to IL-4 and IL-13, there is a long-lasting persistent effect.

Besides Th2, other cytokine pathways are also activated during the evolution of AD. The IL-22–IL-17 pathway is of particular interest, since it, along with IL-4 and IL-13, can inhibit terminal keratinocyte differentiation, including filaggrin expression. Since DC-derived IL-23 enhances IL-22/IL-17 cell differentiation, all of these cytokines are being closely examined for their potential role in AD. It is interesting that IL-4 and Il-13 can enhance IL-23 production by DCs. Furthermore, blockade of the IL-4 and IL-13 pathway, leading to improvement in AD, is also associated with reduced IL-23 and IL-17 expression in AD skin. When acute AD skin lesions become chronic, there is an increase in Th1 cytokines, such as IFN-γ, which potentiates AD skin inflammation. A recent birth cohort study, however, demonstrated that TSLP could be detected in at-risk infants before onset of AD, suggesting that

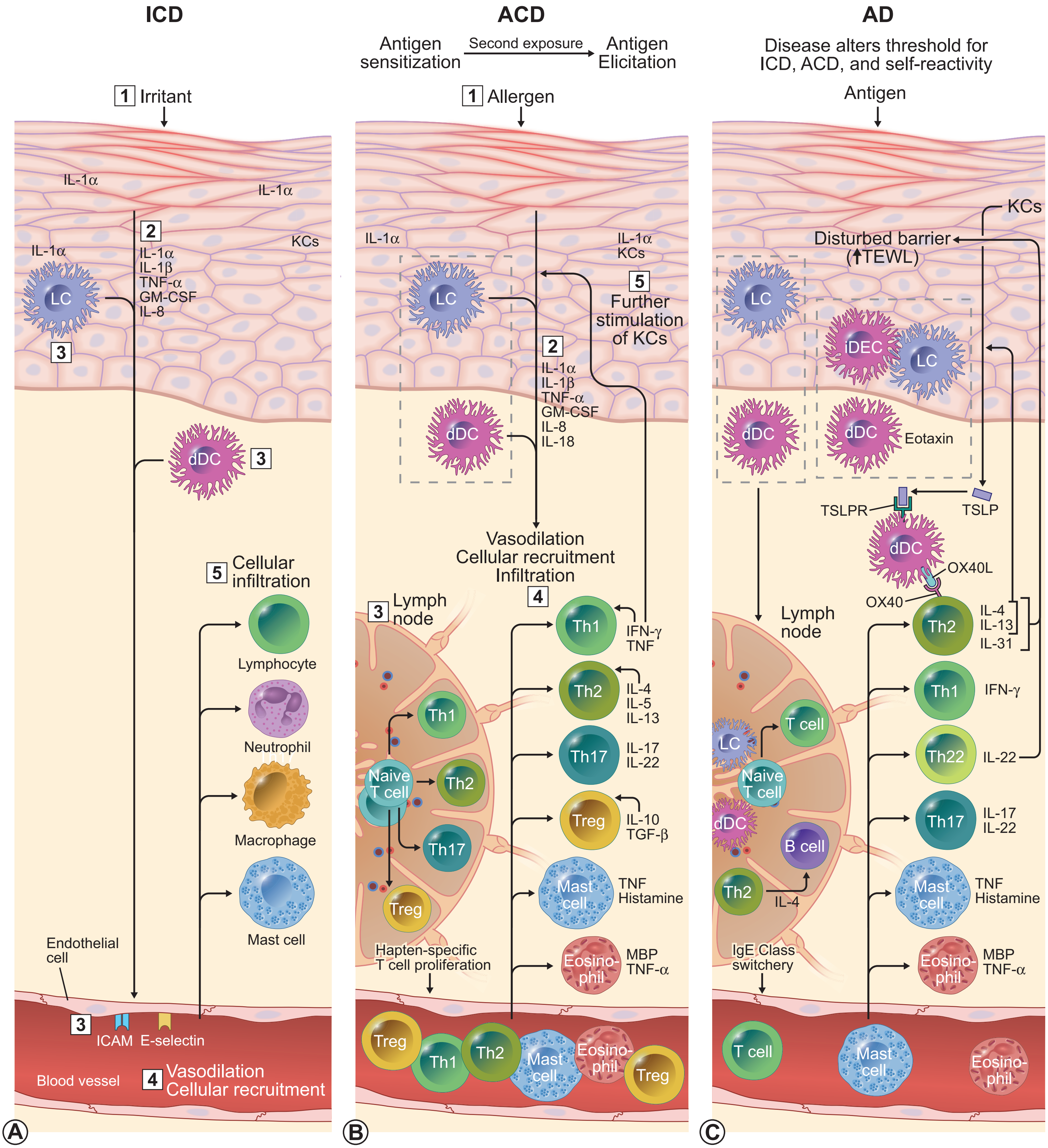

ICD
1 Irritant
IL-1α
IL-1α
KCs
IL-1α
2 IL-1α IL-1β TNF-α GM-CSF IL-8
LC
3
dDC
3
5 Cellular infiltration
Lymphocyte
Neutrophil
Macrophage
Mast cell
Endothelial cell
3
ICAM
E-selectin
Blood vessel
4 Vasodilation Cellular recruitment
A
ACD
Antigen sensitization
Second exposure
Antigen Elicitation
1 Allergen
IL-1α
IL-1α KCs
5 Further stimulation of KCs
LC
2 IL-1α IL-1β TNF-α GM-CSF IL-8 IL-18
dDC
Vasodilation Cellular recruitment Infiltration
4
3 Lymph node
Th1
Naive T cell
Th2
Th17
Treg
Hapten-specific T cell proliferation
Th1
IFN-γ TNF
Th2
IL-4 IL-5 IL-13
Th17
IL-17 IL-22
Treg
IL-10 TGF-β
Mast cell
TNF Histamine
Eosinophil
MBP TNF-α
Treg
Th1
Th2
Mast cell
Eosinophil
Treg
B
AD
Disease alters threshold for ICD, ACD, and self-reactivity
Antigen
KCs
Disturbed barrier (↑TEWL)
LC
iDEC
LC
dDC
dDC
Eotaxin
TSLPR
TSLP
dDC
OX40L
OX40
Lymph node
Th2
IL-4 IL-13 IL-31
Th1
IFN-γ
T cell
LC
Naive T cell
Th22
IL-22
dDC
B cell
Th17
IL-17 IL-22
Th2
IL-4
Mast cell
TNF Histamine
IgE Class switchery
Eosinophil
MBP TNF-α
T cell
Mast cell
Eosinophil
C

FIG 44.1 (A) In patients with irritant contact dermatitis (ICD), exposure to an irritant exerts toxic effects on keratinocytes (KCs), activating innate immunity with release of interleukin (IL)-1α, IL-1β, tumor necrosis factor (TNF)-α, granulocyte macrophage–colony-stimulating factor (GM-CSF), and IL-8 from epidermal keratinocytes. In turn, these cytokines activate Langerhans cells (LCs), dermal dendritic cells (DCs), and endothelial cells, all of which contribute to cellular recruitment to the site of keratinocyte damage. Infiltrating cells include neutrophils, lymphocytes, macrophages, and mast cells, which further promote an inflammatory cascade. (B) In the sensitization phase of allergic contact dermatitis (ACD), similar to ICD, allergens activate innate immunity through keratinocyte release of IL-1α, IL-1β, TNF-α, GM-CSF, IL-8, and IL-18, inducing vasodilation, cellular recruitment, and infiltration. LCs and dermal DCs encounter the allergen and migrate to the draining lymph nodes, where they activate hapten-specific T cells, which include T-helper cell-1 (Th1), Th2, Th17, and regulatory T cells (Tregs). These T cells proliferate and enter the circulation and site of initial exposure, along with mast cells and eosinophils (EOS). On re-encountering the allergen, the elicitation phase occurs, in which the hapten-specific T cells, along with other inflammatory cells, enter the site of exposure and, through release of cytokines and consequent stimulation of keratinocytes, induce an inflammatory cascade. (C) In patients with atopic dermatitis (AD), a disturbed epidermal barrier leads to increased permeation of antigens, which encounter LCs, inflammatory dendritic epidermal cells (IDECs), and dermal DCs, activating Th2 T cells to produce IL-4 and IL-13. DCs then travel to lymph nodes, where they activate effector T cells and induce immunoglobulin E (IgE) class-switching. IL-4 and IL-13 stimulate keratinocytes to produce thymic stromal lymphopoietin (TSLP). TSLP activates OX40 ligand–expressing dermal DCs to induce inflammatory Th2 T cells. Cytokines and chemokines, such as IL-4, IL-5, IL-13, eotaxins, CCL17, CCL18, and CCL22, produced by Th2 T cells and DCs stimulate skin infiltration by DCs, mast cells, and eosinophils. Th2 and Th22 T cells predominate in patients with AD, but Th1 and Th17 T cells also contribute to its pathogenesis. The Th2 and Th22 cytokines (IL-4/IL-13 and IL-22, respectively) have been shown to inhibit terminal differentiation and contribute to the barrier defect in patients with AD. Thus both the barrier defects and immune activation alter the threshold for ICD, ACD, and self-reactivity in patients with AD. MBP, Major basic protein. (Figure adapted from Gittler JK, Krueger JG, Guttman-Yassky E. Atopic dermatitis results in intrinsic and immune abnormalities: implications for contact dermatitis. J Allergy Clin Immunol 2013;131:300–13.)

the TSLP–Th2–ILC2 pathway plays a critical role in initiation of AD.[11]

Epidermal Barrier Dysfunction

The clinically dry skin and increased transepidermal water loss (TEWL) in AD reflects underlying skin barrier dysfunction and loss of natural moisturizing factor that plays an important role in the pathogenesis of AD. However, only a minority of patients have *FLG* null mutations. Other genetic variants in the epidermal differentiation complex (EDC) and tight junctions are even rarer. The majority of patients with AD likely have immune-mediated reduction in epidermal terminal differentiation, which leads to reduced generation of various epidermal structural proteins, filaggrin breakdown products, lipids, and antimicrobial peptides. TSLP, IL-4, and IL-13 are the most potent cytokines downregulating filaggrin expression by keratinocytes.[12] IL-17, IL-22, IL-25, and IL-33 can act synergistically with IL-4 and IL-13 to further downregulate expression of epidermal proteins and lipids. This combination of events, along with activation of proteases and lipases, creates defective epidermal barrier function, alters epidermal acidification, and leads to loss of moisturization in AD, thereby contributing to enhanced allergen and microbial penetration met by the host immune response and clinical appearance of AD.

MANAGEMENT OF ATOPIC DERMATITIS

Recent guidelines address the management of AD.[13,14] Recognition of skin barrier and immune abnormalities suggests the need for barrier repair and maintenance along with antiinflammatory measures.[15]

THERAPEUTIC PRINCIPLES

Atopic Dermatitis (AD)

- Fundamentals of skin care in AD include avoidance of irritants and proven allergens, along with proper skin hydration and use of moisturizers.
- Patients and caregivers have concerns about topical steroids and calcineurin inhibitors and tend to underuse prescribed medications.
- If eczema cannot be cleared and the patients keep relapsing, they may benefit from proactive intermittent antiinflammatory therapy.
- The diagnosis of AD should be reconsidered in patients not responding to conventional therapy, as the differential diagnosis includes a number of immunodeficiencies with eczematous rash.
- Systemic therapies should be reserved for patients with moderate-to-severe AD that does not respond to conventional therapy.

Identification and Elimination of Exacerbating Factors

Irritants

Patients with AD have a lowered threshold of irritant responsiveness. Recognition and avoidance of irritants is integral for successful management of this disease.

Allergens

Identification of clinically relevant allergens involves taking a careful history and doing selective testing, when appropriate. Negative skin tests with proper controls have a high predictive value for ruling out a suspected allergen. Positive skin tests have a lower correlation with clinical symptoms and may reflect sensitization.[16] Patients with severe AD who are avoiding foods based on *in vitro* testing results can tolerate many of the foods previously avoided when they are introduced under supervision. Environmental control measures aimed at reducing dust mite allergen can also improve AD in sensitized patients.

Psychosocial Factors

Recognizing and addressing sleep disturbance in patients is critical in managing a chronic, relapsing disease such as AD. Counseling together with relaxation, behavioral modification, and biofeedback may all be of benefit, especially for patients with habitual scratching.

Patient Education

Patients and caregivers need to be educated regarding the chronic relapsing nature of AD, its natural history, exacerbating factors, and treatment options. Recognizing that normal-appearing skin in patients with AD is, in fact, *not* normal is a difficult concept to understand but has important therapeutic implications. Patients should be counseled about prognosis and receive appropriate vocational counseling.

Hydration

The skin of patients with AD shows enhanced TEWL and lipid abnormalities that result in reduced water-binding capacity, higher TEWL, and decreased water content.[10] Skin hydration by soaking the affected area or bathing and applying an occlusive agent to retain absorbed water can help restore barrier function.[15] Bathing can also remove allergens, reduce *S. aureus* colonization, and act as relaxation therapy.

Moisturizers and Occlusives

Use of moisturizer or occlusive, especially when combined with hydration therapy, helps restore and preserve the skin barrier and can result in decreased need for topical corticosteroids. Twice-daily emollient application has been shown to improve barrier function and protect the skin from *S. aureus* proliferation while preserving microflora biodiversity.[17] Patients with AD have been shown to have a ceramide deficiency of the stratum corneum; barrier repair may be accelerated by increasing the ratio of ceramides, cholesterol, and either the essential fatty acid linoleic acid or the nonessential palmitic or stearic fatty acids.

Corticosteroids

Corticosteroids reduce inflammation and pruritus in acute and chronic AD, acting on multiple resident and infiltrating cells, primarily through suppression of inflammatory genes. Topical corticosteroids have been the mainstay of conventional therapy: when they are appropriately used, side effects are infrequent. Thinning of skin, telangiectasia, bruising, hypopigmentation, acne, striae, and secondary infections may occur. The face, particularly the eyelids, and the intertriginous areas are especially sensitive to these adverse effects. If topical corticosteroids are used on the face, this can lead to perioral dermatitis, characterized by erythema, scaling, and follicular papules and pustules around the mouth, in the alar creases, and sometimes on the upper lateral eyelids.

An important concept with translational applications is the recognition that nonlesional, normal-appearing skin in AD shows evidence of both immunological dysregulation and skin barrier abnormalities.[18] This observation provides a rationale for the use of topical corticosteroids in a proactive manner.[7] If the eczema can be cleared or almost cleared but has a relapsing course, long-term control can be maintained with twice-weekly applications of topical corticosteroid to areas that had previously been involved but now appear normal, with fewer relapses and less need for topical corticosteroids compared with treating eczema in a reactive manner.

In addition to their anti-inflammatory properties, topical corticosteroids can decrease *S. aureus* colonization in patients with AD. Failure to show clinical improvement with topical corticosteroids may be due to inadequate potency or amount of medication used, superinfection, steroid allergy, steroid resistance, or, more commonly, nonadherence with the treatment regimen, emphasizing the need for both education and alternative therapies.

Systemic corticosteroids, including oral prednisone, should be avoided in the management of a chronic disorder such as AD.[13,19] Improvement observed with systemic corticosteroids may be associated with flaring of AD after discontinuation. If a short course of oral corticosteroids is given, topical skin care should be intensified during the taper to suppress rebound flaring of AD.

Topical corticosteroid treatment in patients with AD was recently shown to result in improvements of the AD genomic signature. Cytokine levels (IL-12p40, IL-13, IL-22, CCL17, CCL18, peptidase inhibitor 3 [PI3]/elafin, and S100As) were consistently reduced, with corresponding improvements in epidermal disease markers (keratin 16 and loricrin) in lesional skin from responders. Even low-potency corticosteroids can affect a broad array of immune and barrier responses in patients with AD.

Topical Calcineurin Inhibitors

Tacrolimus ointment (0.03% and 0.1%) and pimecrolimus cream (1%) are nonsteroidal topical calcineurin inhibitors (TCIs) approved for the treatment of AD.[13,14] Both drugs have proven effective, with a good safety profile even when used over extended periods, including in infants treated with pimecrolimus. Treatment with TCIs is not associated with skin atrophy and may also be useful in treating patients with steroid insensitivity. A common side effect with TCIs is burning or stinging sensation of skin. Ongoing surveillance and recent reports have not shown a trend of increased frequency of viral infections or problems with response to childhood vaccinations. Although there is no evidence of a causal link to cancer and the use of TCIs, the US Food and Drug Administration (FDA) issued a "box warning" because of lack of long-term safety data. The labeling states that these drugs are recommended as second-line treatments and that their use in children under the age of 2 years is currently not recommended. However, a review of epidemiological and clinical data concluded that the published data did not demonstrate any causal relationship between TCI use and malignancy or lymphoma risk.

Studies of proactive treatment with tacrolimus ointment in adults and children have shown benefit.[7] A systematic review and meta-analysis of randomized clinical trials comparing TCIs with topical corticosteroids for AD found that both treatment classes showed similar rates of improvement and treatment success.[20]

Antiinfective Therapy

Systemic antibiotic therapy may be necessary when a secondary infection with *S. aureus* is present.[2] Recolonization after a course of anti-staphylococcal therapy occurs rapidly. Maintenance antibiotic therapy should be avoided because it can result in colonization by MRSA. Bleach baths with dilute sodium hypochlorite may reduce skin infections and improve eczema (based on limited data).[13,14]

Patients with disseminated EH require treatment with a systemic antiviral, such as acyclovir.[2] Recurrent cutaneous herpetic infections can be treated with prophylactic oral acyclovir.[21] Superficial dermatophytosis can be treated with topical or, rarely, with systemic antifungals.

Antipruritic Agents

Pruritus is the most common and least tolerated symptom in AD. Mediators other than histamine, including neuropeptides and cytokines, especially IL-31, can contribute to pruritus, and centrally acting agents, such as opioid receptor antagonists, have been shown to be effective against the itch of AD. Treatment with cyclosporine, which results in decreased transcription of a number of proinflammatory cytokines, leads to rapid improvement in pruritus in many AD patients. Monoclonal antibodies (mAbs) against IL-31 and its receptor are undergoing clinical trials in patients with AD,[22] and two phase three trials of a mAb against IL-4 receptor alpha, inhibiting signaling by IL-4 and IL-13 significantly reduced pruritus and enhanced quality of life.[22a] Systemic antihistamines and anxiolytics may be most useful through sedative effects and should be used primarily at bedtime to avoid daytime somnolence. Short-term use of a sedative to allow adequate rest may be appropriate. Treatment with topical antihistamines and topical anesthetics should be avoided because of the potential for allergic sensitization.

RECALCITRANT DISEASE

Hospitalization

Patients who experience erythroderma, appear to have toxicity, or fail on conventional therapies may require hospitalization.[23] Removing the patient from environmental allergens or stressors, intense education, and assurance of adherence with therapy usually result in marked clinical improvement. The patient can also undergo appropriately controlled provocative challenges to identify potential triggering factors.

Wet Wrap Therapy

Wet wrap therapy (WWT), in which a layer of wet clothing or bandages is placed over topical corticosteroid with a dry layer on top, can reduce inflammation and pruritus and acts as a barrier to trauma associated with scratching. WWT can aid epidermal barrier recovery that persists even after this therapy is discontinued, and clinical benefit can be demonstrated after discontinuation. Overuse can result in chilling, maceration of skin, or secondary infection. WWT should be employed for acute exacerbations of AD or for selective use in areas of resistant eczema, rather than as a maintenance treatment. The package inserts for topical calcineurin inhibitors recommend that they should not be used under occlusive dressings.

Systemic Immunosuppressive Agents

Oral cyclosporine has been shown to be effective in placebo-controlled studies in severe AD, although it is not approved for treatment of AD in most countries. A systematic review of randomized controlled trials evaluating systemic immunomodulating treatments for moderate-to-severe AD found that, of 12 different interventions studied in 34 randomized controlled trials, strong recommendations were only possible for the short-term use of cyclosporine.[24] Patients need to be appropriately monitored while on this medication.[19] Other systemic immunosuppressive drugs that have been used to treat severe AD include mycophenolate mofetil (a purine biosynthesis inhibitor), azathioprine (a systemic immunosuppressive agent affecting purine nucleotide synthesis and metabolism), and methotrexate (a folic acid antagonist that interferes with purine and pyrimidine synthesis).[13,19] All of these drugs require careful monitoring of patients for potential serious adverse effects.

Phototherapy and Photochemotherapy

Ultraviolet (UV) light therapy can be a useful treatment for chronic, recalcitrant AD.[19] Potential long-term adverse effects include premature skin aging and cutaneous malignancies. In patients with moderate-to-severe chronic AD treated with narrow-band UVB phototherapy, gene expression and immunohistochemistry studies of both lesional and nonlesional skin showed that Th2, Th22, and Th1 immune pathways were suppressed and that measures of epidermal hyperplasia and differentiation were normalized. Clinical improvement was associated with decrease in Th2/Th22–associated cytokines and chemokines and, importantly, normalized expression of epidermal barrier proteins. UVB phototherapy has also been shown to significantly decrease levels of toxin-producing *S. aureus* on the skin of pediatric patients with AD.

Allergen-Specific Immunotherapy

Allergen-specific immunotherapy (AIT; Chapter 91) has been administered to patients with AD for many years, although there is a paucity of controlled data on this treatment. Based on limited evidence, the updated Practice Parameter for AD states that AIT may be effective for a subset of patients with AD and aeroallergen sensitivity.[13] A Cochrane review of randomized controlled trials of AIT in AD found that the trials were confounded by high loss to follow-up and lack of blinding.[25] The authors were unable to determine by subgroup analyses a particular allergen, age, or level of disease severity where AIT was more successful.

Biologics and Investigational Therapies

Intravenous Gammaglobulin

Because chronic inflammation and T-cell activation appear to play a critical role in the pathogenesis of AD, high-dose intravenous immunoglobulin (IVIG; Chapter 84) could have immunomodulatory effects in this disease. IVIG could also interact directly with infectious organisms or toxins involved in the pathogenesis of AD. IVIG has been shown to contain high concentrations of staphylococcal toxin–specific antibodies that inhibit the *in vitro* activation of T cells by staphylococcal toxins. Treatment of severe refractory AD with IVIG has yielded conflicting results. Studies have not been controlled and have involved small numbers of patients. A systematic review of systemic therapies found that IVIG was not efficacious in the treatment of moderate-to-severe AD.[24]

Omalizumab

Treatment of patients with AD with omalizumab has mainly been reported in case reports and case series, showing both clinical improvement and lack of benefit. A placebo-controlled trial of omalizumab given for 16 weeks did not show significant clinical benefit. No specific markers have been identified that define responders, although a recent study suggested that adult patients with AD that responds to treatment have wild-type *FLG* mutations. In addition, patients receiving omalizumab have been shown to have decreased levels of TSLP, OX40L, thymus and activation-regulated chemokine (TARC), and IL-9 and marked increase in IL-10 compared with placebo.

Rituximab

Rituximab, a chimeric monoclonal anti-CD20 antibody has been given to patients with severe AD, with skin biopsies showing significant improvement in spongiosis and acanthosis and dermal T-cell and B-cell infiltrates decreased as well.[13] Although total serum IgE levels were reduced, allergen-specific IgE levels were not affected. Treatment combining omalizumab and rituximab in patients with severe refractory AD was found to be effective, suggesting that optimal therapy may require use of several biologics.

Dupilumab

Dupilumab is a fully human mAb directed against the IL-4 receptor α subunit that blocks signaling of both IL-4 and IL-13. Results from several randomized, double-blind, placebo-controlled trials involving adults with moderate-to-severe AD showed that treatment with dupilumab resulted in rapid and dose-dependent improvements in clinical indexes, biomarker levels, and disease transcriptome.[22a,26] Of the patients in the dupilumab group, 85% had a 50% reduction in the Eczema Area and Severity Index (EASI) score compared with 35% in the placebo group ($P<0.001$); 40% of patients in the dupilumab group had an investigator's global assessment score of 0–1 (clear or almost clear) compared with 7% in the placebo group ($P<0.001$); and pruritus scores decreased by 55.7% in the dupilumab group versus 15.1% in the placebo group ($P<0.001$). In a combination study, 100% of the patients in the dupilumab group had a 50% improvement in EASI ($P = 0.002$), compared with 50% of those in the placebo group, even though the patients who received dupilumab used less than half the amount of topical steroids compared with those in the placebo group ($P = 0.16$). Adverse events, such as skin infection, occurred more frequently with placebo; nasopharyngitis and headache were the most frequent adverse events with dupilumab. Of note, treatment with dupilumab monotherapy shifted the RNA expression profile of lesional skin to a more nonlesional signature. Dupilumab was well tolerated with no dose-limiting toxicity. Further dose ranging studies have been reported,[27] and studies in adolescents and children are ongoing.

Anti–IL-12/IL-23

Ustekinumab, an mAb used for treatment of psoriasis and psoriatic arthritis, binds to the common p40 subunit of IL-12/IL-23. A recent study of Japanese and Korean patients with AD found overlapping features of both AD and psoriasis.[22] The Asian AD phenotype described presents a blended phenotype, including increased hyperplasia, parakeratosis, higher Th17 activation, and a strong Th2 component. Although these findings need to be reproduced in a larger population of Asian patients, they do suggest that ustekinumab may be of benefit in a subset of patients with similar features.

Other Biologics

A number of other biologics have been evaluated in preliminary studies or are in early development in AD.[22,28] These include anti–IL-13 (lebrikizumab, tralokinumab), anti–IL-22, anti–IL-31, and anti–IL-31-receptor, as well as anti-TSLP and anti–TSLP receptor. New insights into AD phenotypes as well as better understanding of underlying immunopathology will likely lead to improved selection of appropriate patients for more targeted therapies.

Recombinant Human Interferon-γ

Interferon-γ suppresses IgE synthesis and inhibits Th2 cell function. In patients with AD, treatment with subcutaneous recombinant human interferon-γ (rhIFN-γ) resulted in reduced clinical severity and decreased total circulating eosinophil counts.[13] rhIFN-γ may act primarily on the allergic inflammatory response as opposed to IgE synthesis, and a subset of patients treated with rhIFN-γ could respond to individualized titration of their treatment dose. In a case series of pediatric patients with ADEH, IFN-γ and IVIG were thought to be less likely to enhance the cutaneous viral susceptibility of these patients compared with systemic immunosuppressive therapies.[21]

Phosphodiesterase-4 Inhibitors

Apremilast, an oral phosphodiesterase-4 inhibitor approved for the treatment of psoriasis, was shown in a pilot study in adults with AD to reduce pruritus and improve the Dermatology Quality of Life Index (DLQI) and EASI scores.[29] Gene ontological analyses comparing baseline samples with samples obtained during treatment revealed alterations in immune response pathways, especially those related to cyclic adenosine monophosphate (cAMP)–mediated signaling.

Crisaborole is a benzoxaborole phosphodiesterase 4 (PDE4) inhibitor, which has been approved by the FDA as a 2% topical ointment in children and adults with mild to moderate AD.[30] The unique configuration of boron within the crisaborole molecule enables selective targeting and inhibition of PDE4, increasing cAMP levels and reducing inflammation. In addition, the boron molecule allowed for synthesis of a low-molecular-weight compound (251 daltons), thereby facilitating penetration of crisaborole through human skin. *In vitro* experiments have shown that crisaborole inhibits cytokine production by peripheral blood mononuclear cells similar to other PDE4 inhibitors and distinct from corticosteroids. Randomized controlled clinical trials have shown the efficacy and safety of crisaborole topical ointment 2% in children, adolescents, and adults with mild to moderate AD.

Probiotics

Clinical trials in patients with AD show varying results. A meta-analysis of randomized controlled trials that attempted to overcome some of the limitations of prior reviews found a reduction of approximately 20% in the incidence of AD and IgE-associated AD in infants and children with the use of probiotics. Although these results are encouraging, the use of probiotics for the prevention of AD remains investigational.

PREVENTION

Two preliminary studies have suggested that emollient therapy started in early infancy in at-risk neonates can reduce the incidence of AD.[31,32] Both studies were of relatively short-term duration without follow-up off emollient therapy. It is noteworthy that infants at risk for AD at 1 year of age could be identified as early as day 2 of life by TEWL measurement, independent of parental atopy or *FLG* mutation status.[33]

CONTACT DERMATITIS

CD is a skin disorder caused by contact with an exogenous substance that elicits an allergic and/or irritant response. The vast majority of cases of contact-induced skin reactions are attributable to CD. However, there are other less-well-defined contact reactions, including contact urticaria, contact urticaria syndrome, and protein contact dermatitis. ACD affects approximately 7% of the general population, 13.3–24.5% of pediatric patients,[34] and 33–64% of the elderly population. CD was associated with more than 10.6 million physician office visits, with annual direct costs estimated at \$1.6 billion and indirect costs at \$566 million as a result of loss of work time, school time, and productivity.[35] Numerous studies have reported an increased frequency of ACD in patients with AD,[36] likely as a result of the increased exposure to products and chemicals used to treat AD, the barrier defect, and immunological changes in AD, which predispose a patient to both ICD and ACD.[37]

PATHOGENESIS OF ALLERGIC CONTACT DERMATITIS

The Genes

ACD is a multifactorial condition in which genetic background plays an important part, as shown by twin and family studies. An association has been demonstrated between loss of function mutations R501x and 2282del4 in the filaggrin (*FLG*) gene and contact sensitization against nickel II-sulfate, combined with an intolerance to fashion jewelry but not with other contact allergens.[38] *FLG* mutations have further been shown to lower the age of onset of nickel sensitization.[39] Thus FLG defects may be a risk factor for contact sensitization to allergens.

The Allergens

Most contact allergens are *haptens,* that is, simple chemicals that bind to carrier proteins present in skin to form a complete antigen (Chapter 6). To be allergenic, the chemical must be able to penetrate the principal barrier in skin (stratum corneum) and reach the living cells of the epidermis. Only molecules with molecular mass of <500 daltons (Da) are capable of penetrating the stratum corneum. Lipid solubility promotes transit through the stratum corneum. Thus most contact allergens are small, lipophilic molecules. Once in the epidermis, the nature of the protein carrier for the hapten is very important because if the contact sensitizer is complexed to nonimmunogenic carriers, this may induce tolerance rather than sensitization.[38]

The Immune Response

The immune response of ACD requires completion of both an afferent (sensitization) and an efferent (elicitation) phase.

In the afferent limb, the hapten enters the epidermis and activates keratinocytes to release inflammatory cytokines and chemokines, including tumor necrosis factor (TNF)-α, granulocyte macrophage–colony-stimulating factor (GM-CSF), IL-1α, IL-1β, IL-8, IL-10, IL18, and macrophage inflammatory protein-2 (MIP-2). The activation of LCs, other DCs, and endothelial cells can lead to recruitment of even more DCs at the site of antigen contact. The release of IL-1β by epidermal LCs promotes their egress from the epidermis. LCs process the antigen while migrating to regional lymph nodes, where they present it to naïve T cells. This phase is influenced by multiple factors. Defects in the integrity of the stratum corneum allow greater penetration of allergen and increase the chances of activating antigen-presenting cells (APCs) in skin. The availability and viability of APCs in skin, as well as the presence or absence of cytokines produced by keratinocytes, can promote or hinder APC–T-cell engagement.

In the draining lymph nodes, LCs present the peptides to T cells and activate CD4 and CD8 antigen-specific T cells. An important property of LCs and DCs is their ability to present exogenous antigens on both major histocompatibility complex (MHC) class I and class II molecules. This cross-priming leads to the activation of both CD4 and CD8 hapten-specific T cells. Although classic delayed-type hypersensitivity reactions are mediated primarily by CD4 cells, CD to haptens is mediated primarily by CD8 cells with a Th1-type cytokine profile. LCs activate hapten-specific T cells, which include Th1, Th2, Th17, and Treg subsets.[40] This preexisting mix of T-cell subtypes specific for the antigen influence the outcome of this process. The higher the frequency of cells of an effector subtype, the higher the likelihood that dermatitis will result, whereas a higher frequency of cells of a regulatory subtype may limit or prevent the development of dermatitis.

The efferent phase of ACD occurs on subsequent contact of skin with the hapten. Antigen-specific memory T cells and other inflammatory cells leave vessels and enter skin through sequential activation of a number of adhesion molecules by cytokines. Memory T cells constitutively express CLA. E-selectin, the ligand for CLA, is induced on vascular endothelium by inflammatory mediators, such as IL-1 and TNF-α. This interaction causes memory T cells to slow down and roll along the endothelial surface as a prelude to migration to sites of inflammation. Firm adhesion and migration of leukocytes to the endothelium are mediated by T-cell very late antigen-4 (VLA-4)/leukocyte function–associated antigen-1 (LFA-1) and endothelial cell vascular cell adhesion molecule-1 (VCAM-1)/intercellular adhesion molecule 1 (ICAM-1), respectively (Chapter 11). Subsequently, LFA-1β T cells migrate toward ICAM-1β epidermal cells.

Mast cells can also participate in the elicitation phase. In mice, mast cells can be activated by a T-cell–derived antigen-binding factor that induces release of serotonin, producing swelling 2 hours after challenge. In addition, mast cells contain preformed TNF-α, which may regulate the adhesion molecules involved in the early recruitment of Th cells. The net result is an influx of sensitized T cells homing to the hapten-provoked skin site, releasing their inflammatory mediators, IL-2 and IFN-γ, and this results in an enhanced immune response through activation and recruitment of more inflammatory cells, producing spongiosis and the inflammatory dermal infiltrate characteristic of ACD.

Although Th1 cells have been classically considered the primary effector cells of ACD (responses to haptens, such as nickel, were dominated by IFN-γ–producing cells), recent studies have

KEY CONCEPTS

Pathogenesis of Contact Dermatitis

- Allergic contact dermatitis (ACD) is a delayed-type, T cell–mediated response with an afferent limb or sensitization phase and an efferent or elicitation phase.
- Irritant contact dermatitis (ICD) is caused by irritants exerting toxic effects on keratinocytes, causing a direct activation of the innate immune system through hyperproduction of cytokines and chemokines and inducing an inflammatory skin reaction.
- Patients with ICD are more susceptible to the development of ACD, indicating that the activation of innate immunity by irritants likely reduces the threshold for development of ACD.
- There is an increased frequency of ACD in patients with AD likely because of the disturbed skin barrier allowing increased allergen penetration on an already amplified adaptive response in AD.
- Although ACD, ICD, and AD are independently defined diseases, they frequently interact and coexist.

indicated that Th2 cells also participate in the development of contact hypersensitivity.[41]

Recently, both murine models and human studies have suggested the potential role of Th17 cells in the immunopathogenesis of ACD. Increased levels of cytokines favoring a Th17 response (*i.e.,* IL-6)[42] and allergen-specific IL-17–producing T cells were found in sensitized mice. In addition, IL-17–deficient mice showed decreased secretion of cytokines and chemokines, diminished hapten-specific CD4 T-cell responses, and reduced ear swelling. In patients with ACD, Th17-associated mediators, such as IL-17A, IL-17F, IL-22, IL-23, chemokine receptor 6 (CCR6), IL-22 receptor, and the Th17 transcription factor retinoic acid–related orphan receptor γ, were shown to be produced by nickel-specific T cells and were upregulated in ACD lesional skin and positive patch test biopsy specimens.[43] Nickel exposure was reported to induce production of IL-23 by keratinocytes, promoting a Th17-mediated response, as detected by the presence of IL-17–producing T cells in peripheral blood from patients with nickel allergy.[44] The role of IL-17 in ACD lesions includes induction of keratinocyte release of cytokines and chemokines (*i.e.,* IL-8 and IL-6) and promotion of T cell–induced apoptosis of keratinocytes. The significance of IL-17 in ACD is emphasized by the correlation between the increase in IL-17–producing cells with the clinical manifestations of ACD, as well as by their significant percentage (20%) among skin-infiltrating CD4 and CD8 T cells in contrast to their minor representation (only 1.5%) in the regional lymph nodes of allergen-primed mice.[43]

A particular role for IL-22 (produced by Th22 and Th17 T cells)[45] has been suggested in ACD. Significantly higher levels of IL-22 were detected in the serum of patients with ACD to nickel compared with those seen in control subjects. IL-22 cytokine levels have also been shown to be upregulated in inflamed skin of patients challenged with nickel allergy,[44] although its specific contribution to ACD reactions is not yet known.

The skin of patients with AD is at increased risk for the development of ACD, and this may be attributed to multiple factors. First, AD skin is exposed to chemicals used to treat AD, including moisturizers, topical corticosteroids, and TCIs. Second, the disturbed barrier system allows increased permeation of allergens. Third, AD skin has a heightened immunological status with existing activation of innate immunity, increased access of surface antigens to LCs, and selective upregulation of the Th2 adaptive immune response. In patients with AD, cutaneous contact with irritants and allergens leads to amplification of innate immunity and enhanced adaptive immune responses, including Th2 and Th17 in patients with acute AD and Th22 and Th1 in patients with chronic disease. Just as innate immune activation stimulated by an irritant permits a lower threshold of ACD elicitation, the amplified adaptive responses in lesional and nonlesional skin promote increased expression of ACD and ICD in patients with AD.

PATHOGENESIS OF IRRITANT CONTACT DERMATITIS

ICD accounts for 80% of cases of CD. The clinical presentation of ICD is usually restricted to skin that is directly in contact with the offending agent, with little or no extension beyond the site of contact. The inflammatory response is dose- and time-dependent. Any impairment to the epidermal barrier layer (*e.g.,* fissuring, overhydration) renders skin more susceptible to an irritant effect. Contact with agents, such as detergents, solvents, alcohol, creams, lotions, ointments and powders, and environmental factors, such as wetting, drying, perspiration, and temperature extremes, can abrade or irritate skin. After exposure to skin irritants, perturbation of the skin barrier and disorganization of the lipid bilayers with lamellar body lipid extrusion in the epidermis[46] is seen, with an associated increase in TEWL.

Irritants cause direct activation of the innate immune system through hyperproduction of cytokines and chemokines, such as IL-1α, IL-1β, IL-6, IL-8, and TNF-α, which further induce a cytokine cascade and inflammatory reaction with infiltration of inflammatory cells. Epidermal keratinocytes have been identified as key effector cells in the initiation and propagation of contact irritancy. Keratinocytes can release both preformed and newly synthesized cytokines, and can upregulate MHC class II molecules and induce adhesion molecules in response to irritants.[40] These mediators can cause direct tissue damage, activating LCs, dermal DCs, and endothelial cells, which contribute to further cellular recruitment, including neutrophils, lymphocytes, macrophages, and mast cells, which also contribute to the inflammatory cascade. The "final" cellular damage results from inflammatory mediators released by activated nonsensitized T cells.

Application of sodium lauryl sulfate (SLS) exerts direct toxic effects on keratinocytes (an experimental model of ICD) and has been demonstrated to induce LC mobilization and consequent migration to the draining lymph nodes. Other irritants have been shown to cause direct activation of the innate immune system through a set of membranous and intercellular receptors called Toll-like receptor 7 (TLR-7) and nucleotide-binding oligomerization domain (NOD)–like receptors (NLRs) that activate the inflammasome and nuclear factor-κB (NF-κB) pathways, inducing release of many cytokines and chemokines.

Although few immunohistopathological differences have been noted between ICD and ACD, ICD does not require prior sensitization, and immunological memory is not involved. The cellular infiltrate includes CD4 T cells with a Th1-type profile. Irritant reactions predispose to allergic reactions, making patients with ICD more susceptible to the development of ACD. Animal models in which skin was preirritated with 5% SLS showed an increase in skin sensitization rate to paraphenylenediamine from 38% to 78%[47] and enhanced T-cell proliferative responses to 2,4-dinitrochlorobenzene (DNCB). The DNCB-primed ICD reaction conditioned the development and severity of the ACD response. Similarly, in human studies, when patch test sites were pretreated with SLS, the threshold elicitation concentrations of contact allergens, such as cobalt, nickel, and colophony, were significantly decreased, indicating that the activation of innate immunity by irritants likely reduces the threshold for the development of ACD.

In summary, although ACD, ICD, and AD are independently defined diseases, they interact and frequently occur in combination. In patients with AD, if the activation of an innate immune response is preceded by exposure to irritants and allergens, there is altered permeation of contact allergens, which are introduced into an already activated innate immune system, resulting in amplification of innate immunity and enhanced adaptive immune responses. The amplified adaptive responses in AD promote increased ACD and ICD. Conversely, the innate immune activation stimulated by an irritant permits a lower threshold of ACD elicitation (Fig. 44.2).

CLINICAL MANIFESTATIONS OF CONTACT DERMATITIS

The clinical and histological findings of ACD are characteristic, but not diagnostic. ACD should be suspected in patients with

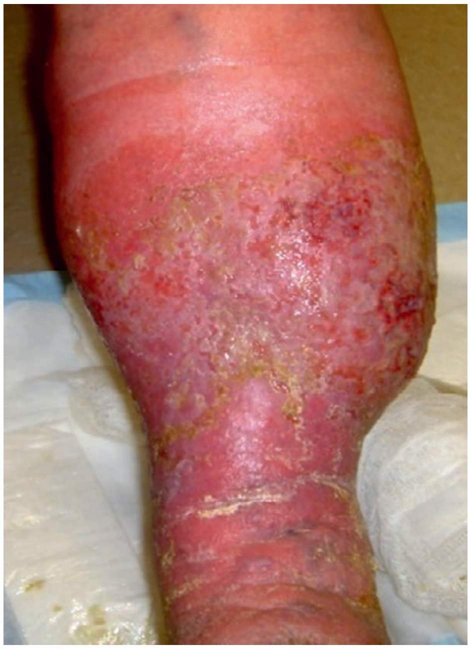


FIG 44.2 Acute allergic contact dermatitis.

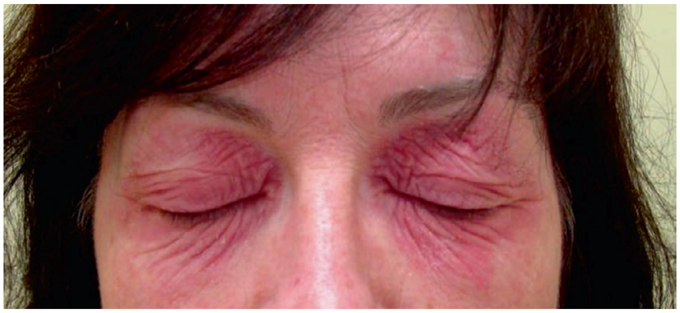


FIG 44.3 Allergic contact dermatitis of the eyelid.

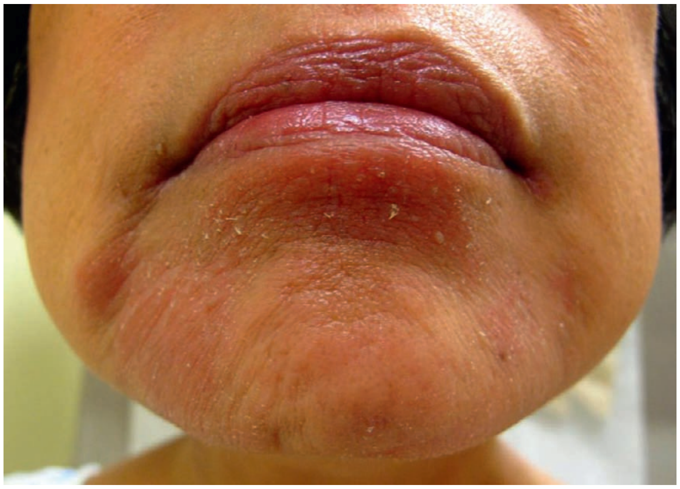


FIG 44.4 Allergic contact dermatitis of the face caused by fragrance.

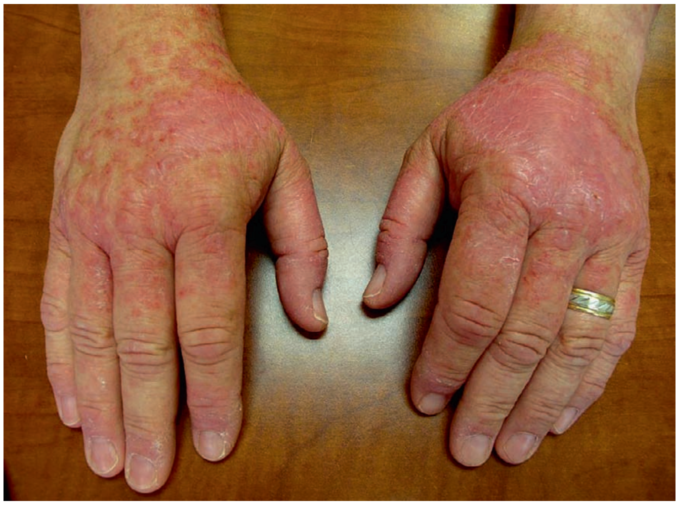


FIG 44.5 Allergic contact dermatitis of the hand caused by rubber accelerators.

acute and chronic eczematous or noneczematous dermatitis and is diagnosed on the basis of the clinical appearance of the lesions, the distribution of the dermatitis, the presence of pruritus, and the absence of other etiologies. Acute CD is characterized by erythematous papules, vesicles, oozing, and crusted lesions (Fig. 44.3). Recurrence and persistence of the dermatitis may lead to subacute and chronic lesions. Subacute CD manifests as erythema, scaling, fissuring, or a parched, scalded appearance and chronic inflammation may have more skin thickening, hardening, scaling, fissuring, and lichenification.

Although the location of the dermatitis serves as an important clue to the source of the offending chemical, multiple factors contribute to the distribution of ACD. Spread from the principal site of exposure can involve distant sites, either by inadvertent contact or by autosensitization. Areas of the scalp, palms, and soles have thicker skin, whereas the eyelid, face, and genital areas have thinner skin that is more sensitive to contact allergens (Figs. 44.4 and 44.5). A geographical approach can be very helpful in identifying the causal allergen, but dermatitis with scattered generalized distribution, which lacks the characteristic distribution that gives a clue as to the possible etiology of ACD, is actually the most common pattern of dermatitis in both children and adults, as reported by the North American Contact Dermatitis Group in 2013, followed by hands and then the face.[48] Systemic contact dermatitis (SCD), specifically the "baboon syndrome," is a diffuse eruption involving flexural and intertriginous areas following oral, intravenous, or transcutaneous exposure to the allergen in a contact-sensitized individual. The most common causes of SCD are *(i)* metals, such as mercury, nickel, and gold; *(ii)* medications, including aminoglycoside antibacterials, topical corticosteroids, and aminophylline; and *(iii)* plants and herbal products, including Compositae and Anacardiaceae families and Balsam of Peru (also known as *Myroxylon pereirae* resin).

Histologically, CD demonstrates intercellular edema of the epidermis (spongiosis) with varying degrees of acanthosis (thickening of the epidermal stratum basale and stratum spinosum) and superficial perivascular, lymphohistiocytic infiltrates. It is often difficult to distinguish ACD (Fig. 44.5) from ICD (Fig. 44.6) on the basis of physical examination or histological findings.

MANAGEMENT OF ALLERGIC CONTACT DERMATITIS

The management of ACD includes identification of the allergen, avoidance, pharmacological intervention, and prevention.

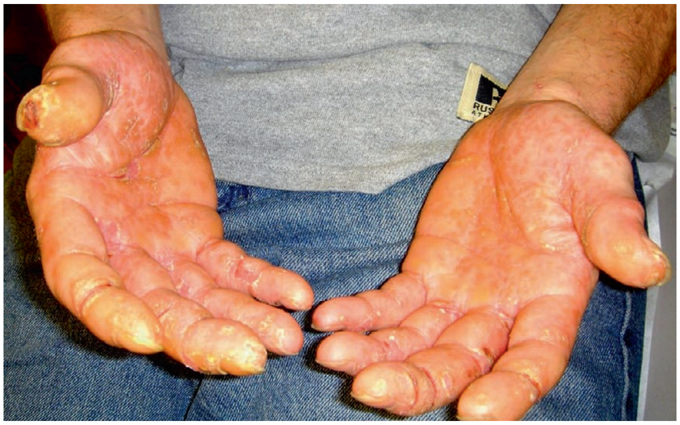


FIG 44.6 Irritant contact dermatitis.

THERAPEUTIC PRINCIPLES

Contact Dermatitis

- Patch testing is the procedure of choice to confirm the diagnosis of allergic contact dermatitis (ACD) and to identify offending contact allergens.
- The interpretation of patch tests requires both experience and judgment on its relevance.
- Once identified, the key to management of ACD is prevention by avoiding substances containing the allergens or irritants that have been identified.
- Medical treatments such as topical and/or systemic corticosteroids can be used to relieve ongoing dermatitis.

Identification of the Allergen

The most widely acceptable and available method for confirming the diagnosis of ACD is patch testing. Nickel remains the most common contact sensitizer and is more common in women than in men, likely because of greater exposure to nickel in jewelry and body piercing practices. In cosmetic products, fragrances, preservatives, and emulsifiers are the most common causative allergens. In addition, paraphenylenediamine (in hair dye), cocamidopropyl betaine (in shampoos and soaps), and medications (*e.g.*, neomycin, benzocaine, corticosteroids) are commonly reported allergens in personal hygiene and medical products. The results of patch tests must be interpreted in the context of the patient's experience; exposure and relevance should be established.[49]

Allergen Avoidance

Once the allergen or irritant has been identified, avoiding further contact with the offending agent would provide improvement of the dermatitis. This requires extensive patient education on what they are allergic to, including synonyms and cross-reactivity, products they must avoid, and, most importantly, products that may be safe for them to use. The dimethylglyoxime test (nickel spot test) and the cobalt spot test (based on disodium-1-nitroso-2-naphthol-3,6-disulfonate) can be used to detect nickel or cobalt released from metal objects and dermal exposure, thus aiding in avoidance of contact in sensitized patients. Use of appropriate skin protection (*i.e.*, gloves) may help with avoidance of exposure, especially in the occupational setting.[50]

Symptomatic Therapy

In addition to avoidance of exposure, appropriate adjunct medical treatment can be prescribed. First-line therapy is with topical corticosteroids, and second-line treatment includes phototherapy, oral retinoids, and immunosuppressant agents. The age of the patient and the severity, location, and acuteness of the dermatitis affect the selection of topical corticosteroids, which may be sufficient for localized lesions. Patients with dermatitis that is acute, extensive (particularly if involving >10% of total body surface), or severe may benefit from systemic therapy.

Contact sensitization to the corticosteroid itself, the vehicle, or other ingredients in the topical corticosteroid should be suspected if symptoms worsen, initially improve but then worsen with continued treatment, or do not respond to treatment at all. Several topical T-cell selective inhibitors (tacrolimus and pimecrolimus) have been used successfully in the treatment of AD, but their efficacy in ACD or ICD has not been established.[51] Other treatments, including cyclosporine, azathioprine, oral retinoids, and phototherapy, have been used for steroid-resistant ACD, such as chronic hand dermatitis.

PERSPECTIVES IN ATOPIC DERMATITIS AND ALLERGIC CONTACT DERMATITIS

Looking into the future, as we define the immune pathways that cause AD, novel biologicals currently in clinical trials may provide more targeted therapy in AD.

ON THE HORIZON

- Novel biologicals currently in clinical trials may provide more targeted therapy in atopic dermatitis (AD).
- Characterization of unique clinical phenotypes of AD may lead to a better understanding of AD pathophysiology and a precision-medicine approach to treatment.
- Identification of "at risk" patients will facilitate preventative strategies.
- A large number of chemicals are in everyday usage: new, rare, and emerging allergens should be considered in the assessment of suspected ACD. Patch testing to these allergens has yet to be standardized and irritant and elicitation concentration to patch testing determined.
- Data about contact allergen sensitization in children with AD are limited and are continually expanding. Frequency and patterns of CD in children with AD and results of patch testing have yet to be determined.

Identification of the endotypes involved in the pathogenesis of unique clinical phenotypes of AD will be required for a precision-medicine approach to treatment. In the absence of a low-cost cure for AD, biomarkers to identify "at-risk" patients could facilitate preventative strategies.

In the area of ACD, a large number of chemicals are being evaluated for their role, and new, rare, and emerging allergens should be considered in the assessment of suspected ACD. Patch testing to these allergens has yet to be standardized; irritant and elicitation concentrations to patch testing need to be determined. Data about contact allergen sensitization in children with AD are limited but are continually expanding. Frequency and patterns of CD in children with AD and the results of patch testing have yet to be determined.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Leung DY, Guttman-Yassky E. Deciphering the complexities of atopic dermatitis: shifting paradigms in treatment approaches. J Allergy Clin Immunol 2014;134:769–79.
2. Boguniewicz M, Leung DY. Recent insights into atopic dermatitis and implications for management of infectious complications. J Allergy Clin Immunol 2010;125:4–13.
3. Odhiambo JA, Williams HC, Clayton TO, et al. Global variations in prevalence of eczema symptoms in children from ISAAC Phase Three. J Allergy Clin Immunol 2009;124:1251–8.e23.
4. Silverberg JI, Hanifin JM. Adult eczema prevalence and associations with asthma and other health and demographic factors: a US population-based study. J Allergy Clin Immunol 2013;132:1132–8.
5. Margolis JS, Abuabara K, Bilker W, et al. Persistence of mild to moderate atopic dermatitis. JAMA Dermatol 2014;150:593–600.
6. Ewald DA, Malajian D, Krueger JG, et al. Meta-analysis derived atopic dermatitis (MADAD) transcriptome defines a robust AD signature highlighting the involvement of atherosclerosis and lipid metabolism pathways. BMC Med Genomics 2015;8:60.
7. Schmitt J, Schwarz K, Baurecht H, et al. Atopic dermatitis is associated with an increased risk for rheumatoid arthritis and inflammatory bowel disease, and a decreased risk for type 1 diabetes. J Allergy Clin Immunol 2016;137:130–6.
8. Irvine AD, McLean WH, Leung DY. Filaggrin mutations associated with skin and allergic diseases. N Engl J Med 2011;365:1315–27.
9. Tamari M, Hirota T. Genome-wide association studies of atopic dermatitis. J Dermatol 2014;41:213–20.
10. Boguniewicz M, Leung DY. Atopic dermatitis: a disease of altered skin barrier and immune dysregulation. Immunol Rev 2011;242: 233–46.
11. Kim J, Kim BE, Lee J, et al. Epidermal thymic stromal lymphopoietin predicts the development of atopic dermatitis during infancy. J Allergy Clin Immunol 2016;137:1282–5.e4.
12. Cabanillas B, Novak N. Atopic dermatitis and filaggrin. Curr Opin Immunol 2016;42:1–8.
13. Schneider L, Tilles S, Lio P, et al. Atopic dermatitis: a practice parameter update 2012. J Allergy Clin Immunol 2013;131:295–9.e1–27.
14. Eichenfield LF, Tom WL, Berger TG, et al. Guidelines of care for the management of atopic dermatitis: section 2. Management and treatment of atopic dermatitis with topical therapies. J Am Acad Dermatol 2014;71:116–32.
15. Boguniewicz M, Leung DY. The ABC's of managing patients with severe atopic dermatitis. J Allergy Clin Immunol 2013;132: 511–2.e5.
16. Bergmann MM, Caubet JC, Boguniewicz M, et al. Evaluation of food allergy in patients with atopic dermatitis. J Allergy Clin Immunol Pract 2013;1:22–8.
17. Bianchi P, Theunis J, Casas C, et al. Effects of a new emollient-based treatment on skin microflora balance and barrier function in children with mild atopic dermatitis. Pediatr Dermatol 2016;33:165–71.
18. Suarez-Farinas M, Tintle SJ, Shemer A, et al. Nonlesional atopic dermatitis skin is characterized by broad terminal differentiation defects and variable immune abnormalities. J Allergy Clin Immunol 2011;127:954–64.e1–4.
19. Sidbury R, Davis DM, Cohen DE, et al. Guidelines of care for the management of atopic dermatitis: section 3. Management and treatment with phototherapy and systemic agents. J Am Acad Dermatol 2014;71:327–49.
20. Broeders JA, Ahmed Ali U, Fischer G. Systematic review and meta-analysis of randomized clinical trials (RCTs) comparing topical calcineurin inhibitors with topical corticosteroids for atopic dermatitis: a 15-year experience. J Am Acad Dermatol 2016;75:410–9.e3.
21. Frisch S, Siegfried EC. The clinical spectrum and therapeutic challenge of eczema herpeticum. Pediatr Dermatol 2011;28:46–52.
22. Noda S, Krueger JG, Guttman-Yassky E. The translational revolution and use of biologics in patients with inflammatory skin diseases. J Allergy Clin Immunol 2015;135:324–36.

22a. Simpson EL, Bieber T, Guttman-Yassky E, et al. Two phase 3 trials of dupilumab versus placebo in atopic dermatitis. N Engl J Med 2016;375:2335–48.

23. Boguniewicz M, Nicol N, Kelsay K, et al. A multidisciplinary approach to evaluation and treatment of atopic dermatitis. Semin Cutan Med Surg 2008;27:115–27.
24. Roekevisch E, Spuls PI, Kuester D, et al. Efficacy and safety of systemic treatments for moderate-to-severe atopic dermatitis: a systematic review. J Allergy Clin Immunol 2014;133:429–38.
25. Tam HH, Calderon MA, Manikam L, et al. Specific allergen immunotherapy for the treatment of atopic eczema: a systematic review. Allergy 2016;71:1345–56.
26. Beck LA, Thaci D, Hamilton JD, et al. Dupilumab treatment in adults with moderate-to-severe atopic dermatitis. N Engl J Med 2014;371: 130–9.
27. Thaci D, Simpson EL, Beck LA, et al. Efficacy and safety of dupilumab in adults with moderate-to-severe atopic dermatitis inadequately controlled by topical treatments: a randomised, placebo-controlled, dose-ranging phase 2b trial. Lancet 2016;387:40–52.
28. Boguniewicz M, Leung DY. Targeted therapy for allergic diseases: at the intersection of cutting-edge science and clinical practice. J Allergy Clin Immunol 2015;135:354–6.
29. Samrao A, Berry TM, Goreshi R, et al. A pilot study of an oral phosphodiesterase inhibitor (apremilast) for atopic dermatitis in adults. Arch Dermatol 2012;148:890–7.
30. Jarnagin K, Chanda S, Coronado D, et al. Crisaborole topical ointment, 2%: a nonsteroidal, topical, anti-inflammatory phosphodiesterase 4 inhibitor in clinical development for the treatment of atopic dermatitis. J Drugs Dermatol 2016;15:390–6.
31. Simpson EL, Chalmers JR, Hanifin JM, et al. Emollient enhancement of the skin barrier from birth offers effective atopic dermatitis prevention. J Allergy Clin Immunol 2014;134:818–23.
32. Horimukai K, Morita K, Narita M, et al. Application of moisturizer to neonates prevents development of atopic dermatitis. J Allergy Clin Immunol 2014;134:824–30.e6.
33. Kelleher M, Dunn-Galvin A, Hourihane JO, et al. Skin barrier dysfunction measured by transepidermal water loss at 2 days and 2 months predates and predicts atopic dermatitis at 1 year. J Allergy Clin Immunol 2015;135:930–5.e1.
34. Heine G, Schnuch A, Uter W, et al. Frequency of contact allergy in German children and adolescents patch tested between 1995 and 2002: results from the Information Network of Departments of Dermatology and the German Contact Dermatitis Research Group. Contact Dermatitis 2004;51:111–17.
35. The Lewin Group, Inc. The burden of skin diseases; 2004. Available at: www.sidnet.org/files/Burden%20of%20Skin%20Diseases%202004%20Final%20Sept%2005. [Accessed 27 May 2014].
36. Mailhol C, Lauwers-Cances V, Rance F, et al. Prevalence and risk factors for allergic contact dermatitis to topical treatment in atopic dermatitis: a study in 641 children. Allergy 2009;64:801–6.
37. Aquino M, Fonacier L. The role of contact dermatitis in patients with atopic dermatitis. J Allergy Clin Immunol Pract 2014;2:382–7.
38. Novak N, Baurecht H, Schafer T, et al. Loss-of-function mutations in the filaggrin gene and allergic contact sensitization to nickel. J Invest Dermatol 2008;128:1430–5.
39. Ross-Hansen K, Menne T, Johansen JD, et al. Nickel reactivity and filaggrin null mutations—evaluation of the filaggrin bypass theory in a general population. Contact Dermatitis 2011;64:24–31.
40. Gittler JK, Krueger JG, Guttman-Yassky E. Atopic dermatitis results in intrinsic barrier and immune abnormalities: implications for contact dermatitis. J Allergy Clin Immunol 2013;131:300–13.
41. Minang JT, Troye-Blomberg M, Lundeberg L, et al. Nickel elicits concomitant and correlated in vitro production of Th1-, Th2-type and

regulatory cytokines in subjects with contact allergy to nickel. Scand J Immunol 2005;62:289–96.

42. Watanabe H, Gehrke S, Contassot E, et al. Danger signaling through the inflammasome acts as a master switch between tolerance and sensitization. J Immunol 2008;180:5826–32.
43. Zhao Y, Balato A, Fishelevich R, et al. Th17/Tc17 infiltration and associated cytokine gene expression in elicitation phase of allergic contact dermatitis. Br J Dermatol 2009;161:1301–6.
44. Larsen JM, Bonefeld CM, Poulsen SS, et al. IL-23 and T(H)17-mediated inflammation in human allergic contact dermatitis. J Allergy Clin Immunol 2009;123:486–92.
45. Nograles KE, Zaba LC, Shemer A, et al. IL-22-producing “T22” T cells account for upregulated IL-22 in atopic dermatitis despite reduced IL-17-producing TH17 T cells. J Allergy Clin Immunol 2009;123:1244–52.e2.
46. de Jongh CM, Jakasa I, Verberk MM, et al. Variation in barrier impairment and inflammation of human skin as determined by sodium lauryl sulphate penetration rate. Br J Dermatol 2006;154:651–7.
47. Magnusson B, Koch G, Nyquist G. Contact allergy to medicaments and materials used in dentistry. I. General principles and diagnostic methods in contact allergy. Identification of contact allergens by animal testing. Odontol Revy 1970;21:287–99.
48. Warshaw EM, Belsito DV, Taylor JS, et al. North American Contact Dermatitis Group patch test results: 2009 to 2010. Dermatitis 2013;24:50–9.
49. Fonacier L. A practical guide to patch testing. J Allergy Clin Immunol Pract 2015;3:669–75.
50. Fonacier L, Bernstein DI, Pacheco K, et al. Contact dermatitis: a practice parameter—Update 2015. J Allergy Clin Immunol Pract 2015;3:S1–39.
51. Belsito D, Wilson DC, Warshaw E, et al. A prospective randomized clinical trial of 0.1% tacrolimus ointment in a model of chronic allergic contact dermatitis. J Am Acad Dermatol 2006;55:40–6.

MULTIPLE-CHOICE QUESTIONS

1. Immunodeficiencies with eczematous rash that may be misdiagnosed as atopic dermatitis include all of the following EXCEPT:
   A. Wiskott-Aldrich syndrome (WAS)
   B. Properdin deficiency
   C. Immunodeficiency with mutations in dedicator of cytokinesis 8 (*DOCK8*)
   D. Immune dysregulation, polyendocrinopathy, enteropathy X-linked (IPEX) syndrome

2. The TRUE statement regarding filaggrin is:
   A. Mutations in the gene encoding filaggrin (*FLG*) occur only in patients of European Caucasian descent.
   B. Filaggrin is an enzyme that regulates breakdown of the epidermal barrier.
   C. T-helper 2 (Th2) cytokines interleukin-4 (IL-4) and IL-13 can downregulate *FLG* gene expression.
   D. Abnormal filaggrin expression in bronchial epithelium leads to increased risk of asthma.

3. Which of the following is a TRUE statement regarding atopic dermatitis?
   A. Downregulation of a number of antimicrobial peptide genes caused by local upregulation of Th2-type cytokines could explain the increased susceptibility of patients with atopic dermatitis (AD) to cutaneous infections.
   B. Serum levels of thymus and activation-regulated chemokine (TARC) rise with successful treatment.
   C. Patients whose AD has been in remission for >5 years are no longer at increased risk for eczema vaccinatum from smallpox vaccine.
   D. Langerhans cells are the primary source of antimicrobial peptides in skin.

4. Which of the following is the MOST common contact allergen?
   A. Neomycin
   B. Nickel
   C. Paraphenylenediamine
   D. Cocamidopropyl betaine

45

Food Allergy

J. Andrew Bird, Stacie Jones, Wesley Burks

Adverse reactions to foods may be mediated by immunological mechanisms (food allergy) and nonimmunological mechanisms (food intolerance). Immune-mediated reactions to foods most commonly involve immunoglobulin E (IgE), an antibody that binds to mast cells and basophils. Upon exposure to allergen, they release inflammatory mediators, such as histamine, prostaglandins, and leukotrienes, resulting in symptoms ranging from localized oral itching to anaphylaxis, a potentially fatal systemic reaction. Non–IgE-mediated immune reactions to foods include such diseases as celiac disease and food protein–induced enterocolitis syndrome (FPIES), while other diseases, such as eosinophilic esophagitis (EoE) and eosinophilic gastroenteritis (EGE), involve both IgE-mediated and non–IgE-mediated mechanisms. Food allergy has a significant negative effect on quality of life, and the burden on the health system amounts to approximately $25 billion US dollars annually.[1] Current management of food allergy relies on avoidance, but ongoing research into treatment for food allergy should provide options for the majority of affected persons in the near future. This chapter will review the basics of food allergy diagnosis, management, and natural history, with a specific focus on IgE-mediated food allergies.

KEY CONCEPTS

Common Characteristics of Most Frequently Encountered Food Allergens

A relatively small molecular weight (<70 kDa)
Abundant source of the relevant allergen
Glycosylation residues
Water solubility
Most are resistant to heat and digestion

PREVALENCE

The true prevalence of food allergy is difficult to establish. This is probably because *(i)* studies have focused on only the most common food allergens; *(ii)* the incidence and prevalence of food allergy may be increasing and changing with time, with studies showing an increasing prevalence over the past 10–20 years; and *(iii)* studies of prevalence are difficult to compare because of inconsistencies and deficiencies in study design.[2] Best estimates suggest a self-reported prevalence of food allergy of 12% and 13% for children and adults, respectively, with confirmed food allergy in approximately 3% of adults and children when assessed by double-blind placebo-controlled food challenge.[3] The most common foods to trigger reactions in children are cow's milk (2.5%), hen's egg (1.3%), peanut (0.8%), wheat (~0.4%), soy (~0.4%), tree nuts (*e.g.*, cashew and walnut) (0.2%), fish (0.1%), and shellfish (0.1%). Prevalent allergens vary in different cultural groups, with milk, egg, peanut, and tree nut allergens topping the lists in the Americas, Australia, and Western Europe, whereas fish and shellfish allergies are more common in Asia.[4] Most food allergies are typically outgrown; however, allergies to some foods, such as peanut, tree nuts, fish, and shellfish, are often more persistent, and these are the food allergens most commonly reported in adults. Food allergy commonly co-occurs with other atopic diseases, such as atopic dermatitis (AD), asthma, and allergic rhinitis.[5]

SPECTRUM OF DISEASE

IgE-Mediated Food Allergies

Immediate reactions to foods compose the largest proportion of food-induced allergic diseases. IgE-mediated food allergy occurs when allergenic proteins cross-link allergen-specific IgE bound to mast cells or basophils, leading to release of histamine and other inflammatory mediators. Symptoms of IgE-mediated food allergy usually occur within minutes after ingestion of the provoking food allergen, and do not start >2 hours later except in very rare circumstances (*e.g.*, delayed anaphylaxis related to red meat ingestion).

Symptoms may be severe (*e.g.*, anaphylaxis) or localized (*e.g.*,. pollen-food allergy syndrome). Characteristic signs of an immediate allergic reaction manifest through skin and subcutaneous tissues (*e.g.*, urticaria and/or angioedema), the respiratory system (*e.g.*, bronchospasm), the gastrointestinal system (*e.g.*, vomiting and/or diarrhea), and/or the cardiovascular system (*e.g.*, increased vascular permeability leading to hypotension). Anaphylaxis is the most severe symptom of food allergy and may result in death (Chapter 42).

Pollen-food allergy syndrome (PFAS, or oral allergy syndrome) symptoms are restricted to the lips, throat, and mouth and are most commonly elicited due to fruit and vegetable proteins cross-reacting with antibodies against pollen proteins in individuals with pollen allergy. Affected individuals will typically complain of pruritus and/or tingling of the lips, tongue, roof of the mouth, and throat with or without swelling. PFAS reactions are unlikely to progress to a systemic reaction.

More than 170 foods have been reported to be allergenic, but most of the food-induced allergic reactions are caused by eight food groups.[5] Cow's milk, hen's egg, and peanut are the most commonly identified food allergens in the United States, followed by tree nuts, wheat, soy, fish, and crustacean shellfish.[6] There is variability across the world and cultural groups with regard to which allergens are the most common, but cows' milk and hen's

TABLE 45.1 Food Allergens by Family

Food Allergen Family	Foods Containing Allergen	Description
Tropomyosins	Crustacean shellfish (*e.g.,* shrimp, lobster, crab), mollusks (*e.g.,* oyster, scallop, squid)	Invertebrate tropomyosins are a family of muscle proteins that share homology across invertebrate species and therefore may act as panallergens. They do not share homology with vertebrate tropomyosins. They are generally heat stable and highly cross-reactive.
Parvalbumins/EF-hand proteins	Vertebrate fish and frogs	Muscle proteins that possess a calcium-binding domain referred to as an EF-hand motif. This is the second-largest family of allergens, and these allergens are considered highly cross-reactive panallergens.
Casein	Mammalian milk	Function to bind calcium and stabilize it in micellar form. There is high sequence homology between cow's milk and other mammalian milks, such as goat's milk and sheep's milk. Other animal milks such as human milk, horse's milk, donkey's milk, and camel's milk have caseins with approximately 60% homology, which may account for less allergenicity than seen with cow's milk.
Prolamin superfamily	Seeds, tree nuts, legumes (including peanut), fruits, vegetables, wheat, corn, rice	This family contains the highest number of plant food allergens and is characterized by rich disulfide bonds and a core of eight conserved cysteine residues, providing stability and resistance to digestion. Families within this superfamily include 2S albumin seed storage proteins, nonspecific lipid transfer proteins, and α-amylase/trypsin inhibitors.
Cupin superfamily	Legumes, nuts, seeds	A large and functionally diverse superfamily of proteins termed seed storage globulins that share a β-barrel structural core domain. Seed storage globulins may be grouped into two families: vicilins and legumins.
Bet v 1 superfamily	Apple, pear, stone fruits, celery, carrot, soybean, peanut	Bet v 1 is the major birch pollen allergen and is a member of the pathogenesis-related protein 10 family within this superfamily. Symptoms of Bet v 1 hypersensitivity typically present with pollen food allergy syndrome (also known as oral allergy syndrome), which is caused by IgE cross-reactivity between Bet v 1 and homologous allergens from plant foods.

From Sampson HA, Aceves S, Bock SA, et al. Food allergy: a practice parameter update-2014. J Allergy Clin Immunol 2014;134:1016–25 e1043.

egg tend to be the most common allergens across varying geographical locations and cultural groups.[7] The most common food and plant allergens are listed in Table 45.1.

Because of protein similarities between allergens, cross-reactivity occurs. Patients allergic to certain foods should be counseled to avoid cross-reactive food proteins. It is common to find cross-reactivity among tree nuts; in particular, cashew and pistachio share common allergen binding sites, as do walnut and pecan. Between 25% and 50% of patients with peanut allergy are also allergic to tree nuts, with particular cross-reactivity noted between peanut allergens and tree nut allergens in almond, walnut, pecan, hazelnut, and Brazil nut. Tropomyosins found in crustacean shellfish are a panallergen,[5] and approximately 75% of individuals allergic to one crustacean (*e.g.,* shrimp) are likely to also react to another crustacean (*e.g.,* lobster and/or crab).[8] Parvalbumins found in vertebrate fish are commonly cross-reactive on testing, but clinical relevance of cross-reactivity varies. Studies have shown that an individual allergic to one fish species has approximately a 50% likelihood of reacting to another species of fish.

Although most reactions occur immediately after ingestion, some individuals may experience delayed anaphylaxis following ingestion of mammalian meat.[5] Delayed allergy to mammalian meats has been linked to the production of IgE to α-gal in susceptible subjects. α-Gal is an immunogenic oligosaccharide, and sensitization is believed to occur via a tick bite. Symptoms of urticaria, angioedema, and anaphylaxis can occur 3–6 hours after eating beef, pork, lamb, and venison. The mechanism(s) underlying the delayed reaction is poorly understood.

Mixed IgE/Non-IgE– and Non–IgE-Mediated Food Allergies

Delayed gastrointestinal (GI) reactions to foods include such diseases as EoE, EGE, FPIES, and eosinophilic proctocolitis (Chapter 46). The most common food proteins involved in these diseases include milk, egg, wheat, and soy. Celiac disease is a non–IgE-mediated food allergy triggered by ingestion of gluten-containing grains (*e.g.,* wheat, barley, and rye). Human leukocyte antigen (HLA) DQ2- or DQ8-restricted CD4 T cells, which recognize gluten selectively in affected persons, are critical to the pathogenesis of celiac disease.

EoE is a clinicopathological diagnosis, based on symptoms of esophageal dysfunction (including dysphagia, vomiting, feeding disorders, and abdominal pain) together with pathological findings of at least 15 eosinophils per high-power field on light microscopy.[9] The precise role of food allergy in EoE is not well defined; IgE-mediated and non–IgE-mediated mechanisms may be involved in the pathogenesis. The most common food allergens implicated in the pathogenesis include milk, egg, wheat, and soy. Numerous other foods have been implicated in the pathogenesis of EoE, and a common approach to treatment is initial elimination of milk, egg, wheat, soy, peanut, tree nuts, fish and shellfish. If dietary elimination is not successful or not feasible for the patient, then topical (swallowed aerosol) inhaled steroids (e.g. fluticasone or budesonide) may be swallowed rather than inhaled to treat the inflammation.

EGE is less common than EoE and, like EoE, it is believed that its pathogenesis involves both IgE-mediated and non-IgE-mediated mechanisms. Common symptoms of EGE include vomiting, abdominal pain, diarrhea, and failure to thrive/weight loss. Multiple food allergens are often implicated, although response to dietary elimination of the most common food allergens (milk, egg, wheat, soy, peanut, tree nuts, fish, and shellfish) is typically less successful than dietary elimination reported in patients with EoE. Topical (swallowed aerosol) inhaled steroids (e.g. fluticasone and budesonide) may provide some benefit; however, systemic steroids are often necessary for disease control.

FPIES is a non–IgE-mediated disease usually occurring in infants.[10] Characteristic symptoms of FPIES manifest as repetitive emesis with or without diarrhea accompanied by lethargy occurring 2–4 hours after ingestion of the offending food protein. The risk of abrupt volume loss, hypotension, and potential for bowel perforation makes this a medical emergency. Treatment is reliant on rehydration. Ondansetron may be helpful in managing acute FPIES reactions. Milk and soy are most commonly implicated, along with less common food allergens, such as rice, oats, fruits, or vegetables. FPIES is outgrown in the majority of affected children by 3 years of age but may be protracted for many years in a smaller subset of patients.

Celiac disease is an immune-based reaction to gluten, a storage protein for wheat, barley, and rye. The small intestine is typically affected in genetically predisposed individuals, and symptoms resolve with gluten avoidance. Symptoms of celiac disease are variable and may include diarrhea, steatorrhea, weight loss, bloating, flatulence, abdominal pain, and also non-GI symptoms, such as abnormal results on liver function tests, iron deficiency anemia, bone disease, and skin disorders. Celiac disease is detected with serological testing of celiac-specific antibodies and confirmed by duodenal mucosal biopsy, both of which should be performed while the patient is on a gluten-containing diet.[11]

PATHOPHYSIOLOGY

Food allergy results from a breakdown of oral tolerance (or failure to develop it); foods which are ordinarily harmless may then trigger an immune response that results in harmful adverse symptoms upon exposure. Maintaining tolerance requires a delicate balanced effort from multiple arms of the immune system. Deviation from the protective response may result in the development of an allergic response.

PROPERTIES OF FOOD ALLERGENS

An intact GI mucosal barrier is required to maintain tolerance. The first line of defense against the mucosal immune system is a hydrophobic layer of mucin oligosaccharides, which serve to trap antigen.[12] Secretory IgA is also a part of the outer layer of intestinal defense against dietary antigens. Dietary antigens must then penetrate the intestinal epithelium, which is maintained by epithelial junction complexes (adherens junctions) and tight junctions. Intestinal epithelial barrier dysfunction may play a role in food allergen sensitization. Alterations in the integrity of junctional complexes may be induced by calcineurin inhibitors, and this can result in food allergen sensitization. Genetic defects, such as those in individuals with filaggrin mutations, a protein that binds to keratin and is important for epithelial cell integrity, may predispose individuals to increased risk of EoE. Other factors that have been shown to affect intestinal permeability include viruses, alcohol, and nonsteroidal anti-inflammatory drugs (NSAIDs). These environmental exposures may alter intestinal epithelial integrity, thus allowing antigen to interact with the next layer of defense, mucosa associated lymphoid tissue (MALT; Chapter 20).

MALT is composed of lymphocytes, antigen-presenting cells (APCs), stromal cells, and other immune cells in the lamina propria. It is within MALT that dendritic cells (DCs) interact with dietary antigens.

There are several common characteristics among the most commonly allergenic foods: *(i)* relatively small molecular weight, generally less than 70 kilodaltons (kDa); (ii) an abundant source of the relevant allergen; *(iii)* glycosylation residues; *(iv)* water solubility; and *(v)* resistance to heat and digestion. These characteristics allow the proteins to stay intact until reaching the small intestine, where they initiate a T-helper cell-2 (Th2) response that results in production of specific IgE and eventual allergic disease. *Glycosylation* refers to the reaction by which carbohydrates are attached to molecules: in food allergens, the carbohydrate is most often attached to a protein. Carbohydrate residues surrounding proteins may be important in initiating the immune response. For example, interaction with DC-specific intercellular adhesion molecule–grabbing nonintegrin (DC-SIGN), a c-type lectin expressed on APCs that identifies conserved carbohydrate residues, has been shown to mediate recognition of the major peanut protein, Ara h 1. This interaction allows DC activation and Th2 skewing of naïve human T cells.

KEY CONCEPTS

Risk Factors for Fatal Food-Induced Allergic Reactions

Peanut and/or tree nut allergy
Delay in administration of autoinjectable epinephrine
Preexisting and/or poorly controlled asthma
Concomitant use of beta-blocker medications
Teen and young adult age groups

Once the Th2 response is initiated, it is strengthened through the induction of interleukin-4 (IL-4) signaling. IL-4 signals B cells to undergo class-switch recombination and begin producing IgE. Basophils have been implicated as a likely contributor of early IL-4 production and may play an important role in priming the T-cell response to allergens.

The Allergic Response

Allergenic food proteins that survive the initial stages of digestion are taken up by the APCs in MALT. Mucosal DCs may encounter antigen through *(i)* extending dendrites through the paracellular space between epithelial cells to sample luminal contents; *(ii)* directly interacting with the epithelial cells; and *(iii)* taking up antigen in the Peyer patch. Once contact with the antigen is established, the antigen is processed and loaded onto major histocompatibility complex (MHC) class II molecules on the cell surface, costimulatory molecules necessary for T-cell activation are upregulated, and chemotaxis to the draining lymph node occurs. Once a DC encounters a T-cell receptor with the same specificity as the peptide antigen, an immune response ensues. In the presence of cytokines, such as IL-4, IL-5, and IL-13, the responding T cell is programmed to be a Th2 cell. The Th2 cell will then signal B cells to generate IgE antibodies.

Soluble IgE that is produced by B cells circulates and binds to the surface of mast cells and basophils. Mast cells are found in skin, the gut, and the respiratory tract and are located adjacent to nerves and blood vessels. When an allergen is encountered and recognized by cell-bound IgE, calcium influx ensues, activating the mast cell. Once activated, the mast cell degranulates and releases vasoactive compounds and proteases, including histamine, platelet-activating factor, tryptase, chymase, carboxypeptidase, and heparin, resulting in the characteristic symptoms of an allergic reaction: urticaria, angioedema, flushing, nausea, vomiting, abdominal pain, diarrhea, wheezing, coughing/bronchospasm, rhinorrhea, and hypotension/syncope. Symptoms may occur alone or in combination and typically appear within minutes of ingestion.

Natural History

The majority of food allergies are outgrown. Allergies to peanut, tree nuts, fish, and shellfish are more likely to persist.[13] Clinical characteristics and laboratory measures may help predict which food allergies will be outgrown and which are more likely to be lifelong.

CLINICAL PEARLS

Important Diagnostic Considerations

The patient's history should support a diagnosis of immunoglobulin E (IgE) –mediated food allergy before performing either serum-specific IgE testing or the skin prick test.

95% predictive probability cutoffs have been established for only a few foods, including cow's milk, hen's egg, and peanut.

Even with negative specific IgE serum testing or skin testing, the patient could be allergic if he or she has a convincing history. In this case a physician-supervised oral food challenge may be required to confirm the presence or absence of an IgE-mediated food allergy.

Testing for food allergies should be limited to the food(s) in question, since positive IgE tests are not always clinically relevant. Unnecessary avoidance of foods may lead to nutritional deficiencies.

Cow's milk is typically one of the first foods introduced to infants in the form of infant formula: it is present in diets across cultural groups and is one of the most common allergens globally. Fortunately, cow's milk allergy is typically outgrown without intervention. Studies of natural resolution vary but about 50% of children with milk allergy develop tolerance between 5 and 10 years of age.[13] High levels of milk-specific IgE generally indicate a higher likelihood of persistent disease, but as many as 60% of children whose milk-specific IgE level peaks at over 50 kU/L will achieve natural tolerance by age 18 years. Up to 75% of children with reactions to uncooked milk can tolerate baked milk products. Consumption of heated milk products has been associated with accelerated acquisition of tolerance.[14]

Hen's egg allergy is another common food allergen across cultural groups. Most of the allergenic proteins in hen's egg are in the egg white. Allergy typically develops in the first year of life, while in some children, especially those with atopic dermatitis, it may develop before 4 months of age.[13] Similar to cow's milk allergy, egg allergy usually resolves during childhood without intervention. Roughly 50% of individuals with egg allergy in infancy develop natural tolerance between 6 and 9 years of age. Baked egg is tolerated by approximately 70% of children with egg allergy.[15] Individuals who can tolerate baked egg are likely to develop tolerance to lightly cooked egg sooner compared with individuals unable to tolerate baked egg.[16]

Although the majority of individuals allergic to peanut will remain reactive throughout life, approximately 20% of individuals with peanut allergy may develop natural tolerance. Favorable prognostic factors include low levels of peanut-specific IgE antibodies in the first 2 years of life and decreasing levels of IgE sensitization by 3 years. Those with peanut-specific IgE ≥3 kU/L and skin prick test (SPT) wheal diameter >6 mm before 2 years of age are more likely to have persistent peanut allergy.[17]

Allergy persistence, regardless of the food allergen, has been associated with the following factors: *(i)* earlier age at diagnosis; *(ii)* concomitant presence of other allergic diseases (*e.g.,* allergic rhinitis, asthma, and eczema); *(iii)* severity of those allergic diseases; *(iv)* symptom severity after ingestion; and *(v)* lower threshold dose required to elicit a reaction. The higher the food-specific IgE level, the more likely it is that the allergy will persist. Allergy persistence has also been associated with the rate of change of food-specific IgE levels or SPT wheal sizes. In clinical practice, food-specific IgE levels are typically checked yearly except in patients whose specific IgE levels remain high and unchanged over several years.

DIAGNOSIS

The diagnosis of food allergy begins with obtaining a detailed medical history.[6] Food-induced allergic reactions result in reproducible characteristic symptoms, as described above. Validated testing modalities are only available for IgE-mediated food allergies and celiac disease. If the clinical history does not support either diagnosis, then food allergy serum or skin testing should not be done, since there is a risk of finding sensitization to allergens which are not clinically relevant and multiple studies have shown the dangers of unnecessary dietary avoidance. When the clinical history does support a diagnosis of food allergy, this can be confirmed by SPTs and detection of specific IgE in serum.

SPTs to food allergens can be performed in the office setting and are both safe and effective, with results being available within minutes. A positive result of the SPT reflects the presence of specific IgE bound to the surface of cutaneous mast cells, but as with serum IgE testing, a positive test result does not always indicate clinical reactivity. A positive test result is generally interpreted as 3 mm larger than the negative SPT control, and the larger the SPT mean wheal diameter, the more likely it is indicative of a clinically relevant response. Negative SPT results have been associated with a high negative predictive value and may lead the physician either to offer an observed challenge or to counsel the patient on dietary reintroduction, depending on the clinical history and circumstances.

Serum-specific IgE testing is useful in providing an objective measure of food-specific IgE antibody, especially if the patient cannot stop antihistamine therapy or has extensive skin disease making it impossible to perform the SPT, and may be helpful in counseling patients on the natural history of their food allergy. Predictive values have been established for a limited number of foods. Higher specific IgE levels are more likely to be associated with clinical reactivity, but the predictive value of specific IgE levels varies across patient populations and is affected by such factors as the patient's age, ethnicity, and time since last ingestion of allergen. Specific IgE levels may also help physicians decide when an oral food challenge is or is not appropriate.

The component-resolved diagnostic (CRD) test uses allergenic proteins derived from recombinant DNA technology or purification from natural sources to identify the patient's specific IgE reactivity to individual allergenic proteins rather than to the whole allergen. Diagnostic accuracy can be enhanced in specific circumstances (*e.g.,* for peanut and hazelnut).[18] However, CRD is not routinely used for diagnosis and has not been shown to provide significant additional clinical information for most allergens. CRD for peanut provides additional diagnostic information that is helpful to the clinician, but standardized decision-making cutoffs have not yet been established.

The basophil activation test (BAT) uses flow cytometry to detect upregulation of cell-surface molecules, such as CD63 and CD203c, after stimulation with allergen.[19] BAT has been reported to be superior to the SPT, the CRD test, and the whole-allergen-specific IgE test for diagnosis of peanut allergy; however, testing has not been standardized. Further research is needed to standardize the BAT and validate results with various food allergens.

The oral food challenge (OFC) remains the gold standard for the diagnosis of food allergy.[20] OFCs can be conducted in an open manner; with a placebo control, where the patient is blinded to the product being given; or in a double-blinded manner, with both the physician and patient blinded to the food being given to the patient. An open OFC is most commonly performed in clinical practice, while the double-blind, placebo-controlled food challenge is considered the diagnostic standard typically reserved for research studies. During the OFC, a standard serving size of the allergen is divided into 4–7 servings and administered over 60–90 minutes, with each dose being given 15–20 minutes apart. The initial amount fed to the patient is typically a very small proportion of the total serving, and each successive dose administers a larger amount of protein. At the first sign of an objective reaction, the OFC is stopped and appropriate treatment administered. In cases where anxiety or subjective symptoms may affect the interpretability of the OFC, a single-blind or double-blind placebo-controlled food challenge may be preferred.

MANAGEMENT

The patient with a food allergy must maintain strict avoidance of the food allergen to prevent an allergic reaction. Avoidance requires constant vigilance. Accidental ingestion is common, with reports showing that as many as 50% of children with a peanut allergy may experience an adverse reaction in a 2-year period and up to 75% over 10 years.[6] Individuals with food allergies and their caregivers must read ingredient labels closely, prevent cross-contact, communicate with restaurant staff when eating outside of the home, and be prepared to treat a reaction, when necessary.

Food allergy labeling laws in the United States require that the presence of the most common allergens (milk, egg, peanut, tree nuts, wheat, soy, fish, and crustacean shellfish) must be declared in simple English on the ingredient labels of all packaged foods. Individuals allergic to foods other than the eight most common allergens may have more difficulty with interpretation of ingredient labels. Ingredient labels may report “spices” or “natural flavors,” which could include a multitude of foods or food products not covered by food allergy labeling laws. Statements such as “may contain [allergen]” and “manufactured on shared equipment with [allergen]” are voluntary and not regulated. Allergen content in such products is unknown, and it is typically recommended that individuals with allergies avoid products with “may contain” labeling.

Children with milk allergy or with two or more food allergies have been shown to be at particular risk of growth deficiency. Nutritional counseling with a registered dietitian is encouraged for these patients. A registered dietitian will help educate the patient and his or her family on avoidance of food allergens, in addition to providing guidance on nutrient supplementation to avoid potential dietary deficiencies.

TREATMENT OF A REACTION

An acute reaction must be recognized and treated expeditiously. Food-induced fatalities are most commonly reported from ingestion of peanut and tree nuts, but any food allergen can induce a severe reaction. Fatalities have been associated with delay in administration of autoinjectable epinephrine, preexisting and/or poorly controlled asthma, and concomitant use of beta-blocker medications; there is increased mortality in teen and young adult age groups. Intramuscular autoinjectable epinephrine must be readily available to patients with IgE-mediated food allergies and is the first-line treatment for a food-induced allergic reaction. Patients with food allergies are encouraged to have a written emergency action plan that lists the signs and symptoms of an allergic reaction and details treatment of those symptoms.

PREVENTION OF FOOD ALLERGY

Exposure to antigen early in life is likely very important for shaping the appropriate immune response to foods. Primary exposure through the oral route is believed to predispose to the development of a tolerogenic response, whereas primary exposure through skin may result in sensitization.[21] Support for the theory of prevention through primary oral exposure has been strongly supported through epidemiological studies showing that some cultural groups that introduce peanuts to their children in the first year of life have a lower incidence of peanut allergy.[22] This theory has been strengthened through recent evidence demonstrating that children identified as being at high risk of developing peanut allergy (severe atopic dermatitis and/or egg allergy) are substantially protected against the development of peanut allergy if they regularly ingest peanut, from between 4 and 11 months of age through 60 months of age, compared with matched controls who avoid peanut.[23]

Evidence in favor of a compromised skin barrier as the primary site of sensitization was shown in the Avon Longitudinal Study of Parents and Children birth cohort study, which reported that infants who had peanut allergy were more likely to have had severe AD in the first 6 months of life and also to have been treated with application of peanut oil to the skin.[24] In children with AD, the outermost layer of skin (stratum corneum) contributes to skin barrier function. Patients with AD have increased transepidermal water loss (TEWL). Increased TEWL not only allows water loss through skin but also facilitates allergen penetration and resultant sensitization. Further work is required to understand whether emollient use may reduce the risk of AD development and possibly reduce the risk of developing food allergy.

The role of the microbiome is an intense area of study with regard to food allergy development. Microbial products in the gut flora interact with innate immune receptors, such as Toll-like receptors (TLRs) and relay signals implicated in the activation of regulatory T cells (Tregs), which are important in the promotion of tolerance. Activation of a specific TLR using nonpathogenic bacteria (probiotics) could conceivably prevent allergic disease. Unfortunately, studies investigating the role of probiotics in preventing allergic disease have not shown promise, probably because of the difficulty in identifying which specific strains of gut bacteria are beneficial. Encouragingly, new techniques, such as deep-sequencing technologies, are allowing better characterization of the gut flora. As techniques improve and our understanding deepens, there is hope that new therapeutic targets will emerge for food allergy prevention.

EXPERIMENTAL INTERVENTIONAL THERAPIES

The standard of care for IgE-mediated food allergy is avoidance of the potentially triggering allergen, treatment of a reaction with autoinjectable epinephrine, and dietary supplementation of potentially deficient nutrients in the diet of the patient with a food allergy. Allergen-specific immunotherapies are currently under investigation utilizing the oral, sublingual, and epicutaneous routes for the application of the allergen. Although currently no therapies have been approved by the US Food and Drug Administration for the treatment of food allergy, several of those being investigated are promising.

Oral Immunotherapy

Oral immunotherapy (OIT) is accomplished by mixing the allergenic food into a vehicle food, initially giving doses below the level that would trigger reactions in an allergic individual and gradually increasing the amount of protein ingested over time. The buildup phase of therapy typically lasts several months; once a maintenance dose of allergen is achieved, the patient has to ingest the allergen for a certain period (typically ≥1 years, possibly indefinitely) to maintain a protected, desensitized state. Most studies have focused on achievement of desensitization, which refers to a temporary increase in the threshold of allergen required to elicit a reaction and is dependent on regular exposure to the allergen.

OIT will induce significant desensitization in most patients who are able to tolerate therapy. However, sustained protection against an allergic reaction independent of ongoing allergen exposure (sustained unresponsiveness [SU]) has not been adequately measured; only a minority of individuals achieved SU in the few studies measuring this outcome.[25] Most individuals undergoing OIT will experience adverse reactions. Oral pruritus and transient abdominal pain are the most common problems; reactions are typically mild and do not require any treatment. Severe reactions, such as anaphylaxis, may develop during therapy; predisposing factors include infection, exercise, and allergen coexposure. GI symptoms are the most common reason for participants withdrawing from OIT trials, and EoE has occasionally been documented. Further work is needed to determine which patients are most likely to develop SU, who will tolerate OIT with few dose-limiting adverse events, and the mechanisms underlying the development of desensitization and SU.

The initial immune response detected in desensitization includes an increase in food-specific IgG4, decreased basophil and mast cell responsiveness, and an initial increase in allergen-specific IgE. Allergen-specific IgE then decreases gradually over time. After 6–12 months of therapy there appears to be a shift away from Th2 cytokine production in response to allergen toward a Th1 profile. Treg upregulation occurs later in the course of OIT, with studies showing increased function of antigen-specific $CD4^+CD25^+$ $FOXP3^+$ Tregs. Epitope mapping typically changes over time indicating different antigen-specific responsiveness. Unfortunately, there are no biomarkers that consistently predict successful desensitization or sustained unresponsiveness.

Sublingual Immunotherapy

Sublingual immunotherapy (SLIT) utilizes a food protein dissolved in a liquid medium and delivered beneath the tongue. The oral mucosa contains tolerogenic APCs: SLIT is thought to rely on these cells to induce a desensitized state. SLIT dosing utilizes microgram to milligram quantities of protein, whereas OIT protocols utilize gram quantities of protein. Increasing the amount of allergen given is limited by the concentration of available extracts and the volume of liquid that can be held sublingually.

Few studies have compared SLIT with OIT; current evidence indicates that SLIT has fewer side effects compared with OIT, but SLIT does not appear to induce a similar level of desensitization or achieve SU as often as OIT. Ongoing compliance with SLIT has also been reported as challenging. Additional studies are needed to reveal whether adjuvant therapy with SLIT increases efficacy and to understand whether SLIT could be combined with OIT to improve its safety.

THERAPEUTIC PRINCIPLES

Avoidance

Read ingredient labels closely. The eight most common food allergens are required to be disclosed on ingredient labels of foods manufactured and sold in the United States.

Minimize cross-contact with food allergens during meal preparation.
- Use utensils, cutting boards, and pans that have been thoroughly washed with soap and water.
- If you are preparing several foods, make the allergy-safe food first.
- Wash hands with soap and water before touching anything else if you have handled a food allergen.
- Wash counters and table with soap and water after making meals.

When eating at a restaurant, inform the waiter and cooking staff about food allergens.

Avoid buffets.

Treatment

Advise patients at risk of anaphylaxis to carry two autoinjectable epinephrine devices at all times.

Recommend a medical identification bracelet.

Provide an anaphylaxis emergency plan, and review indications for administration of autoinjectable epinephrine.

Demonstrate the appropriate use of autoinjectable epinephrine with a trainer device at the physician's clinic visits.

Epicutaneous Immunotherapy

Epicutaneous immunotherapy (EPIT) delivers allergen to the skin through application of an allergen-containing patch. Langerhans cells in the skin are activated and effector cell responses are downregulated. Peer-reviewed published data are currently lacking, but preclinical studies have demonstrated potential for clinical efficacy. The only published trial utilizing EPIT has reported an increased threshold of reactivity after 3 months of therapy in children with milk allergy.[26] Anaphylaxis has not been reported with EPIT; the most common side effect appears to be an eczematous response at the site of patch application.

CONCLUSIONS

ON THE HORIZON

Allergen-specific desensitization therapies are currently investigational but may be available for clinical use in the near future.
- Oral immunotherapy (OIT) exposes the allergic patient to progressively larger doses of ingested allergen in an effort to induce a desensitized state. Gram quantities of allergen are typically administered.
- Epicutaneous immunotherapy (EPIT) applies microgram amounts of allergen directly to the allergic patient's skin, resulting in an effort to increase the threshold of reactivity. EPIT patch is typically kept on skin for up to 24 hours at a time, and a new patch is applied daily. Few published studies to date have reported its efficacy.
- Sublingual immunotherapy (SLIT) involves sublingual administration of milligram quantities of allergen solubilized in a liquid formulation. Systemic reactions are rare; however, studies to date have not consistently shown benefit in tested subjects.

Diagnosis of food allergy is reliant on an understanding of its pathogenesis and proper application of available diagnostic tools. Management of food allergy requires education about avoidance of the allergenic food(s), dietary supplementation of missing nutrients, and recognition and early treatment of any allergic reactions. Although food allergy has increased in prevalence over the past two decades, preventive strategies, including early introduction of allergenic solids, may help curb this as yet unexplained epidemic. As our knowledge of food allergy grows, we can expect significant changes in our approach to the treatment of those affected in the years to come.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Gupta R, Holdford D, Bilaver L, et al. The economic impact of childhood food allergy in the United States. JAMA Pediatr 2013;167: 1026–31.
2. Boyce JA, Assa'ad A, Burks AW, et al. Guidelines for the diagnosis and management of food allergy in the United States: summary of the NIAID-sponsored expert panel report. J Allergy Clin Immunol 2010;126:1105–18.
3. Gupta RS, Springston EE, Warrier MR, et al. The prevalence, severity, and distribution of childhood food allergy in the United States. Pediatrics 2011;128:e9–17.
4. Koplin JJ, Mills EN, Allen KJ. Epidemiology of food allergy and food-induced anaphylaxis: is there really a Western world epidemic? Curr Opin Allergy Clin Immunol 2015;15:409–16.
5. Sampson HA, Aceves S, Bock SA, et al. Food allergy: a practice parameter update-2014. J Allergy Clin Immunol 2014;134: 1016–25.e1043.
6. Bird JA, Lack G, Perry TT. Clinical management of food allergy. J Allergy Clin Immunol Pract 2015;3:1–11, quiz 12.
7. Fiocchi A, Brozek J, Schunemann H, et al. World Allergy Organization (WAO) Diagnosis and Rationale for Action against Cow's Milk Allergy (DRACMA) guidelines. Pediatr Allergy Immunol 2010;21:1–125.
8. Sicherer SH. Clinical implications of cross-reactive food allergens. J Allergy Clin Immunol 2001;108:881–90.
9. Dellon ES, Liacouras CA. Advances in clinical management of eosinophilic esophagitis. Gastroenterology 2014;147:1238–54.
10. Nowak-Wegrzyn A. Food protein-induced enterocolitis syndrome and allergic proctocolitis. Allergy Asthma Proc 2015;36:172–84.
11. Rubio-Tapia A, Hill ID, Kelly CP, et al. ACG clinical guidelines: diagnosis and management of celiac disease. Am J Gastroenterol 2013;108:656–76, quiz 677.
12. Vickery BP, Chin S, Burks AW. Pathophysiology of food allergy. Pediatr Clin North Am 2011;58:363–76, ix-x.
13. Savage J, Sicherer S, Wood R. The natural history of food allergy. J Allergy Clin Immunol Pract 2016;4:196–203, quiz 204.
14. Kim JS, Nowak-Wegrzyn A, Sicherer SH, et al. Dietary baked milk accelerates the resolution of cow's milk allergy in children. J Allergy Clin Immunol 2011;128:125–31.e122.
15. Leonard SA, Caubet JC, Kim JS, et al. Baked milk- and egg-containing diet in the management of milk and egg allergy. J Allergy Clin Immunol Pract 2015;3:13–23, quiz 24.
16. Leonard SA, Sampson HA, Sicherer SH, et al. Dietary baked egg accelerates resolution of egg allergy in children. J Allergy Clin Immunol 2012;130:473–80.e471.
17. Ho MH, Wong WH, Heine RG, et al. Early clinical predictors of remission of peanut allergy in children. J Allergy Clin Immunol 2008;121:731–6.
18. Grabenhenrich L, Lange L, Hartl M, et al. The component-specific to total IgE ratios do not improve peanut and hazelnut allergy diagnoses. J Allergy Clin Immunol 2016;137:1751–60.e1758.
19. Santos AF, Lack G. Basophil activation test: food challenge in a test tube or specialist research tool? Clin Transl Allergy 2016;6:10.
20. Nowak-Wegrzyn A, Assa'ad AH, Bahna SL, et al. Work Group report: oral food challenge testing. J Allergy Clin Immunol 2009;123:S365–83.
21. du Toit G, Tsakok T, Lack S, et al. Prevention of food allergy. J Allergy Clin Immunol 2016;137:998–1010.
22. Du Toit G, Katz Y, Sasieni P, et al. Early consumption of peanuts in infancy is associated with a low prevalence of peanut allergy. J Allergy Clin Immunol 2008;122:984–91.
23. Du Toit G, Roberts G, Sayre PH, et al. Randomized trial of peanut consumption in infants at risk for peanut allergy. N Engl J Med 2015;372:803–13.
24. Lack G, Fox D, Northstone K, et al. Factors associated with the development of peanut allergy in childhood. N Engl J Med 2003;348: 977–85.
25. Wood RA. Food allergen immunotherapy: current status and prospects for the future. J Allergy Clin Immunol 2016;137:973–82.
26. Dupont C, Kalach N, Soulaines P, et al. Cow's milk epicutaneous immunotherapy in children: a pilot trial of safety, acceptability, and impact on allergic reactivity. J Allergy Clin Immunol 2010;125:1165–7.

MULTIPLE-CHOICE QUESTIONS

1. A 3-year-old boy is seen in his primary care physician's office after having a reaction at a recent birthday party. His mother states that she noticed lip swelling and coughing within minutes after eating off a plate containing a variety of foods. The meal consisted of a hamburger with mustard, lettuce, onion, and pickles; potato chips; fruit punch juice; and a piece of candy containing chocolate, pecans, and caramel. Which of these foods the child ate most likely contributed to his reaction?
A. Mustard
B. Red dye #5 in the fruit punch
C. Chocolate
D. Pecan
E. Onion

2. A 14-year-old boy with a diagnosis of atopic dermatitis and multiple food allergies is seen in his primary care physician's office for his yearly well-child physical. His mother reports that he has been avoiding eggs, peanuts, tree nuts, fish, and shellfish since he was 2 years old. He has a history of anaphylaxis to peanuts, and documented sensitivity to eggs that resulted in worsening of his atopic dermatitis as an infant. Which of the following allergens is most likely to be tolerated in this child who is now 14 years old?
A. Egg
B. Peanut
C. Cashew
D. Tilapia
E. Shrimp

3. An 11-year-old male has recently been diagnosed with eosinophilic esophagitis (EoE). Which of the following is an appropriate long-term treatment option for him?
A. Daily oral prednisone
B. Avoidance of milk, egg, wheat, soy, nuts, and seafood
C. Monthly injections of omalizumab
D. Daily montelukast
E. Oral desensitization to milk, egg, and wheat

4. A new mother brings her 5-month-old child with severe atopic dermatitis into your office for evaluation. Is there anything you could tell her to decrease the likelihood of the child developing a food allergy?
A. Introduce egg products after 2 years of age.
B. Introduce wheat products after 2 years of age.
C. Introduce sweet potatoes at 6 months of age.
D. Introduce peanut products at 5 months of age.
E. Introduce shrimp at 3 years of age.

46

Eosinophil-Associated Gastrointestinal Disorders

Alex Straumann

The human gastrointestinal (GI) tract is the largest host-environment interface of the body, where the epithelial surface is exposed to an overwhelming load of diverse microorganisms as well as to dietary products. Commensal microorganisms, the so-called "intestinal microbiota," are currently regarded as an integrated part of the body involved in the regulation of several physiological functions (Chapter 14). In contrast, potentially pathogenic microorganisms of the luminal content must be recognized and, if possible, eliminated by the mucosal immune system. Several mechanisms are involved in maintaining a physical and functional barrier to stabilize body homeostasis. Besides the physical barriers, the mucosal immune system is of particular importance, with the evolutionarily ancient "innate" and the more specific and more diverse "adaptive" immunity (Chapter 20). Its task is to initiate an inflammatory response against invading microorganisms while maintaining a state of nonresponsiveness or tolerance to innocuous substances such as commensal bacteria or food antigens.

Each system of immunity relies on different cell types and gene products with characteristic modes of action. Eosinophils—part of the innate immune system—are present in the peripheral blood of healthy individuals and account for approximately 1–3% of peripheral leukocytes with an upper limit of the normal range of 350 cells per cubic millimeter of blood.[1] Tissue-dwelling eosinophils reside in the hematopoietic and lymphatic organs, such as the bone marrow, spleen, lymph nodes, and thymus. However, based on a comprehensive analysis of normal human tissues from almost all body organs, we know that the GI tract is the only nonhematopoietic organ showing significant numbers of tissue-dwelling eosinophils *under healthy conditions*.[2] The distribution of eosinophils is not homogeneous throughout the digestive tract, with the highest density of cells in the cecal and appendiceal region.[2] Under physiological conditions, the esophagus is the only segment of the digestive tract that lacks tissue-dwelling eosinophils.[2,3] Besides the resident eosinophils, a marked accumulation of eosinophils in the digestive tract can be seen under inflammatory conditions. Because eosinophils are nonspecific late-phase inflammatory cells, infiltration can occur in the context of any type of inflammation—*e.g.*, bacterial and parasitic infections, celiac disease, Crohn disease, and ulcerative colitis—independently of its cause.[1]

This chapter focuses on three idiopathic eosinophil-associated GI disorders, in particular idiopathic eosinophilic esophagitis (EoE), idiopathic eosinophilic gastroenteritis (EGE), and idiopathic hypereosinophilic syndromes (HESs) with GI manifestation (Table 46.1). The term "eosinophil-associated" denotes that the histological inflammatory response is characterized by an eosinophil-predominant tissue infiltration. Moreover, it is likely that eosinophils play a crucial role in the pathogenesis of these disorders. However, each of these conditions has different properties and probably has its own pathogenesis; to subsume them into one single category is quite arbitrary and based exclusively on descriptive features. EoE is definitely an esophageal-restricted disease with a benign long-term prognosis.[4] In patients with EGE, the inflammatory process can involve several segments of the GI tract; nevertheless, in clinical terms this chronic inflammation is also a benign disorder.[5] In contrast, HES is a multisystem disorder that may involve several organ systems, including the digestive tract, sometimes with a fatal outcome.[6] Despite these differences, all three conditions share the common features of being idiopathic, chronic, and eosinophilic, and although our understanding of the underlying mechanisms remains fragmentary, we may take the liberty of classifying them into one single category.

EoE is by far the most common eosinophil gastrointestinal disorder (EGID). Because of its clinical relevance, EoE is discussed in greater depth here than the other two eosinophil-associated disorders.

EOSINOPHILIC ESOPHAGITIS

Definition

In the latest consensus recommendations, the following conceptual definition was published: EoE represents a chronic, immune/antigen-mediated esophageal disease, characterized clinically by symptoms related to esophageal dysfunction and histologically by an eosinophil-predominant inflammation.[4] This definition shows that the diagnosis of EoE is based on clinical and on histopathological findings.

Epidemiology

Epidemiological information is essential to determine the clinical and socioeconomic impact of a disease. Therefore knowledge of the epidemiological parameters of a disease is crucial for identifying risk factors as well as pathogenetic mechanisms, for planning preventive measures, and for determining optimal treatment approaches.

Demographic Cornerstones

Cases of EoE have been reported worldwide and identified in those with a variety of ethnic backgrounds, including Caucasians, African Americans, Hispanics, and Asians. However, as there are no controlled data about geographical variations of prevalence, it remains unclear whether EoE is associated with any particular

TABLE 46.1 Eosinophil-Associated Gastrointestinal Disorders

Eosinophilic esophagitis (EoE)
Eosinophilic gastroenteritis (EGE)
Hypereosinophilic syndromes (HESs)

CLINICAL PEARLS

Eosinophil-Associated Gastrointestinal Disorders (EGID)

- All more common in males

Eosinophilic Esophagitis (EoE)

- Persistent solid food dysphagia/impaction, often with coexisting allergic diseases
- Structural changes may be present at endoscopy
- Diagnosis confirmed by histology
- Treatment includes:
  - Corticosteroids (swallowed or systemic)
  - Allergen-specific diets
  - Dilatation treatment for strictures

Eosinophilic Gastroenteritis (EGE)

- Nonspecific gastrointestinal symptoms due to eosinophilic infiltration of segments of GI tract; can affect all layers of the intestinal wall
- Diagnosis based on histology and exclusion of other causes of intestinal eosinophilia
- Treatment of serosal EGE is with systemic steroids; optimal treatment for other forms not yet defined

Hypereosinophilic Syndromes (HESs)

- Persistent peripheral blood eosinophilia >6 months
- Eosinophilic infiltration of several organ systems with related symptoms
- Gut and cardiac involvement associated with a poor prognosis
- Treatment strategies for patients with HESs include:
  - Corticosteroids
  - Cytotoxic agents
  - Interferon-α
  - Imatinib mesylate (for *PDGFRA*-associated HESs)

ethnic or racial groups, especially as most of the published studies have analyzed data from primarily Caucasian patients. EoE can be found in all age groups, but most studies report an average age between 34 and 42 years with a male-to-female risk ratio of 3:1.[7,8] EoE therefore affects mainly middle-aged male individuals with an atopic background.

Interestingly, age at diagnosis does not correlate with onset of EoE-attributed symptoms, which can be considered as the onset of disease. Several studies report a substantial time lag between onset of symptoms and time of diagnosis (diagnostic delay), which in some cases can be attributed to unawareness of sentinel features at endoscopy (doctors' delay). Often patients develop specific eating strategies, with careful chewing and avoidance of dry and rough food; despite a substantial impairment of quality of life, they bypass medical care (patient delay). An average diagnostic delay of 5 to 6 years has been reported from both North America and Europe.[7-9]

Incidence and Prevalence of Eosinophilic Esophagitis

Although considered a rare disease, several epidemiological studies from geographically confined regions indicate an increasing incidence and prevalence of EoE in the adult population during recent years.[7-9]

Probably the most conclusive epidemiological data originate from the geographically confined district of Olten, Switzerland, where a prospective population-based assessment strategy has been used with consistent diagnostic and enrollment procedures for evaluation of EoE. Between 1989 and 2009, EoE was diagnosed in 46 patients (76% males; mean age 41 ± 16 years). An average annual incidence rate of 2.45/100,000 was calculated. In the face of a constant diagnostic delay and the lack of EoE awareness programs, significantly more EoE cases were diagnosed from 2000–2009 compared with 1989–1999. Because EoE is clinically a benign disease affecting mainly younger individuals, the cumulative EoE prevalence rose to 42.8/100,000 in 2009. Extrapolating from this study, in North America and in Europe there is one patient with diagnosed EoE among every 2–3000 inhabitants. Current prevalence and incidence data of EoE are comparable with those of other chronic inflammatory diseases of the GI tract, for instance Crohn disease.

KEY CONCEPTS

Differential Diagnosis of Eosinophilic Esophagitis

- Gastroesophageal reflux disease (GERD)
- Infectious esophagitis (Herpes, *Candida*)
- Parasitic infection
- Drug-induced esophagitis
- Autoimmune disease (vascular and connective tissue diseases)
- Eosinophilic gastroenteritis with involvement of the esophagus
- Hypereosinophilic syndrome with involvement of the esophagus
- Crohn disease with involvement of the esophagus

Pathophysiology

Eosinophils' Natural Lifecycle

Eosinophils reside predominantly in three anatomical compartments: the bone marrow, blood vessels, and organs with mucosal surfaces. Eosinophils originate in the bone marrow from pluripotent stem cells. Their differentiation process is orchestrated mainly under the influence of three cytokines: interleukin (IL)-3, IL-5, and granulocyte macrophage–colony-stimulating factor (GM-CSF), leading to a fully granulated state before they migrate to the vascular space.[1] In particular, IL-5 is very specific for the eosinophil lineage; it stimulates the release of eosinophils from the bone marrow and extends their survival once they are in target tissue. Mice lacking IL-5 show a significant reduction in tissue eosinophilia, whereas mice overexpressing IL-5 show markedly increased peripheral eosinophilia.[10] A multistep process mediated by adhesion molecules on endothelial cells and corresponding ligands on eosinophils (P-selectin and β-1 and β-2 integrins) enables migration from the vascular space into tissues. This is orchestrated by T-helper cell-2 (Th2) cytokines (IL-4 and IL-13) that induce expression of cell surface ligands of the β-integrin family, such as very late antigen (VLA)-4 (β-1 integrin) on the surface of eosinophils and their counterligands on endothelial cells including vascular cell adhesion molecule (VCAM)-1.[9] Various chemoattractants—released within local mucosal environments—provoke eosinophil migration, including leukotriene B_4, platelet activating factor, chemokines, and bacterial products.[9,11] Eotaxin is important, as it binds the chemokine receptor CCR-3 on eosinophils. Eosinophils are absent from the gastrointestinal tract in mice lacking eotaxin-1.[12]

The Role of IgE in Eosinophilic Esophagitis

In up to 70% of EoE patients, there is a history of concomitant atopic diseases such as allergic rhinitis, bronchial asthma, and atopic dermatitis; elevated total serum immunoglobulin E (IgE) levels are found in about 70% of EoE patients.[7-9] Therefore both food and airborne allergens have been implicated as factors in inducing and maintaining the eosinophilic inflammation.[13] Children with esophageal eosinophilia not responding to pharmacological and/or surgical antireflux therapy showed marked improvements after being given elemental formulas.[7-9] This observation suggests that EoE could represent an IgE-mediated allergic disease in which food proteins play an important role.

On the other hand, clinical trials of targeted food elimination diets, or treatment with IgE-blocking agents, have failed to show an IgE-mediated mechanism.[7-9] Moreover, the identification of causative foods based on empiric elimination diets correlated poorly with the findings of IgE-sensitization patterns to food allergens as determined by skin prick tests (SPTs) and/or blood tests for food-specific IgE.

In summary, the current understanding is that IgE plays at most a subsidiary role in the pathogenesis of EoE and that neither IgE-based diagnostic tools nor IgE-targeted treatment is helpful.[13]

Th2-Mediated Immune Response

EoE is characterized by infiltration with cells having a characteristic Th2-type inflammatory pattern (Fig. 46.1) (Chapter 16).[14] In esophageal biopsy specimens of EoE patients, increased numbers of lymphocytes (T cells and B cells) and mast cells are found as well as increased expression of the cytokines IL-5 and IL-13.[14] The crucial role of IL-5 and IL-13 in the pathogenesis of EoE has been demonstrated in various experimental models.[15] As a proof-of-concept study of aeroallergen-induced esophageal eosinophilia, intranasal allergen challenge with *Aspergillus fumigatus* resulted in an infiltration of eosinophils in both the esophagus and the bronchi, whereas tissue eosinophilia did not develop in IL-5–deficient mice. Direct delivery of IL-13—the other important cytokine of the Th2 immune response—into the pulmonary tree induced esophageal eosinophilia, which can be blocked by antihuman IL-13 antibody.[15] Interestingly, esophageal epithelial cells produce the important eosinophil chemoattractant eotaxin-3 after IL-13 stimulation, through a transcriptional mechanism dependent on *STAT6*.[10] Esophageal eotaxin-3 strongly correlates with tissue eosinophilia, and eotaxin-3 is one of the most induced genes in patients with EoE.[12] Moreover, a single nucleotide polymorphism in the eotaxin-3 gene (*SNP* 2496 T→G) is linked to increased disease susceptibility for EoE.[12]

Accumulation and activation of eosinophils in the esophagus is followed by release of various granule proteins and cytokines. These mediators can exert cytotoxic effects (*e.g.,* major basic protein [MBP], eosinophil-derived neurotoxin [EDN], eosinophil peroxidase [EPO]), or they can contribute to perpetuation of the inflammatory response through activation of a wide range of other inflammatory cytokines, such as IL-1, IL-3, IL-4, IL-5, IL-13, IL-15, GM-CSF, TNF-α, RANTES, MIP1-α, and transforming growth factor (TGF)-β, highlighting the complexity of the pathophysiological mechanisms.[8,9]

Esophageal Remodeling

Unbridled eosinophilic inflammation leads to fibrosis and angiogenesis with an ensuing loss of elasticity of the esophageal wall and luminal narrowing. This phenomenon is called tissue remodeling and has been described in clinical studies as well as in animal models.[8,16,17] Although the exact mechanism is not known, eosinophils and subepithelial mast cells expressing tryptase are critically involved in this process via secretion of TGF-β1. TGF-β1 appears to be involved in numerous processes relevant to allergic inflammation, including regulation of profibrotic processes and modulation of smooth muscle contraction with increased contractility, thereby potentially contributing to clinical symptoms such as dysphagia and bolus obstruction in EoE patients. Furthermore, there is growing evidence in murine models that IL-5 and IL-13, the two crucial cytokines involved in tissue eosinophilia, also induce esophageal remodeling.[17]

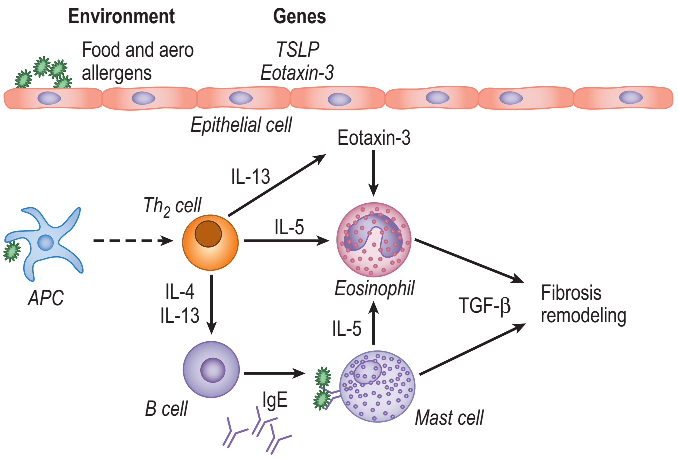


FIG 46.1 **Immunopathogenesis of Eosinophilic Esophagitis.** The current understanding of cellular and molecular mechanisms in the pathogenesis of eosinophilic esophagitis (EoE) involves a complex interaction of genetic and environmental factors. EoE disease susceptibility is linked to single nucleotide polymorphisms (SNPs) in the eotaxin-3 and thymic stromal lymphopoietin (*TSLP*) gene. A sensitization to food and aeroallergens has been noted. Two possible pathways can lead to accumulation of eosinophils in the esophageal wall. First, allergens can be recognized by antigen-presenting cells (APCs), such as dendritic cells, and can activate a type 2 helper T-lymphocyte (Th2) immune response to induce the release of eosinophil-specific cytokines such as interleukin (IL)-5 and IL-13 and subsequently the production of eotaxin-3, which is a potent chemoattractant for eosinophils. Second, allergen exposure can result in the cross-linking of immunoglobulin E (IgE) receptors on mast cells, leading to release of specific inflammatory mediators, which boost directly or through stimulation of epithelial cells' accumulation of eosinophils in the esophageal layers. Both eosinophils and mast cells produce transforming growth factor (TGF)-β, a potent cytokine involved in fibrosis and smooth muscle contraction that promotes tissue remodeling, leading to loss of elasticity of the esophageal wall and luminal narrowing.

Application of the topical corticosteroid budesonide markedly decreased expression of fibrosis-related markers such as TGF-β1 and tenascin C in esophageal tissue, suggesting that antiinflammatory treatment might interrupt or even reverse remodeling of the esophagus.[7]

Clinical Manifestation of Eosinophilic Esophagitis

The predominant symptom of adult EoE is dysphagia for solids, often leading to long-lasting food impaction requiring endoscopic bolus removal. Because EoE patients rapidly develop specific

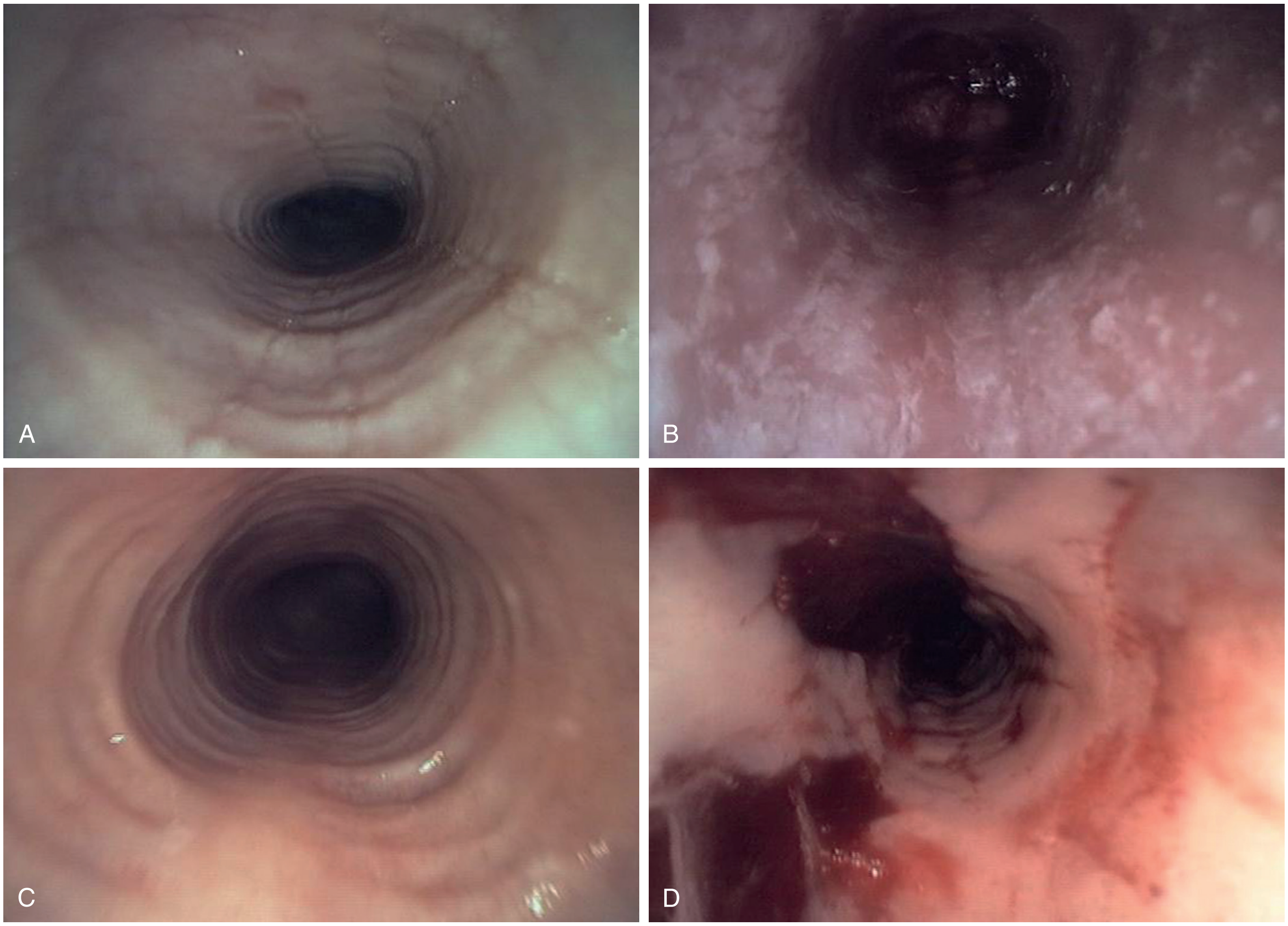


FIG 46.2 **Endoscopic Findings in Eosinophilic Esophagitis (EoE).** Representative endoscopic pictures of active eosinophilic esophagitis showing loss of vascularity and red furrows (A), whitish exudates (B), corrugated rings (C), and a long-distance stricture with a deep laceration after dilatation (D).

eating strategies, a careful evaluation of changes in eating habits is required.[18] EoE patients tend to chew carefully and avoid dry or rough food. Dysphagia and food impaction are reported increasingly with age. A minority of patients report symptoms akin to gastroesophageal reflux disease (GERD) swallowing-unrelated chest pain, and upper abdominal pain. In contrast, food refusal and failure to thrive may be observed in children, as dysphagia is less easily expressed in this age group. Children often report GERD-like symptoms such as heartburn and reflux, vomiting, and abdominal pain; less frequently, they report diarrhea (Table 46.2).[7]

TABLE 46.2 **Symptoms of Eosinophilic Esophagitis**

Adults	Children
Dysphagia	Food aversion/food refusal
Bolus obstruction	Vomiting/regurgitation
Nonswallowing-related chest pain	Abdominal pain
	Failure to thrive
	Dysphagia
	Heartburn/reflux
	Bolus obstruction
	Diarrhea

Endoscopy and Histology

Endoscopy shows a heterogeneous mix of abnormalities: subtle reddish, longitudinal furrows and white exudates may occur, reflecting local edema and acute inflammation, respectively. Transient or fixed corrugated rings as well as crêpe paper mucosa may be observed due to loss of mucosal elasticity and, most likely, chronic eosinophilic inflammation (Fig. 46.2A–D). The five major endoscopic signs of EoE are summarized in an EoE Endoscopic Reference Score (EREFS), a validated scoring instrument that should be used for the structured assessment of endoscopic activity.[19] Of note, because endoscopic features may be minimal in EoE, histological examination is recommended in every patient presenting with dysphagia.[7,8] The diagnosis of EoE is based on typical esophageal symptoms and characteristic histological findings (≥15 eosinophils/high power field [×400]), after exclusion of other esophageal diseases associated with an eosinophilic infiltrate of the mucosa (*e.g.*, GERD).[4] Several biopsies should be taken in EoE: tissue eosinophils can be patchily distributed in EoE, whereas eosinophils appear superficially along the luminal surface.[4] Because white exudates correspond to eosinophil aggregations and microabscesses (Fig. 46.3), biopsies should preferably be taken from these areas.

Treatment

As a result of recent advances in the understanding of the natural course of the disease, at least three reasons have been determined to treat EoE: (a) improvement of quality of life after resolution of the swallowing disturbances; (b) reduction of the risk of severe esophageal injury by preventing long-lasting food impactions; and (c) prevention of esophageal remodeling.[8] The current treatment options can be summarized as the three D's (Drugs, Diet, and Dilatation) and are presented in Table 46.3.

TABLE 46.3 Current Treatment Options for Eosinophilic Esophagitis: The Three D's (Drugs, Diet, and Dilatation)

THERAPEUTIC OPTIONS
Drugs
• Swallowed topical corticosteroids (budesonide, fluticasone)
• Systemic corticosteroids (prednisone)
• Biologicals (monoclonal antibodies against interleukin [IL]-13, IL-5)
• Immunosuppressants (azathioprine, 6-mercaptopurine)
• Antiallergic agents (CRTH2 blockers)
Diets
• Targeted elimination diets (individualized based on results of allergic testing)
• Empiric "6-Food" elimination diet (milk, soy, wheat, nuts, eggs, seafood)
• Elemental diet (protein-free formulas)
Dilatation
• For strictures

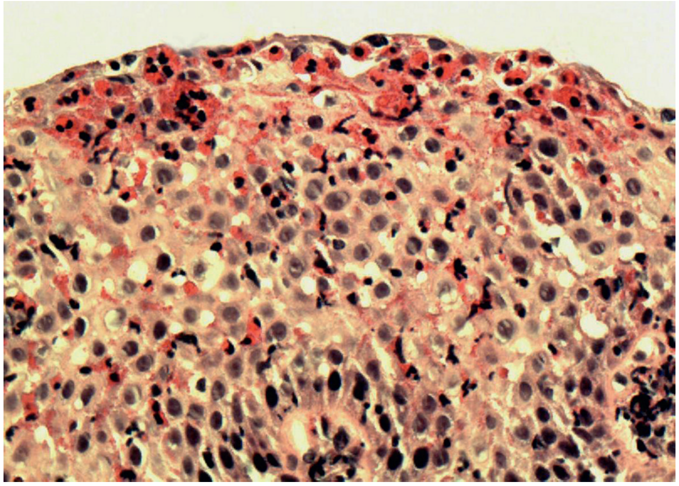


FIG 46.3 **Histopathological Findings in Eosinophilic Esophagitis (EoE).** Representative histological pictures of active eosinophilic esophagitis showing a dense infiltration of the squamous epithelium with eosinophils forming microabscesses.

Drugs

Acid suppression with proton pump inhibitors (PPI) is not usually effective in relieving symptoms and resolving eosinophilic inflammation in patients with EoE. However, patients with concomitant GERD, and a subgroup of EoE patients with "PPI-responsive" EoE, may respond to treatment with PPI. Therefore PPI therapy should not be considered as first-line treatment but instead used as cotherapy in patients with secondary or coexisting GERD.[4,8]

Systemic and topical corticosteroids show comparable effectiveness in resolving signs and symptoms of active EoE in both children and adults.[8] As topical steroids have fewer side effects, they are recommended as first-line therapy. Short-term use of systemic corticosteroids may be limited to emergent cases, such as dysphagia requiring hospitalization, patients with dehydration due to swallowing difficulties, or patients with symptoms refractory to topical steroids. Discontinuation of topical and systemic corticosteroids is usually followed by recurrence of the disease within a few weeks.[8]

Antiallergic drugs have been found largely ineffective in EoE treatment. Cromolyn sodium has no apparent therapeutic effect, and although leukotriene receptor antagonists have been shown to induce symptomatic relief, they do not affect esophageal eosinophilia.

Only limited data are available for targeted therapy with novel biological agents or immunosuppressants. The most impressive effect on symptoms, inflammation, and molecular abnormalities has been demonstrated with QAX576, a monoclonal anti-IL-13 antibody. Mepolizumab (anti-IL-5) significantly reduced esophageal eosinophils in adult EoE patients with only minimal effects on clinical improvement.[7,8] Treatment of EoE patients with the anti-TNF-α antibody infliximab did not reduce eosinophilic tissue infiltration or improve symptoms, even though massive expression of TNF-α has been shown in the squamous epithelium of the esophagus in active EoE. The immunosuppressants azathioprine and 6-mercaptopurine were effective in inducing and maintaining remission in three corticosteroid-refractory EoE patients. However, further evaluation of these alternatives for corticosteroid-refractory EoE is needed before they can be implemented into daily clinical practice.

Diet

Several prospective uncontrolled trials have been conducted to assess the potential of three different diets in treating EoE. Individually tailored so-called targeted elimination diets, empirical six-food elimination diets (removal of the six most common allergenic foods such as dairy, eggs, wheat, soy, peanuts, fish/shellfish), and a protein-free elemental diet have all shown efficacy in active EoE.[7-9] The choice of a specific dietary therapy requires that the patient's lifestyle and family resources be considered. So far, dietary treatment has been more effective in children than in adults. Overall, the value and the feasibility of dietary therapy require further evaluation.

Dilatation

Esophageal dilatation leads to long-lasting symptom relief but does not influence the underlying inflammation. It should therefore be reserved for patients who present with functional esophageal narrowing (strictures, stenosis) that is refractory to drug therapy. Endoscopic dilatation should be performed with caution, as it carries a risk of esophageal injury, although recently published large series suggest this risk is not as great as originally thought.[7,8]

IDIOPATHIC EOSINOPHILIC GASTROENTERITIS

Definition and Classification

EGE is a rare, heterogeneous, and poorly defined condition. It is characterized by chronic-recurrent tissue infiltration of the GI tract with eosinophils.[5] The diagnosis is based on the following three criteria: *1) nonspecific GI symptoms; 2) eosinophilic infiltration of one or more areas of the GI tract; and 3) exclusion of other causes for the intestinal eosinophilia.*[5] No consensus has yet been reached regarding the histological criteria for diagnosing EGE. This may be related to the fact that healthy gastric and intestinal mucosa harbors eosinophils under physiological conditions and so it is difficult to establish "normal" cell numbers.

EGE can be subclassified either according to the segments of the GI tract affected by the process or according to the depth of the eosinophilic infiltration.[5] The latter (or Klein) classification system distinguishes between mucosal, muscular, and serosal forms of EGE.[5] Parasitic infections, inflammatory bowel disease, connective tissue diseases, side effects of drugs, and lymphoproliferative malignancies must be ruled out before a diagnosis of EGE can be established.[5]

Epidemiology and Natural History

EGE predominantly affects male children and adults. In contrast to EoE, EGE is a very rare disease. During the 70 years since it was first described by Kaijser in 1937,[19a] approximately 200 cases have been reported in the literature. Although EGE is likely a chronic disorder, its natural course is still poorly defined.

Clinical Presentation

The clinical manifestation of EGE depends on its location within the GI tract and on its depth of infiltration of the intestinal wall. Mucosal involvement is typically associated with vomiting, diarrhea, abdominal pain, weight loss, failure to thrive, and either occult or frank bleeding. Involvement of the muscular layers may lead to signs and symptoms of intestinal obstruction, whereas patients with serosal involvement typically complain of bloating and can present with ascites. Peripheral blood eosinophilia is observed in approximately two-thirds of EGE patients.[5]

Diagnostic Measures

The mucosal form of the disease can be visualized by endoscopy. The findings are usually not spectacular and include thickening of the intestinal folds with deformation of the luminal configuration, diminished peristalsis, and an erythematous and friable mucosa. Moreover, endoscopy enables representative biopsy samples to be taken for histological confirmation of the diagnosis.[5] Patients with suspected serosal disease should be evaluated by laparoscopy; findings for this form of EGE include ascites, whitish nodules, and thickening of the parietal and visceral peritoneum.[5]

The muscular form of EGE can be detected by CT scan or conventional radiological examinations. Diagnostic confirmation is hampered by the difficulty in obtaining histological samples. Most reported cases of muscular EGE have been diagnosed after surgical resection of an intestinal obstruction or suspected malignancy.

Treatment

EGE is an uncommon disease, and therefore no therapeutic trials have been done. Case reports and small case series have reported responses to a variety of agents, including corticosteroids, PPIs, mast cell stabilizers, antihistamines, and leukotriene antagonists as well as surgical resection of strictured segments.[5,20]

Patients with the serosal type typically respond to steroid therapy; however, optimal treatment is not yet defined for other forms of EGE. Based on these limited data, initial treatment should be with PPIs, with a switch to systemic corticosteroids in the case of nonresponse. Unfortunately, the relapsing nature of the disease means long-term treatment may be needed with oral corticosteroids, with the inevitable risk of side effects. Leukotriene inhibitors should be evaluated in patients with frequent flare-ups.

HYPEREOSINOPHILIC SYNDROMES WITH GASTROINTESTINAL INVOLVEMENT

Definition and Classification

The HESs are a heterogeneous group of rare disorders characterized by 1) persistent peripheral blood eosinophilia with more than 1500 cells/mm^3 for a period of >6 months; 2) no known cause of eosinophilia; and 3) signs and symptoms of organ involvement.[6,21] Recent efforts have led to reclassification of these disorders according to the recognition of several clinical subtypes and new biomarkers. Currently recognized subtypes include fip1-like1/platelet-derived growth factor receptor α (*FIP1L1-PDGFRA*)-associated HESs, lymphocytic variant HESs, chronic eosinophilic leukemia, and familial eosinophilia.[22,23] Independent of this subclassification, the eosinophilic infiltration can affect the cardiovascular system (90%), the peripheral and central nervous systems (90%) including the retina, the coagulation system (80%), the skin (55%), the respiratory system (50%), the liver and spleen (35%), and the GI tract (25%).

Epidemiology and Natural History

HESs, similar to the other idiopathic eosinophilic disorders, predominantly affect middle-aged males, with a male-to-female ratio of 3 : 1. Onset of the disease is commonly between 20 and 50 years of age. Data about its prevalence or incidence are not available.[6,21]

Clinical Presentation

HES has a gradual onset, presenting initially with general symptoms such as anorexia, fatigue, weight loss, fever, abdominal pain, and night sweats. Throughout the course of the disease, the clinical manifestations vary according to which organs are involved. With GI involvement, abdominal pain and diarrhea with malabsorption have been reported, but the leading sign is hepatosplenomegaly, either due to eosinophilic infiltration or secondary to congestive heart failure. During its long-term course, extraintestinal manifestations or lymphoproliferative conditions may occur and facilitate the definite diagnosis. The ultimate prognosis depends on the extent of end organ damage. Intestinal manifestations and cardiac involvement are associated with poor prognosis and risk of fatal outcome, whereas patients with skin disease generally have a milder course.[6,21]

Immunopathogenesis

HESs encompass several disease processes with different pathogenic mechanisms. One form is characterized by a significant elevation of the eosinophilopoietic cytokine IL-5 in serum. Recently an IL-5–independent form of HES has been described in which a tyrosine kinase is overexpressed due to gene fusion.[22,24] Notwithstanding the heterogeneity of underlying mechanisms, the clinical manifestation of HES is similar, with eosinophil-mediated end organ tissue damage.

Treatment

Like EGE, HES is a rare disorder, and no prospective treatment studies have been carried out to date. Treatment strategies for patients with HESs aim to lower the eosinophil tissue infiltration. Corticosteroids have been used for decades to treat HESs and remain first-line therapy, except for *PDGFRA*-associated HESs. Cytotoxic agents, such as hydroxyurea, vincristine, and chlorambucil, have been successfully used in corticosteroid-refractory patients. Interferon-α is reported to be useful in patients with mucosal ulcerations.[23] The tyrosine kinase inhibitor imatinib mesylate has been shown to induce prolonged remissions in patients with *PDGFRA*-associated HESs. Future studies will show whether monoclonal antibodies against IL-5 could prove to be a valuable approach, given that patients with HESs often have high levels of IL-5.[25]

SUMMARY

ON THE HORIZON

- Clinical research is needed to:
  - Improve diagnostic criteria
  - Implement new diagnostic tools
  - Achieve early diagnosis
- Goals of clinical research are early treatment to improve quality of life and avoid tissue remodeling in patients with eosinophilic esophagitis (EoE).
- Multicenter clinical studies are needed to evaluate alternative treatment strategies (biologicals, immunosuppressants) for steroid-refractory EoE.
- Translational medicine will help to identify cellular and molecular factors involved in fibrosis of the esophageal wall to understand the molecular mechanisms of tissue remodeling.

Much ground still remains to be covered in diagnosing and treating patients suffering from the three eosinophil-associated GI disorders:

EoE is a common disease but challenging to diagnose due to its confusing endoscopic features. Patients, usually males, present with recurrent dysphagia and food impaction as well as coexisting allergic airway diseases. Untreated EoE can induce structural changes of the esophagus, eventually leading to an extremely fragile and rigid esophageal wall structure with increased risk for complications, including perforation. Effective treatments for adult patients with EoE include swallowed topical (or, seldom, systemic) corticosteroids, hypoallergenic diets, and esophageal dilatation. Though quality of life is substantially diminished in this chronic disorder, life expectancy is not affected.

EGE, a rather rare disorder affecting predominantly males, may involve the entire GI tract, or it may be restricted to isolated segments and affect mucosal, muscular, and serosal layers of the intestinal wall. Clinical manifestations depend on location and depth of the infiltration. Mucosal involvement, the most frequent subtype, is typically associated with vomiting, diarrhea, abdominal pain, and weight loss. The muscular form may mimic intestinal obstruction or acute abdomen. Patients with the serosal form complain mainly of bloating and ascites and usually respond dramatically to corticosteroids.

HESs are a heterogeneous group of rare disorders, characterized by persistent peripheral blood eosinophilia with more than 1500 cells/mm^3 for longer than 6 months, no known cause of eosinophilia, and signs and symptoms of organ involvement. Prognosis depends on the degree and location of the eosinophil-mediated end-organ tissue damage.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Rothenberg ME. Eosinophilia. N Engl J Med 1998;338:1592–600.
2. Kato M, Kephart GM, Talley NJ, et al. Eosinophil infiltration and degranulation in normal human tissue. Anat Rec 1998;252:418–25.
3. Straumann A, Simon HU. The physiological and pathophysiological roles of eosinophils in the gastrointestinal tract. Allergy 2004;59:15–25.
4. Liacouras CA, Furuta GT, Hirano I, et al. Eosinophilic esophagitis: updated consensus recommendations for children and adults. J Allergy Clin Immunol 2011;128:3–20.
5. Talley NJ, Shorter RG, Phillips SF, et al. Eosinophilic gastroenteritis: a clinicopathological study of patients with disease of the mucosa, muscle layer, and subserosal tissues. Gut 1990;31:54–8.
6. Fauci AS, Harley JB, Roberts WC, et al. NIH conference. The idiopathic hypereosinophilic syndrome. Clinical, pathophysiologic, and therapeutic considerations. Ann Intern Med 1982;97:78–92.
7. Straumann A, Aceves SS, Blanchard C, et al. Pediatric and adult eosinophilic esophagitis: similarities and differences. Allergy 2012;67:477–90.
8. Furuta GT, Katzka DA. Eosinophilic esophagitis. N Engl J Med 2015;373:1640–8.
9. Rothenberg ME, Mishra A, Collins MH, et al. Pathogenesis and clinical features of eosinophilic esophagitis. J Allergy Clin Immunol 2001;108:891–4.
10. Mishra A, Rothenberg ME. Intratracheal IL-13 induces eosinophilic esophagitis by an IL-5, eotaxin-1, and STAT6-dependent mechanism. Gastroenterology 2003;125:1419–27.
11. Blanchard C, Stucke EM, Rodriguez-Jimenez B, et al. A striking local esophageal cytokine expression profile in eosinophilic esophagitis. J Allergy Clin Immunol 2011;127:208–17.
12. Blanchard C, Wang N, Stringer KF, et al. Eotaxin-3 and a uniquely conserved gene-expression profile in eosinophilic esophagitis. J Clin Invest 2006;116:536–47.
13. Simon D, Cianferoni A, Spergel JM, et al. Eosinophilic esophagitis is characterized by a non-IgE-mediated food hypersensitivity. Allergy 2016;71:611–20.
14. Straumann A, Bauer M, Fischer B, et al. Idiopathic eosinophilic esophagitis is associated with a T(H)2-type allergic inflammatory response. J Allergy Clin Immunol 2001;108:954–61.
15. Blanchard C, Mingler MK, Vicario M, et al. IL-13 involvement in eosinophilic esophagitis: transcriptome analysis and reversibility with glucocorticoids. J Allergy Clin Immunol 2007;120:1292–300.
16. Straumann A, Spichtin HP, Grize L, et al. Natural history of primary eosinophilic esophagitis: a follow-up of 30 adult patients for up to 11.5 years. Gastroenterology 2003;125:1660–9.
17. Aceves SS, Newbury RO, Dohil R, et al. Esophageal remodeling in pediatric eosinophilic esophagitis. J Allergy Clin Immunol 2007;119:206–12.
18. Schoepfer AM, Straumann A, Panczak R, et al. Development and validation of a symptom-based activity index for adults with eosinophilic esophagitis. Gastroenterology 2014;147:1255–66.
19. Hirano I, Moy N, Heckman MG, et al. Endoscopic assessment of the oesophageal features of eosinophilic oesophagitis: validation of a novel classification and grading system. Gut 2013;62:489–95.

19a. Kaijser R. Zur Kenntnis der allegischen Affektioner desima Verdauungskanal von Standpunkt desima Chirurgen aus. Arch Klin Chir 1937;188:36–64.

20. Neustrom MR, Friesen C. Treatment of eosinophilic gastroenteritis with montelukast. J Allergy Clin Immunol 1999;104:506.
21. Hardy WR, Anderson RE. The hypereosinophilic syndromes. Ann Intern Med 1968;68:1220–9.
22. Owen WF, Rothenberg ME, Petersen J, et al. Interleukin 5 and phenotypically altered eosinophils in the blood of patients with the idiopathic hypereosinophilic syndrome. J Exp Med 1989;170:343–8.
23. Klion AD, Bochner BS, Gleich GJ, et al. Approaches to the treatment of hypereosinophilic syndromes: a workshop summary report. J Allergy Clin Immunol 2006;117:1292–302.
24. Simon HU, Rothenberg ME, Bochner BS, et al. Refining the definition of hypereosinophilic syndrome. J Allergy Clin Immunol 2010;126:45–9.
25. Plotz SG, Simon HU, Darsow U, et al. Use of an anti-interleukin-5 antibody in the hypereosinophilic syndrome with eosinophilic dermatitis. N Engl J Med 2003;349:2334–9.

MULTIPLE-CHOICE QUESTIONS

1. Eosinophilic esophagitis:
   - **A.** Is the most common entity among the eosinophil gastrointestinal disorders (EGIDs)
   - **B.** During the natural course of the disease, often involves stomach and small bowel
   - **C.** Is mainly triggered and driven by food-antigens
   - **D.** Immunoglobulin E (IgE)-based testing plays a crucial role in the identification of causative foods
2. Eosinophilic esophagitis:
   - **A.** Must be treated only when a patient's quality of life is markedly impaired
   - **B.** Can be treated with drugs, diet, or dilatation
   - **C.** Dilatation is regarded as first-line treatment due to the rapid symptom resolution
   - **D.** Swallowed topical steroids and elimination diet have a comparable efficiency
3. Eosinophilic gastroenteritis:
   - **A.** Is a rare disease with unclear pathogenesis
   - **B.** Can be diagnosed only after exclusion of infections and intestinal allergies
   - **C.** Can involve each segment of the gastrointestinal tract
   - **D.** Swallowed topical steroids are currently considered as first-line drugs
4. Hypereosinophilic syndromes:
   - **A.** Have an unforeseeable and sometimes even fatal natural course
   - **B.** Involve mainly the gastrointestinal tract and seldom other organ systems
   - **C.** Encompass a group of entities with different immune pathogeneses
   - **D.** Require a comprehensive hematological and multisystem examination

47

Allergic Disorders of the Eye

Virginia L. Calder, Amirah Mohd-Zaki, Valerie Saw

Allergic disorders of the eye range from the mild conditions of seasonal and perennial allergic conjunctivitis (SAC, PAC), due to immunoglobulin E (IgE)–mediated mast cell and histamine-related inflammation, to the clinically more severe diseases associated with T cell–mediated chronic inflammation, vernal and atopic keratoconjunctivitis (VKC, AKC). The ocular allergic responses in SAC and PAC are confined to the conjunctival tissues, whereas in both VKC and AKC the cornea also can become involved, potentially leading to impaired vision. Research to identify the inflammatory cells and molecules involved in allergic responses at the ocular surface and their role in the pathogenesis of allergic eye diseases has increased in the past few years. Extensive immunohistochemical studies have identified distinct expression profiles of adhesion molecules, chemokine receptors, cytokines, and immune cell types for each subtype of allergic conjunctivitis, while immune pathways are being studied via novel experimental models. This chapter reviews recent advances in our understanding of the cellular and molecular processes occurring in ocular allergy.

KEY CONCEPTS

Predominating Cell Types in the Conjunctival Tissues

In seasonal and perennial allergic conjunctivitis: mast cells
In vernal keratoconjunctivitis: eosinophils, T cells, neutrophils, mast cells
In atopic keratoconjunctivitis: T cells, eosinophils, neutrophils

SEASONAL ALLERGIC CONJUNCTIVITIS

Ocular allergy, principally SAC, is the commonest ocular disorder seen in primary care. SAC is a recurrent conjunctivitis apparent only during the pollen season; the timing (spring, summer, or autumn) is dependent on which particular pollen or mold is allergenic for that individual. The disorder can occur at any age but is more frequently seen in children and young adults and tends to lessen in severity in the older age groups.

Patients present with seasonally recurring itching, watering, redness, and swelling of the eyes and lids, and there may be increased conjunctival discharge. An associated rhinitis is common. A personal or family history of atopy very often can be elicited. During the allergen season, the ocular signs range from mild to dramatic and are usually bilateral and symmetrical. Edema of the lids and conjunctiva may be mild and often outweighs the degree of hyperemia, giving a milky or pink appearance to the eye, but the swelling may be gross, with inability to open the lids and a ballooning out of the conjunctiva termed *chemosis,* particularly after exposure to high aeroallergen concentrations or after rubbing of the eye. Eversion of the lids to reveal the tarsal conjunctiva demonstrates some hyperemia and mild infiltration (loss of transparency and thickening) of the conjunctiva with diffuse small inflammatory excrescences known as papillae. Because there is no serious limbal disease or conjunctival scarring and the cornea is not involved, the visual acuity remains normal. Outside the pollen season, the eye examination is normal.

Conjunctival Immunostaining in SAC

SAC involves an immediate, rapid early-phase hypersensitivity response due to allergens binding and cross-linking IgE molecules attached to tissue mast cells. The acute response is rapid, usually taking place within 30 min. The late-phase response occurs 6–12 h after antigen exposure and involves the infiltration of inflammatory cells, primarily mast cells, and some eosinophils and neutrophils (Fig. 47.1).[1] Primarily mast cells and their degranulation products (histamine, proteases, leukotrienes, chemokines, and cytokines) orchestrate the inflammatory response in SAC, although neutrophils and eosinophils also secrete a wide range of proinflammatory mediators that amplify the inflammation. Immunohistochemical studies of tarsal and bulbar conjunctival biopsy specimens demonstrated increased expression of intercellular adhesion molecule (ICAM)-1 and E-selectin adhesion molecules in SAC in comparison with controls, but only inseason,[2] and expression returns to baseline outside the pollen season. This pattern of expression is correlated with the degree of neutrophil or eosinophil infiltration in the bulbar tissue, suggesting a mast-cell–mediated cell recruitment process.

Therapy

General measures may provide effective relief, particularly when the disease is not severe. Limiting pollen exposure by keeping windows shut and remaining indoors, especially on high-pollen-count days, is common sense. Cold compresses and ocular lubricants (artificial tears) relieve symptoms, particularly itching.

Numerous classes of agents, both systemic and topical, have been used to manage the signs and symptoms of allergic conjunctivitis.[3] Topical mast cell inhibitors, the prototype being sodium cromoglycate, are the most useful drugs for non-sight-threatening ocular allergy. Sodium cromoglycate eye drops are available over the counter. They have a number of effects (inhibition of other leukocytes, direct mediator antagonism) in addition to inhibiting mast cell degranulation (one of the earliest mechanisms of the acute ocular allergic response), which accounts

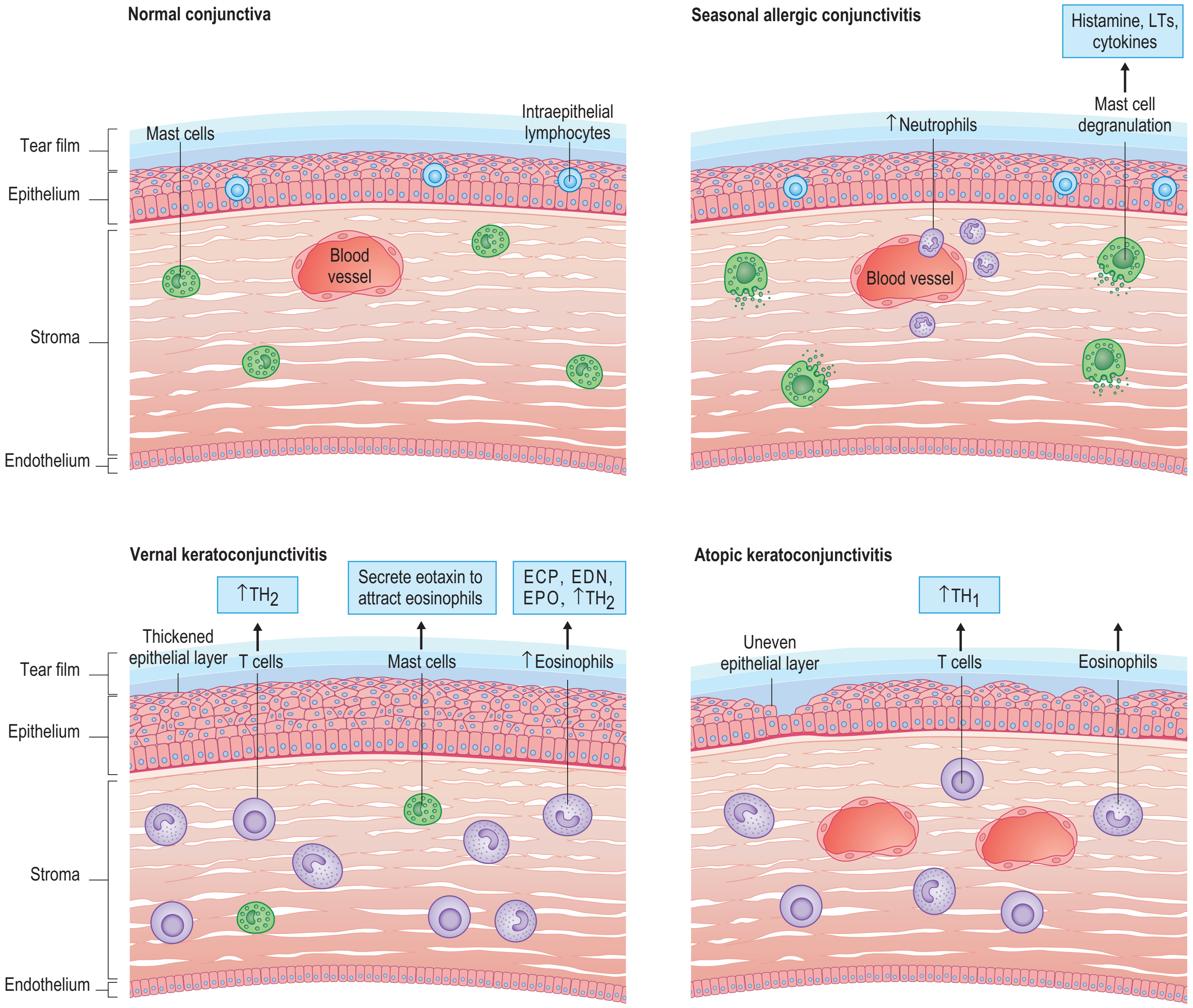

Normal conjunctiva
Tear film
Epithelium
Stroma
Endothelium
Mast cells
Intraepithelial lymphocytes
Blood vessel
Seasonal allergic conjunctivitis
Histamine, LTs, cytokines
↑Neutrophils
Mast cell degranulation
Blood vessel
Vernal keratoconjunctivitis
↑TH2
Secrete eotaxin to attract eosinophils
ECP, EDN, EPO, ↑TH2
Thickened epithelial layer
T cells
Mast cells
↑Eosinophils
Tear film
Epithelium
Stroma
Endothelium
Atopic keratoconjunctivitis
↑TH1
Uneven epithelial layer
T cells
Eosinophils

FIG 47.1 A schematic diagram of a cross-section of conjunctival tissue, showing the cell processes involved in normal vs. SAC- vs. VKC- vs. AKC-affected tissues. *AKC,* atopic keratoconjunctivitis; *ECP,* eosinophil cationic protein; *EDN,* eosinophil-derived neurotoxin; *EPO,* eosinophil peroxidase; *SAC,* seasonal allergic conjunctivitis; *VKC,* vernal keratoconjunctivitis.

for their potency and their ability to offer a preventative action if used early and continued throughout the allergen season. One limitation is the slower onset of symptom relief compared with antihistamines, although this is less of an issue with the newer, high-potency preparations (lodoxamide, nedocromil). They are extremely safe.

If topical mast cell inhibitors do not provide sufficient relief, dual-acting topical antiallergic compounds (olopatadine, epinastine, and ketotifen) combine antihistaminic properties (H1 and/or H2 receptor antagonism) with mast cell stabilization. They offer the advantage of rapid onset of action with the preventative and long-term effect of inhibiting mast cell mediator release. Olopatadine and epinastine in particular offer a clinical advantage over oral and topical antihistamines and compare well with mast cell inhibitors in allergen challenge models and, at least in some studies, in SAC.[4]

Anti-H1/H2 histamine drugs, both oral and topical, are used widely and provide rapid symptom reduction; however, they are not preventative and have limited potency, as they address only one arm of the inflammatory mediator response.[5] Oral antihistamines also treat any rhinitis, but the onset of ocular action is slower, the local effective concentration is less than for topical therapy, and they expose the patient to the risk of unwanted effects. They often are used for children when drop use can be difficult, plus at night any sedative effect can be advantageous. Long-term use of oral antihistamines leads to chronic dry eye. Topical antihistamines are available as higher-potency antihistamine-only preparations (*e.g.,* levocabastine, emedastine, azelastine), which provide rapid onset of symptom relief and more prolonged action after drop instillation. Lower-potency topical antihistamines in combination with a vasoconstrictor (*e.g.,* antazoline-naphazoline) are generally not recommended, especially for long-term use, due to rebound vasodilatation of the conjunctival vessels with long-term vasoconstrictor use. They can cause contact allergic conjunctivitis, sedation, or excitability, and they are slower and shorter-acting than high-potency topical antihistamines.

CLINICAL PEARLS

Seasonal Allergic Conjunctivitis

- Seasonal response to allergen (*e.g.*, pollen)
- Mast cell–mediated
- Antihistamines, mast cell inhibitors

Of the several topical nonsteroidal antiinflammatory drugs (NSAIDs) assessed in ocular allergy, ketorolac is the only one approved for ocular allergy. NSAIDs are not used as first-line therapy for ocular allergy. They offer the advantage of safety, but their efficacy is not well proven against other established remedies, and cost is a concern, as is discomfort upon instillation and potentially reduced patient adherence. The use of NSAIDs is more common outside the United Kingdom.

Topical steroids have a minimal role in SAC because of their potential for ocular adverse effects (see below). They occasionally are used short-term for severe disease to gain control of inflammation, but their use must be supervised by an ophthalmologist.

PERENNIAL ALLERGIC CONJUNCTIVITIS

PAC is the second most common ocular allergy. It bears many similarities to SAC, but because the allergens in PAC are present for most or all of the year, the time course of the disease is not seasonal. The disorder is again most severe and most frequently seen in children and young adults. House dust mite (*Dermatophagoides pteronyssinus*) is the most common sensitizing allergen, but animal hair and dander, molds, and other antigens may be responsible.

The symptoms are perennial ocular itch, discomfort, watering, redness, and some discharge. Patients may be able to correlate symptomatology with exposure to, for example, pets or a particular location. House dust mite allergy sufferers may give a history of symptoms worse in the morning. Approximately one-third have an associated rhinitis, and a family and/or personal atopic history is very common. The clinical appearance is of mild conjunctival inflammation, and clinical signs may be very slight. The bulbar conjunctiva (on the eye globe) may be slightly red and edematous, and the tarsal conjunctiva shows mild to moderate hyperemia, infiltration, and fine papillae. Lid edema is usually mild. As with SAC, there is neither conjunctival scarring nor corneal or serious limbal involvement, so normal visual acuity is maintained.

Immunohistological Studies in PAC

As with SAC, PAC involves an immediate-type hypersensitivity response, but because the allergens are present continuously, the resultant inflammation is more chronic. Therefore the immunopathology of PAC differs slightly from that of SAC. Increased numbers of mast cells are detected in both the tarsal conjunctival epithelium and the substantia propria, with both mucosal and connective-tissue-type mast cell phenotypes.[2] In addition, mast cells, eosinophils, neutrophils, and T cells are present, suggesting that several cell-mediated processes are involved and contribute to the chronic inflammation found in PAC.[7]

Therapy

In the case of house dust mite sensitivity, advice should be provided on mite reduction; allergy support groups can be helpful for this and for accessing equipment. Potential mite reduction measures include removing soft furnishings (*e.g.*, carpet, curtains) from the bedroom; using special vacuum cleaner filters; intensive and regular vacuuming of bedrooms, including the curtains and mattress; using a mite-impermeable mattress and pillow covers; washing bed linen and curtains at mitocidal temperatures (>131 °F or >55 °C); and using mitocidal chemicals. For mold sensitivity, dehumidifying devices and mold-killing chemicals may help.

The drug therapy of PAC is essentially as recommended for SAC (see above), but as the disease is usually prolonged, the continuing use of mast cell inhibitors assumes more importance, whereas antihistamines tend to be used intermittently for flare-ups of inflammation.

Allergen-specific immunotherapy (AIT) can be effective for patients with severe allergic rhinoconjunctivitis who have specific IgE antibodies to allergens (Chapter 91). Increasing doses of the allergen are administered by subcutaneous immunotherapy (SCIT) or sublingual immunotherapy to achieve hyposensitization mediated by production of interleukin (IL)-10 and transforming growth factor (TGF)-β1. Side effects, including anaphylaxis, are known to occur with this form of therapy. SIT requires a long-term commitment from the patient and/or caregiver in terms of time (SCIT requires monthly injections for 3 years) and expense. This may not be practical for the SAC or PAC patient who does not suffer from other comorbidities.[8]

Experimental Models of Allergic Conjunctivitis

Conjunctival Allergen Challenge

This model involves challenging the ocular surface with an allergen to artificially provoke an ocular allergic response in sensitized individuals.[9] Symptoms are similar to those seen in SAC, and so the model is useful for investigating the early- and late-phase allergic responses at the ocular surface. During the early-phase response (20 mins post challenge), increased levels of histamine and tryptase can be detected in tears, suggesting that the active cell population is predominantly mast cells. At 6 hours, histamine and eosinophil cationic protein levels are increased, but not tryptase, suggesting that basophils and eosinophils are infiltrating and active during the late phase.[10] T cells are also increased and can be seen in bulbar biopsy specimens. The allergen challenge model is often used for testing the efficacy of eye drops.

Experimental Allergic Conjunctivitis

A mouse model of experimental allergic conjunctivitis (EAC)[11] was established by an initial footpad sensitization with short ragweed pollen (SRW), followed 7–10 days later by conjunctival SRW challenge. An infiltration of mast cells, neutrophils, and eosinophils occurs, as well as increased conjunctival chemosis and lid edema. Interestingly, although there is no significant infiltration of T cells, expression of the model is attenuated in IL-12 knock-out mice. In BALB/c mice, SRW-induced EAC was interferon (IFN)-γ–dependent and could not be induced in IFN-γ knock-out mice,[12] supporting a role for Th1 pathways. A role for thymic stromal lymphopoietin (TSLP) has also been demonstrated in EAC, with TSLP expression upregulated in allergen-challenged corneal and conjunctival epithelium.[13]

VERNAL KERATOCONJUNCTIVITIS

VKC is a serious ocular allergy of childhood, comprising 0.1–0.5% of ocular disease in the developed world but more common and much more severe in hot, dry countries, especially the Middle East, West Africa, and the Mediterranean. In the United Kingdom, VKC is an unusual, self-limiting, often seasonal ocular allergy that affects children and young adults, especially males (85%) many of whom have a personal or family history of atopy. The link with atopy and seasonality is less clearly defined in cooler climates.

The symptoms are worse in the spring and summer but last all year in severe disease. Patients complain of marked itching, discomfort or pain, photophobia, stringy discharge, blurred vision, and difficulty opening the eyes in the morning. The ocular signs may be very asymmetrical. Conjunctival signs are maximal in the superior tarsal conjunctiva and limbus, and the heavily inflamed lid may droop (ptosis). The conjunctival surfaces are hyperemic, edematous, and infiltrated, and a stringy mucoid discharge is present. The tarsal conjunctiva is densely infiltrated, with papillae that are often giant (>1 mm in diameter, also known as *cobblestone papillae*). The limbus may show discrete swellings or, less often, diffuse hyperemia and infiltration. The presence of small white-chalky deposits (Trantas' dots) is typical of vernal limbitis. In the later stages, fine reticular white scarring may be seen or, in some cases, sclerosing VKC with sclerosis of the upper limbus over 90–120° of arc (from ~10–2 o'clock).

Visual acuity can be affected by involvement of the cornea (keratopathy), which is most marked in the upper third of the cornea due to greater exposure to toxic inflammatory mediators, not mechanical rubbing by the papillae. At its mildest, there is a punctate disturbance of the epithelium, which may coalesce to form a discrete epithelial defect (macroerosion). Deposition of mucus, fibrin, and inflammatory debris can then result in the formation of a shallow oval plaque (or shield) ulcer, which repels the hydrophilic tears and the epithelial healing response (Fig. 47. 2A). Herpetic and bacterial (staphylococcal) corneal infection may occur as a result of impaired cell-mediated immunity in these patients. In the later stages, scarring of the cornea can lead to permanent visual reduction. Complications related to steroid treatment, along with (because of the young age group) sensory-deprivation amblyopia, can contribute to the potential for long-term visual loss.

Immunological Studies in VKC

Several studies have identified innate and adaptive immune cells to be activated during VKC. T lymphocytes and eosinophils predominate, with mast cells, neutrophils, and other cell types infiltrating the conjunctival epithelium and stroma (Fig. 47.1). Studies of tarsal conjunctival tissue specimens have found increased numbers of activated CD4 T cells, mainly localized to the subepithelial layer of the affected tissue, and increased HLA-DR expression compared with normal subjects.[14] There is evidence that dendritic cells migrate to the lymph node to activate T cells using a CCR7-mediated process.[15] Increased numbers of Langerhans cells and activated macrophages ($CD68^+$) have also been observed. Conjunctival T cells, cloned from VKC tissues, have been functionally characterized as T-helper cell-2 (Th2) type, whereas *in situ* hybridization staining demonstrated upregulation of IL-3, IL-4, and IL-5 messenger ribonucleic acids (mRNAs) in VKC in areas of maximum T-cell infiltration, consistent with Th2 cell involvement.[16] Th1 and Th17 cells also can be involved.[17]

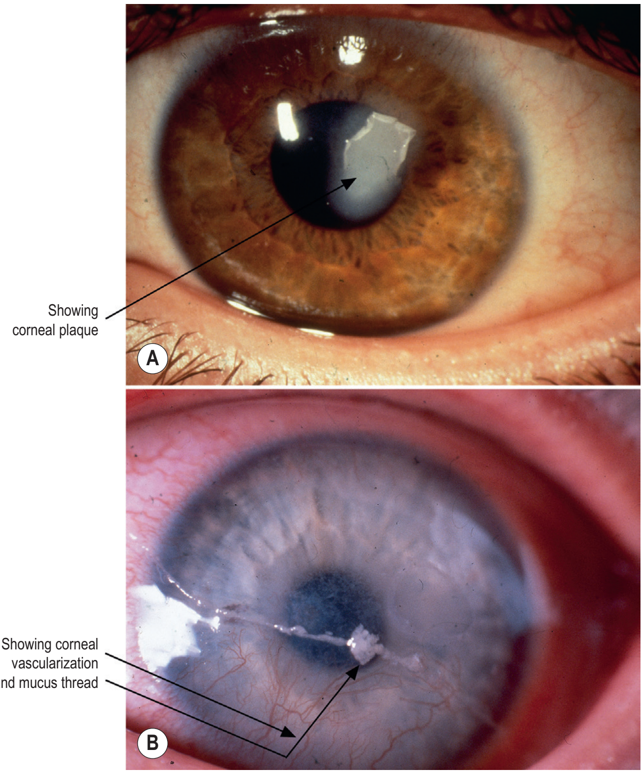


FIG 47.2 Two clinical pictures showing features of (A) vernal keratoconjunctivitis, corneal plaque formation, formation of giant papillae; (B) atopic keratoconjunctivitis, mucus thread, corneal vascularization.

VKC tear fluids were found to have increased levels of IL-4, IL-10, IFNγ, eotaxin, and tumor necrosis factor (TNF)-α compared with controls.[18] Compared with controls, conjunctival biopsy specimens from VKC patients showed increased expression of RANTES, eotaxin, monocyte chemotactic protein (MCP)-1, and MCP-3,[19] reflecting the range of inflammatory cells present.

Tissue remodeling, giant cobblestone papillae formation, limbal stem cell deficiency, and various degrees of superficial corneal opacification are further consequences of the chronic inflammation typical of VKC. Factors that promote fibroblast proliferation include Th2 cytokines, growth factors such as TGF-β, bFGF, PDGF, and also histamine.[20] These growth factors also increase integrin expression, which in turn promotes cellular infiltration and proliferation in VKC. Dendritic cells activating profibrotic T lymphocytes also may be relevant.[21]

Management of VKC

The treatment goals of VKC are to obtain adequate symptom control and to prevent complications of disease and treatment that might permanently reduce visual acuity; one must bear in mind that the disease is likely to remit spontaneously before adulthood. In cases where there appears to be pollen sensitivity, advice on pollen avoidance should be given, similar to that for SAC (see above). Simple measures such as cold compresses, ocular lubricants, and mucolytic drops may help. Oral and topical antihistamines have a very limited role in the disease.

Topical mast cell inhibitors are effective in VKC and should be maintained throughout the time of active inflammation, two to four times daily, depending on the severity of disease and which preparation is used. Olopatadine twice daily is effective and safe to use in VKC. Lodoxamide and nedocromil are more potent than sodium cromoglycate and are worth trying if olopatadine does not appear to control symptoms. Those with mild disease may be able to discontinue therapy during the winter months. It is important to emphasize to patients and parents that mast cell inhibitors are safe and must be continued when using steroids to minimize the dose of steroids required and the risk of steroid complications.

CLINICAL PEARLS

Vernal Keratoconjunctivitis

- Chronic conjunctival inflammation with seasonal exacerbations
- Eosinophils and T-helper 2 (Th2) cells infiltrate the conjunctival tissues
- Usually affects young males
- Cornea can be affected
- Increased expression of Th2 cytokines, adhesion molecules, eotaxin
- Often requires steroids and topical cyclosporine

Steroids are highly potent controllers of multiple features of allergic inflammation and are frequently required in VKC. Unfortunately they carry a significant risk of ocular adverse effects, including ocular hypertension, glaucoma, and cataract, and they can worsen infective keratitis. This is a particular concern in children, where examination to detect these complications can be difficult (*e.g.*, tonometry for intraocular pressure) and also where iatrogenic adverse effects may have long-term visual consequences beyond the time when the disease has spontaneously regressed. To minimize the risk of adverse effects, topical cyclosporine should be used, and then steroids can be prescribed in short, sharp, rapidly tapering doses during episodes of high disease activity or significant keratopathy. In addition, the use of surface-acting preparations with a reduced intraocular action (*e.g.*, fluorometholone, rimexolone, loteprednol) is advisable, although they are not available in preservative-free formulations that are preferred for high-frequency use. Systemic steroids also are sometimes used but expose the patient to numerous potential adverse effects. Supratarsal injections of steroids, either long-acting (triamcinolone) or short-acting (*e.g.*, dexamethasone), may also be used to great effect, but these are not surface-acting agents and therefore do carry significant risk of local side effects; unlike drops, neither the treatment effect nor any adverse effects can be terminated suddenly if problems arise.

Cyclosporine is a specific T-cell inhibitor but also has a number of other inhibitory effects (*e.g.*, on eosinophils, mast cells) that are likely to contribute to its effectiveness in ocular allergic therapy. Topical cyclosporine 2% dissolved in oil (usually corn) has been used to great effect in VKC and is particularly effective in treating corneal complications and in acting as a steroid-sparing agent.[22] It has no systemic adverse effects and none of the serious ocular complications of steroids, so generally it can be used safely long-term. Commercially available topical cyclosporine 0.1% in a cationic emulsion (Ikervis in UK; Restasis in US) has recently been licensed for use in dry eye. A randomized study in Europe evaluating 4×/day topical cyclosporine 0.1% in children with VKC is ongoing.

Surgical interventions are sometimes required in VKC for the corneal manifestations; surgical or excimer laser superficial keratectomy may be used in conjunction with medical therapy for plaque ulcer; rarely, corneal grafting may be required for scarring. Surgical removal of giant papillae or conjunctival reconstruction is not generally recommended.

Experimental Model of VKC

In genetically susceptible rats and mice, experimental blepharoconjunctivitis (EBC) is inducible by subcutaneous immunization with SRW followed by conjunctival allergen challenge at day 10. In this model a significant eosinophilia is found 24 hours after challenge, and EBC has therefore been used as a model for studying eosinophil infiltration as found in VKC. It also can be induced by adoptively transferring Th2, but not Th1, T cells, consistent with our understanding that eosinophil infiltration is Th2-mediated.[23]

ATOPIC KERATOCONJUNCTIVITIS

AKC is the least common but most serious of the ocular allergies. It is a life-long condition that affects adults who have systemic atopic disease, particularly atopic dermatitis. It usually starts in the late teens, but unlike VKC, the disease is persistent and may be relentlessly progressive; occasionally the disease can begin in childhood. AKC is a highly symptomatic disorder with severe itching, pain, watering, stickiness, and redness of the eyelids and eye.[24]

CLINICAL PEARLS

Atopic Keratoconjunctivitis

- The most severe form of allergic eye disease
- Affects adults with atopic dermatitis or asthma
- Predominant infiltration of T cells expressing interferon (IFN)γ in severe cases
- Cornea can be affected, often due to secondary infections
- Requires steroids and cyclosporine

There is usually facial atopic dermatitis involving the eyelids. The lid margins show severe blepharitis (chronic inflammation of the lash follicles and Meibomian glands) and are thickened and hyperemic, posteriorly rounded, and sometimes keratinized; the lid anatomy may be distorted with ectropion (outwardly turning eyelid), entropion (inwardly turning eyelid), trichiasis (in-turning lashes), loss of lashes, and notching. The whole conjunctiva is affected and shows intense infiltration, papillae (which may be giant), and sometimes scarring with linear and reticular white scar tissue, lid-to-conjunctiva adhesions (symblepharon), shrinkage or loss of the conjunctival sac, and secondary lid distortion. Marked limbal inflammation can develop, and Trantas' dots may occur. The disease may not affect the cornea at all, in which case it is sometimes called atopic blepharoconjunctivitis (ABC); in this situation, the overall inflammation is generally less severe.

The cornea may be affected directly during inflammation, or it may be damaged secondarily after extensive changes to the protective ocular surface by continual mechanical trauma, reduced lid protection, or severe loss of conjunctival tear production. Significant visual acuity reduction due to corneal involvement occurs in 40–70% of cases. Keratopathy may consist of punctate and macroscopic epithelial defects, filamentary keratitis, plaque ulcer, progressive scarring, neovascularization (with or without lipid deposition), thinning, and secondary corneal infections (herpetic, bacterial, and fungal) (Fig. 47.2B). There are recognized associations between AKC and eye rubbing, keratoconus, atopic cataract, and retinal detachment.

Immunological Studies in AKC

In AKC the predominant cell types infiltrating the conjunctival tissues are T cells, eosinophils, and neutrophils (Fig. 47.1). As observed in VKC, increased numbers of activated CD4 T cells, HLA-DR expression, and cells of the monocyte/macrophage lineage are found in conjunctival biopsy specimens in AKC,[14] as well as mRNA expression of IL-3, IL-4, and IL-5 in the stroma. However, in AKC there is a significant increase in the expression of IL-2 mRNA and in the numbers of IFN-γ–expressing T cells, suggesting a Th1-mediated inflammation in this disease.[16] Collagen deposition and conjunctival tissue remodeling are considerable in VKC and AKC, so the production of proinflammatory cytokines by infiltrating cells may be the mechanism whereby tissue-resident cells such as conjunctival fibroblasts become involved.

Experimental Model of AKC

To investigate the immunopathogenic mechanisms involved during AKC, a new clinically more severe model of EAC has been developed.[25] In this model, EAC was induced in *C57BL/6* mice by systemic immunization and topical daily challenge with ovalbumin (OVA). OVA-challenged mice exhibited more severe clinical symptoms compared with other mouse models described above, with chemosis and hyperemia. Because this model demonstrated clinical symptoms similar to AKC, it has been used to show that mast cell, CD4 T cell, and dendritic cell numbers are increased during disease, while IL-9, a growth factor for mast cells, was also upregulated (Mohd Zaki et al, *submitted;* Fig. 47.3).

Therapy

Therapy for AKC not only aims to control symptoms but also attempts to modify and reduce serious sight-threatening sequelae; therefore it should be aggressive. The topical treatment of the ocular surface is similar to VKC in that some general therapy may help, antihistamines are not useful, mast cell inhibitors are continued long-term, and steroid and cyclosporine drops are often required. However, the disease is generally less episodic than VKC, so long-term steroid use is often needed, and steroid-related complications are more problematic. In particular, herpetic keratitis, which is more common and severe in AKC, can be potentiated by topical steroids. Topical tacrolimus is more potent than topical cyclosporine and well tolerated.[26] Facial and lid dermatitis should be actively managed, if necessary in conjunction with a dermatologist, whereas lid margin inflammation (blepharitis) should be treated with hot compresses followed by lid hygiene, topical antibiotic, and/or steroid preparations and systemic low-dose antibiotics (especially tetracyclines), all of which will lessen the need for antiinflammatory and immunosuppressive therapy.

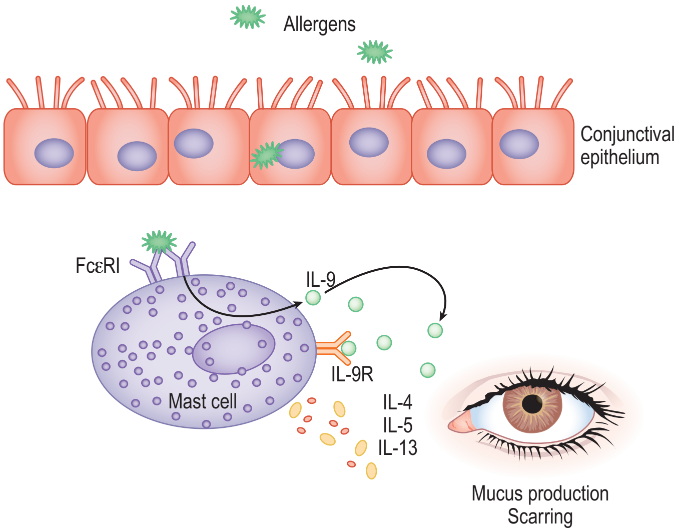


FIG 47.3 Hypothesis of a role for interleukin-9 in allergic conjunctivitis.

Systemic therapy can be necessary in severe cases, particularly when surgical therapy is undertaken, and includes steroids, cyclosporine, and sometimes other immunosuppressive agents such as mycophenolate mofetil. All of these carry risks of serious side effects and also affect the general atopic picture, so consultation with the patient's physician is advisable.

A significant number of patients require ocular surgery, either as a consequence of the disease or because of associated keratoconus. Surgery for AKC includes both elective procedures and emergency interventions and may consist of corneal gluing, patch grafts, corneal transplants (partial or full thickness), conjunctival reconstruction, amniotic membrane grafts, and limbal transplantation. These are generally high-risk procedures and often require support with systemic immunosuppression.

In summary, allergic disorders of the eye range in severity and duration. Recent studies have increased our understanding of the molecular mechanisms involved. VKC and AKC require immunosuppressive therapy, which can have serious side effects. Although the therapeutic options for treating mast cell–mediated forms of allergic conjunctivitis have improved, there is still a need to find alternative, safer therapies for the more severe and chronic forms.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Anderson DF, MacLeod JD, Baddeley SM, et al. Seasonal allergic conjunctivitis is accompanied by increased mast cell numbers in the absence of leucocyte infiltration. Clin Exp Allergy 1997;27(9):1060–6.
2. Bacon AS, McGill JI, Anderson DF, et al. Adhesion molecules and relationship to leukocyte levels in allergic eye disease. Invest Ophthalmol Vis Sci 1998;39(2):322–30.
3. Castillo M, Scott NW, Mustafa MZ, et al. Topical antihistamines and mast cell stabilisers for treating seasonal and perennial allergic conjunctivitis. Cochrane Database Syst Rev 2015;(6):CD009566.
4. Bilkhu PS, Wolffsohn JS, Naroo SA. A review of non-pharmacological and pharmacological management of seasonal and perennial allergic conjunctivitis. Cont Lens Anterior Eye 2012;35(1):9–16.
5. del Cuvillo A, Sastre J, Montoro J, et al. Allergic conjunctivitis and H1 antihistamines. J Investig Allergol Clin Immunol 2009;19(Suppl. 1):11–18.
6. Dart JK, Buckley RJ, Monnickendan M, et al. Perennial allergic conjunctivitis: definition, clinical characteristics and prevalence. A comparison with seasonal allergic conjunctivitis. Trans Ophthalmol Soc U K 1986;105(Pt 5):513–20.
7. Leonardi A, De Dominicis C, Motterle L. Immunopathogenesis of ocular allergy: a schematic approach to different clinical entities. Curr Opin Allergy Clin Immunol 2007;7(5):429–35.
8. Kari O, Saari KM. Updates in the treatment of ocular allergies. J Asthma Allergy 2010;3:149–58.
9. Friedlaender MH. Objective measurement of allergic reactions in the eye. Curr Opin Allergy Clin Immunol 2004;4(5):447–53.

10. Bacon AS, Ahluwalia P, Irani AM, et al. Tear and conjunctival changes during the allergen-induced early- and late-phase responses. J Allergy Clin Immunol 2000;106(5):948–54.
11. Magone MT, Whitcup SM, Fukushima A, et al. The role of IL-12 in the induction of late-phase cellular infiltration in a murine model of allergic conjunctivitis. J Allergy Clin Immunol 2000;105(2 Pt 1):299–308.
12. Stern ME, Siemasko K, Gao J, et al. Role of interferon-gamma in a mouse model of allergic conjunctivitis. Invest Ophthalmol Vis Sci 2005;46(9):3239–46.
13. Zheng X, Ma P, de Paiva CS, et al. TSLP and downstream molecules in experimental mouse allergic conjunctivitis. Invest Ophthalmol Vis Sci 2010;51(6):3076–82.
14. Metz DP, Bacon AS, Holgate S, et al. Phenotypic characterization of T cells infiltrating the conjunctiva in chronic allergic eye disease. J Allergy Clin Immunol 1996;98(3):686–96.
15. Saban DR. The chemokine receptor CCR7 expressed by dendritic cells: a key player in corneal and ocular surface inflammation. Ocul Surf 2014;12(2):87–99.
16. Metz DP, Hingorani M, Calder VL, et al. T-cell cytokines in chronic allergic eye disease. J Allergy Clin Immunol 1997;100(6 Pt 1):817–24.
17. Reyes NJ, Saban DR. T helper subsets in allergic eye disease. Curr Opin Allergy Clin Immunol 2014;14(5):477–84.
18. Leonardi A, Curnow SJ, Zhan H, et al. Multiple cytokines in human tear specimens in seasonal and chronic allergic eye disease and in conjunctival fibroblast cultures. Clin Exp Allergy 2006;36(6):777–84.
19. Abu El-Asrar AM, Struyf S, Al-Kharashi SA, et al. Chemokines in the limbal form of vernal keratoconjunctivitis. Br J Ophthalmol 2000;84(12):1360–6.
20. Vichyanond P, Pacharn P, Pleyer U, et al. Vernal keratoconjunctivitis: a severe allergic eye disease with remodeling changes. Pediatr Allergy Immunol 2014;25(4):314–22.
21. Dale SB, Saban DR. Linking immune responses with fibrosis in allergic eye disease. Curr Opin Allergy Clin Immunol 2015;15(5):467–75.
22. Gonzalez-Lopez JJ, Lopez-Alcalde J, Morcillo Laiz R, et al. Topical cyclosporine for atopic keratoconjunctivitis. Cochrane Database Syst Rev 2012;(9):CD009078.
23. Fukushima A, Ozaki A, Jian Z, et al. Dissection of antigen-specific humoral and cellular immune responses for the development of experimental immune-mediated blepharoconjunctivitis in C57BL/6 mice. Curr Eye Res 2005;30(4):241–8.
24. Foster CS, Calonge M. Atopic keratoconjunctivitis. Ophthalmology 1990;97(8):992–1000.
25. Schlereth S, Lee HS, Khandelwal P, et al. Blocking CCR7 at the ocular surface impairs the pathogenic contribution of dendritic cells in allergic conjunctivitis. Am J Pathol 2012;180(6):2351–60.
26. Erdinest N, Solomon A. Topical immunomodulators in the management of allergic eye diseases. Curr Opin Allergy Clin Immunol 2014;14(5):457–63.

MULTIPLE-CHOICE QUESTIONS

1. Which patient with ocular allergy would you treat with specific immunotherapy?
   A. Seasonal allergic conjunctivitis (SAC)
   B. Perennial allergic conjunctivitis (PAC)
   C. Vernal keratoconjunctivitis (VKC)
   D. Atopic keratoconjunctivitis (AKC)

2. Anti-IgE is indicated in which forms of ocular allergy?
   A. SAC
   B. PAC
   C. VKC
   D. AKC

3. Are topical cyclosporine and/or tacrolimus indicated in ocular allergy?
   A. Yes
   B. No
   C. Unknown

48

Drug Hypersensitivity

Mariana Castells, Rafael Bonamichi-Santos

KEY CONCEPT

Main Presentations of Drug Hypersensitivity

1. Type I reactions with the involvement of mast cells and basophil triggering by IgE and/or non-IgE mechanisms and ranging from urticaria to anaphylaxis
2. Type IV delayed hypersensitivity reaction presenting as delayed rash; drug-specific T cell–specific clones may be identifiable
3. Severe cutaneous reactions, including Stevens-Johnson syndrome (SJS), toxic epidermal necrolysis (TEN), drug-induced reactions with eosinophilia and systemic symptoms (DRESS), and acute generalized exanthematous pustulosis (AGEP), in which HLA genotypes and viral infections play major roles

Reactions to drugs are an important source of morbidity and mortality and constitute a major hazard in the practice of medicine. We have an incomplete understanding of the mechanisms of the reactions, few tools for their diagnosis and/or prevention, and limited treatment options.[1,2] With the introduction of new and better medications to treat cancer, chronic inflammatory diseases, and infections, there has been a parallel increase in instances of drug allergy and hypersensitivity. Although these new, targeted therapies offer increased survival and quality of life, repeated exposures to small molecules and monoclonal antibodies (mAbs) with immunogenic and allergenic capacity have increased the rate of sensitizations and reactions. Drug hypersensitivity prevents patients from receiving first-line therapy, and doctors have few guidelines on how to diagnose and address these potentially life-threatening reactions.[3,4]

Identifying drug-hypersensitive patients protects them from reexposure to culprit medications and permits access to desensitization procedures when needed and indicated. Identifying patients who are not in fact hypersensitive means they can avoid receiving second-line therapies with their associated increased costs and resource requirements.[5] The vast majority of patients with a history of penicillin allergy have negative results on skin testing and do not react during a challenge or during further courses of penicillin,[6] but there is a lack of standardized reagents approved by the US Food and Drug Administration (FDA) for the diagnosis of allergy to the majority of antibiotics and other drugs. Central to the evaluation of patients with drug hypersensitivity is an understanding of the pathophysiology of the reactions. Genotype studies of patients with severe hypersensitivity reactions, including drug-induced reactions with eosinophilia and systemic symptoms (DRESS), Stevens-Johnson syndrome (SJS), toxic epidermal necrolysis (TEN), have uncovered predisposing specific HLA alleles permitting identification of targeted populations and protection against these reactions.[7] Systematic utilization of available markers of immune cell activation, such as tryptase (the major mast cell protease release during anaphylactic IgE- and non-IgE–mediated reactions) can provide useful diagnostic information and help guide the treatment and management of drug reactions.[8]

EPIDEMIOLOGY OF DRUG HYPERSENSITIVITY

Drug hypersensitivity has a spectrum of presentations and can be fatal[9]: the most common presentation is skin rashes, observed in about 2–3% of hospitalized patients.[10,11] Any drug can elicit hypersensitivity reactions: antibiotics, chemotherapy drugs, mAbs, antihypertensives, radio contrast media, and antiepileptics are among the most frequent causes.[11] The risk of sensitization and of a reaction depends on the drug chemistry and structure, the host immune system, the dose, the route, the duration of treatment, the patient's gender and specific HLA alleles. In a study from Brazil, >50% of the anaphylactic reactions treated in the emergency department were induced by nonsteroidal anti-inflammatory drugs (NSAIDs), the most frequently implicated medications.[12] A recent US study of drug-induced anaphylaxis found that most patients were female, few received appropriate treatment with epinephrine, and 29% had concomitant asthma and allergic rhinitis. A high proportion of these patients were receiving antidepressants, angiotensin-converting enzyme (ACE) inhibitors and/or angiotensin II blockers, confirming the association of these medications with severe reactions.[13]

IMMUNE DRUG RECOGNITION AND GENETIC BASIS OF DRUG HYPERSENSITIVITY

Drugs are recognized by the immune system and initiate immune responses, depending on their structure, size, and multimeric/multivalent presentation. Most drugs are small molecules, which require haptenization, or binding to carrier proteins, to interact with antigen-presenting cells (APCs), including B cells. Once B cells are activated, epitope-specific antibodies are made, including IgE if interleukin-4 (IL-4) is present.[14] Beta-lactam antibiotics, such as penicillin and cephalosporins, are haptens with unstable beta-lactam rings that acylate lysine residues in proteins, forming covalent bonds and leading to the major penicilloyl epitope responsible for urticarial reactions. Carboxyl and thiol groups in beta-lactams can also haptenate. These hapten–carrier complexes are called *minor allergens* because they are found in only a small proportion of cases; they have been implicated in anaphylactic IgE-mediated reactions to penicillin.[15] Large molecules, such as succinylcholine, which has quaternary ammonium epitopes, can act as complete allergens. Epitope-specific antibodies can elicit type I, II, and III hypersensitivity reactions (see below). Type IV (T cell–mediated) reactions can occur through several mechanisms. Currently at least three models are proposed—the hapten/prohapten model, the p-I model, and the altered peptide repertoire model (Fig. 48.1).[16] Drugs or their reactive metabolites undergo intracellular processing to generate a pool of peptides, which then are presented in the context of MHC as foreign peptides at the surface of the APC and elicit a T-cell response (see Fig. 48.1A). The three current models for T-cell interactions with small drugs are shown in Fig. 48.1B. The prohapten/hapten model is based on the drug binding to a protein for processing and generating haptenated peptides recognized by T cells as neoantigens. In the p-I model, small drug molecules bind directly to the HLA molecule or T-cell receptor (TCR) in a noncovalent manner and stimulate T cells. In the altered repertoire model, small molecules bind directly to the MHC-binding cleft to alter the specificity of peptide binding. This results in the presentation of novel peptide ligands that elicit an immune response. Fig. 48.1C shows the altered repertoire model in the case of abacavir (green), which binds to HLA-B*57:01 (gray). The neoepitopes generated by abacavir are thought to be responsible for T-cell activation and the symptoms of hypersensitivity.

At least 24 HLA alleles have been identified in association with drug hypersensitivity, providing an insight into both the complexity of the immune response to drugs and our incomplete

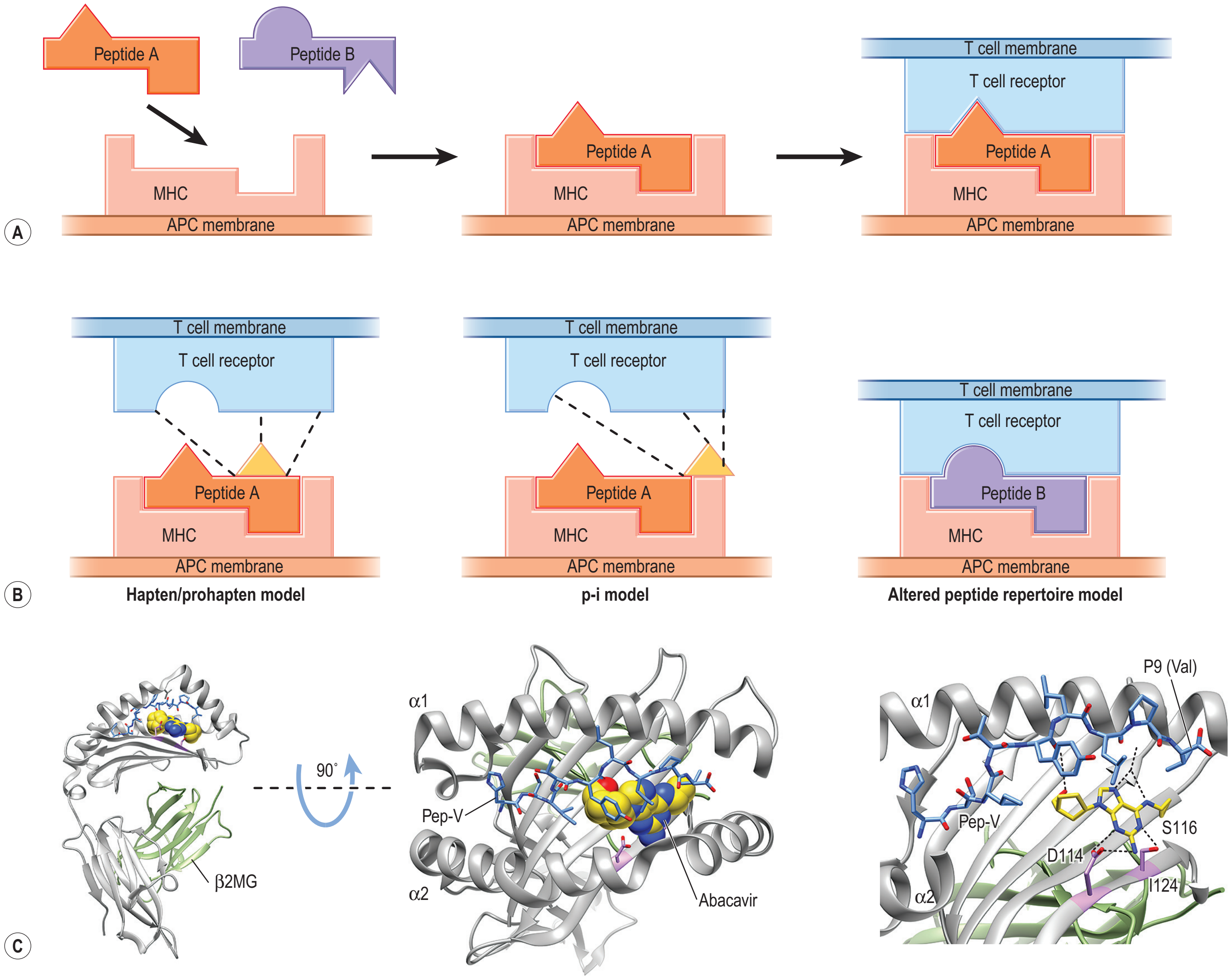

Peptide A
Peptide B
MHC
APC membrane
T cell membrane
T cell receptor
Peptide A
MHC
APC membrane
T cell membrane
T cell receptor
Peptide A
MHC
APC membrane
A
T cell membrane
T cell receptor
Peptide A
MHC
APC membrane
T cell membrane
T cell receptor
Peptide A
MHC
APC membrane
T cell membrane
T cell receptor
Peptide B
MHC
APC membrane
B
Hapten/prohapten model
p-i model
Altered peptide repertoire model
β2MG
90°
α1
Pep-V
α2
Abacavir
P9 (Val)
α1
Pep-V
S116
D114
α2
I124
C

FIG 48.1 Models of T-cell activation by small molecules. (A) Peptide selection and presentation by MHC. (B) Mechanisms of T-cell activation by small molecules. (C) The abacavir–HLA-B*57:01 interaction demonstrates the altered peptide repertoire model. (From White KD, et al. Evolving models of the immunopathogenesis of T cell-mediated drug allergy: the role of host, pathogens, and drug response. *J Allergy Clin Immunol.* 2015; 136: 219–34.)

understanding of the pathogenesis of these reactions.[17] Although different organ systems are affected, typically the targets are the liver and skin; other organs, such as muscles (statin-induced myopathy) and neutrophils (clozapine-induced agranulocytosis), can be affected, indicating a heterogeneity determined not just by the associated HLA allele but also by the expression of homing receptors and T-cell clonotypes.[16] Striking ethnic variations, such as the high prevalence of HLA-B*15:02 in carbamazepine-induced SJS/ TEN have been uncovered in Han Chinese, Thai, and Malay populations, but not in Japanese, Korean, or Northern Europeans (Table 48.1). Some HLA alleles are associated with reactions to unrelated compounds, such as B*57:01, which has been linked to reactions to abacavir and flucloxacillin, with different organ target effects. In contrast, specific HLA alleles may predispose to a particular type of organ injury regardless of the drug involved in the reaction, as seen with HLA-DRB1*15:01–associated liver injury with lumiracoxib and co-amoxiclav. Although HLA-associated reactions occur within hours to days of drug exposure, additional factors must influence these interactions because rechallenge with the drug may not reproduce the symptoms, and time lags have been observed. Such factors can be influenced by the dose. High doses of phenytoin can increase the rate of cutaneous eruptions, which suggests that genetic polymorphisms influence drug metabolic pathways leading to a potential increase in the risk for idiosyncratic reactions. A recent genome-wide association study has linked phenytoin reactions to allelic variations of cytochrome P450 isoform CYP2D9*3, with resulting decreased detoxification activity of CYP2C9.[18] Evidence for the association of specific HLA alleles in penicillin/beta-lactam or sulfonamide reactions is less clear, perhaps because these medications are metabolized into several antigenic and allergenic components that interact with multiple HLA types. The recently reported association between penicillin-induced type I reactions and HLA-DRA indicates the need for strong phenotype association with the culprit drug and adequate sample size.[19] The strength of the association between HLA alleles and certain drug reactions can be very strong, as seen with abacavir and HLA-B*57:01. Patients expressing the relevant allele can be patch tested before exposure to the drug. Prospective DNA screening has a 100% negative predictive value, eliminating all potential reactors.

RISK FACTOR FOR DRUG ALLERGY: HOST, PATHOGENS, AND DRUG RESPONSE

Females are at higher risk for drug reactions, but the basis of this propensity has not been studied.[20] Quaternary ammonium compounds present in cosmetics are thought to be cross-reactive with neuromuscular-blocking agents and may increase the risk of drug reactions during general anesthesia. Patients with cystic fibrosis and those with multiple exposures to latex during surgery were found to have latex IgE sensitization and presented severe reactions, including asthma and anaphylaxis. Decrease in aerosolized latex particles has effectively decreased reactions.[21] Because patients with cystic fibrosis have recurrent infections, these patients have an increased exposure to antibiotics and an increased rate of type I IgE-mediated reactions,[22] including anaphylaxis. Increased rates of drug reaction have been observed in patients with human immunodeficiency virus (HIV) infection, especially to trimethoprim-sulfamethoxazole. Low CD4 T cell counts as well as active immune suppression as a result of antiretroviral therapy are thought to increase the risk of reactions.

Heterologous responses to viral infections, in the context of specific HLA alleles, have been implicated in severe skin reactions. T cell–mediated immunity can be modified by heterologous immune responses that are induced by pathogen-specific T-effector-memory cells. Microbial pathogens, such as human herpes virus (HHV) and cytomegalovirus (CMV), among others, can establish lifelong infections and cellular latency. From time to time, latent virus can activate transcriptional programs, resulting in viral proteins that stimulate virus-specific T-cell memory.

The induction of this type of drug hypersensitivity requires an HLA risk allele, a primary infection with HHV or other viruses in the context of that HLA risk allele, polyclonal expansion of CD8 T cells, and the induction of memory T cells. Upon exposure to the drug, the interaction with HLA induces neoantigens, and the peptide–MHC complex is recognized by the TCR, triggering activation of T cells (Fig. 48.2).[16]

About 25% of women who are treated with carboplatin for gynecological cancers become sensitized after several exposures and then have type I IgE-mediated reactions, most of which are anaphylactic. Although carboplatin triggers cell death at the cancer tissue level, it can also bind to proteins as a hapten and then be presented to T cells, inducing sensitization and expansion of B cells producing IgE that recognizes the hapten. Because women with *BRCA1* and *BRCA2* mutations seem more prone to carboplatin reactions after fewer exposures,[23] it is possible that these genetic mutations facilitate T-cell presentation. The presence of IgE antibodies in carboplatin-sensitized women can be detected in blood and by skin test. The IgE levels may be transient, being lost within 6–12 months of exposure. However, memory T and B cells remain active in these patients, allowing clonal expansion upon reexposure months to years later. Patients with connective tissue disorders have an increased rate of reactivity to mAbs.[24] Similarly, women exposed to endogenous and exogenous progesterone can develop progesterone hypersensitivity.[25]

CLASSIFICATION OF DRUG HYPERSENSITIVITY REACTIONS

Drug reactions can be classified according to whether they occur as a result of the mode of action (on target) or in some other way (off target) (Fig. 48.3).[16] Type A (on-target) reactions account for over 80% of drug reactions. These are predictable, dose dependent, and mirror the pharmacological action of the drug. Type B (off-target) reactions comprise 20% of drug reactions. These are not predictable and may be immunological or nonimmunological. Immunological reactions include B cell–dependent reactions, such as antibody-mediated immediate and immune complex reactions (Gell and Coombs types I and III) and T cell–mediated reactions (type IV). Nonimmunological hypersensitivity reactions include idiosyncratic reactions caused by individual predisposition (*e.g.*, enzymatic defect) and activation of effector cells through off-target receptors, such as the MRGPRX2 mast-cell receptor, a transmembrane G-coupled protein that can be activated by quinolone medications, such as icatibant, and neuromuscular blocking agents, such as atracurium.[26]

Drug hypersensitivity can cause a spectrum of hypersensitivity diseases. These were originally classified by Gell and Coombs into four types: antibody-dependent type I, II and III reactions and T cell–mediated type IV reactions. More recently, in the light of developments in T-cell biology and to take account of

TABLE 48.1 HLA–Associated Drug Hypersensitivity Reactions

Drug	Adverse Drug Reaction	Associated HLA Alleles	Positive Predictive Value	Negative Predictive Value	Populations
Abacavir	HSS	B*57:01	55%	100%	European, African
Carbamazepine	SJS/TEN	B*15:02	3%	100%	Han Chinese, Thai, Malaysian, Indian
		B*15:11			Korean, Japanese
		B*15:18, B*59:01, C*07:04			Japanese
		A*31:01			Japanese, Northern European, Korean
	DRESS	8.1 AH (HLA A*01:01, Cw*07:01, B*08:01, DRB1*03:01, DQA1*05:01, DQB1*02:01)			White
		A*31:01	0.89%	99.98%	European
		A*31:01	0.59%	99.97%	Chinese
		A*31:01			Northern European, Japanese, Korean
		A*11 and B*51 (weak)			Japanese
	MPE	A*31:01	34.9%	96.7%	
Allopurinol	SJS/TEN, DRESS	B*58:01 (or B*58 haplotype)	3%	100% in Han Chinese	Han Chinese, Thai, European, Italian, Korean
Oxcarbazepine	SJS/TEN	B*15:02 and B*15:18			Han Chinese, Taiwanese
Lamotrigine	SJS/TEN	B*15:02 (positive), B*15:02 (no association)			Han Chinese
Phenytoin	SJS/TEN	B*15:02 (weak), Cw*08:01, DRB1*16:02, CYP2C9*3			Han Chinese
	DRESS/MPE	B*13:01 (weak), B*5101 (weak), CYP2C9*3			Han Chinese
Nevirapine	SJS/TEN	C*04:01			Malawian
	DRESS	DRB1*01:01 and DRB1*01:02 (hepatitis and low CD41)	18%	96%	Australian, European, and South African
		Cw*8 or Cw*8-B*14 haplotype			Italian and Japanese
		Cw*4			Black, Asian, White, and Han Chinese
		B*35, B*35:01, B*35:05	16%	97%	Asian
	Delayed rash	DRB1*01			French
		Cw*04			African, Asian, European, and Thai
		B*35:05, rs1576*G CCHCR1 status			Thai
Dapsone	HSS	B*13:01	7.8%	99.8%	
Efavirenz	Delayed rash	DRB1*01			French
Sulfamethoxazole	SJS/TEN	B*38			European
Amoxicillin-clavulanate	DILI	DRB1*15:01, DRB107 (protective), A*02:01, DQB1*06:02, and rs3135388, a tag SNP of DRB1*15:01-DQB1*06:02			European
Lumiracoxib	DILI	DRB1*15:01-DQB1*06:02-DRB5*01:01- DQA1*01:02 haplotype			International, multicenter
Ximelagatran	DILI	DRB1*07 and DQA1*02			Swedish
Diclofenac	DILI	HLA-B11, C-24T, UGT2B7*2, IL-4 C-590-A			European
Flucloxacilin	DILI	B*57:01, DRB1*01:07-DQB1*01:03	0.12%	99.99%	European
Lapatinib	DILI	DRB1*07:01-DQA2*02:01-DQA1*02:01			International, multicenter

DILI, drug-induced liver injury; DRESS, drug-induced reactions with eosinophilia and systemic symptoms; HSS, hypersensitivity syndrome; MPE, maculopapular eruption; SJS, Stevens-Johnson syndrome; SNP, single nucleotide polymorphism; TEN, toxic epidermal necrolysis.

(From White KD, et al. Evolving models of the immunopathogenesis of T cell-mediated drug allergy: the role of host, pathogens, and drug response. J Allergy Clin Immunol. 2015; 136: 219–34.)

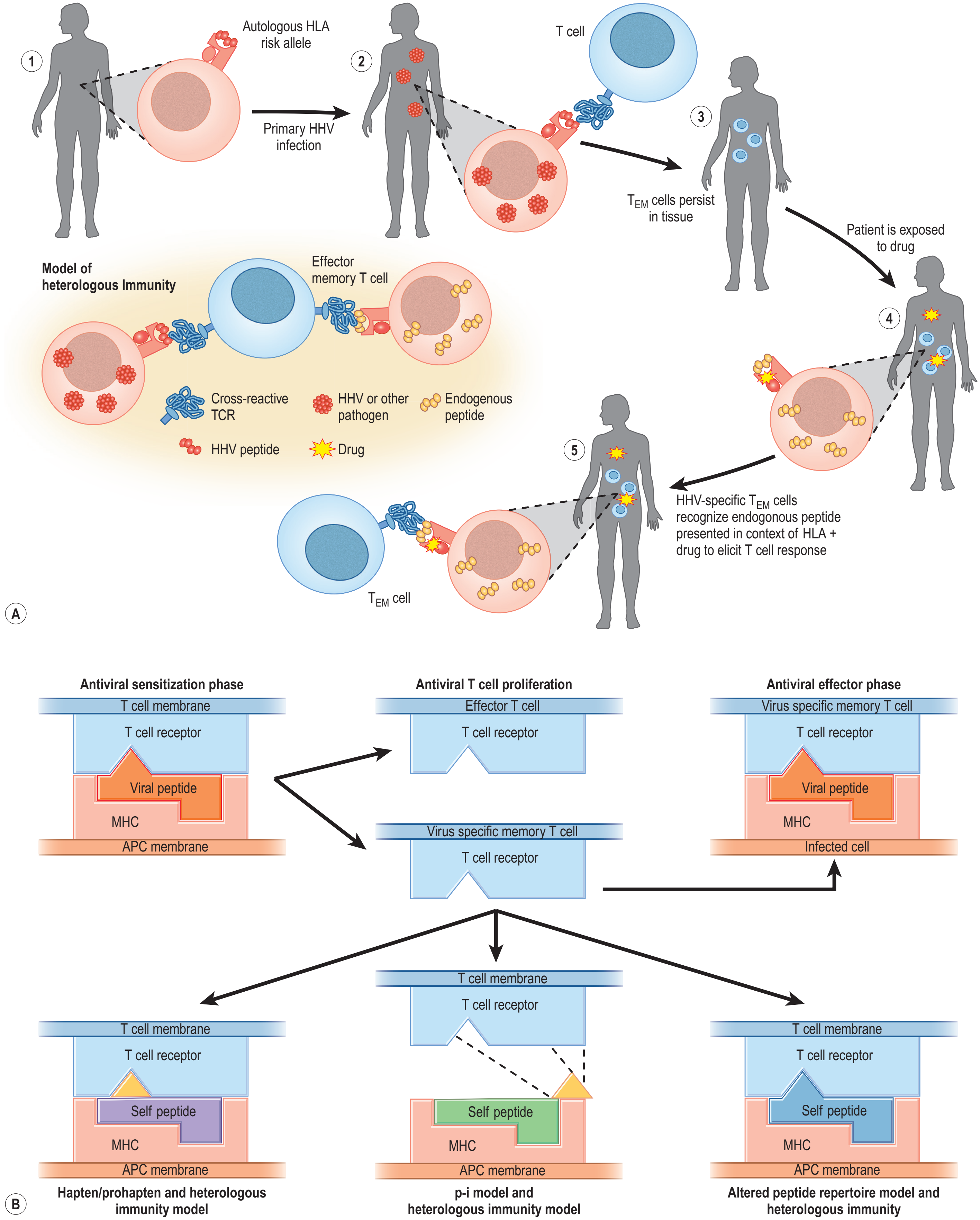

1
Autologous HLA risk allele
Primary HHV infection
2
T cell
3
T_{EM} cells persist in tissue
Patient is exposed to drug
4
Model of heterologous Immunity
Effector memory T cell
Cross-reactive TCR
HHV or other pathogen
Endogenous peptide
HHV peptide
Drug
5
HHV-specific T_{EM} cells recognize endogonous peptide presented in context of HLA + drug to elicit T cell response
T_{EM} cell
A
Antiviral sensitization phase
T cell membrane
T cell receptor
Viral peptide
MHC
APC membrane
Antiviral T cell proliferation
Effector T cell
T cell receptor
Virus specific memory T cell
T cell receptor
Antiviral effector phase
Virus specific memory T cell
T cell receptor
Viral peptide
MHC
Infected cell
T cell membrane
T cell receptor
T cell membrane
T cell receptor
Self peptide
MHC
APC membrane
Self peptide
MHC
APC membrane
T cell membrane
T cell receptor
Self peptide
MHC
APC membrane
B
Hapten/prohapten and heterologous immunity model
p-i model and heterologous immunity model
Altered peptide repertoire model and heterologous immunity

FIG 48.2 Generation of heterologous immune responses that contribute to the pathogenesis of T cell–mediated adverse drug reactions (ADRs). (A) Timeline of the generation of ADRs. (B) Integration of the models of T-cell activation by small molecules and heterologous immunity. (From White KD, et al. Evolving models of the immunopathogenesis of T cell-mediated drug allergy: the role of host, pathogens, and drug response. *J Allergy Clin Immunol.* 2015; 136: 219–34.)

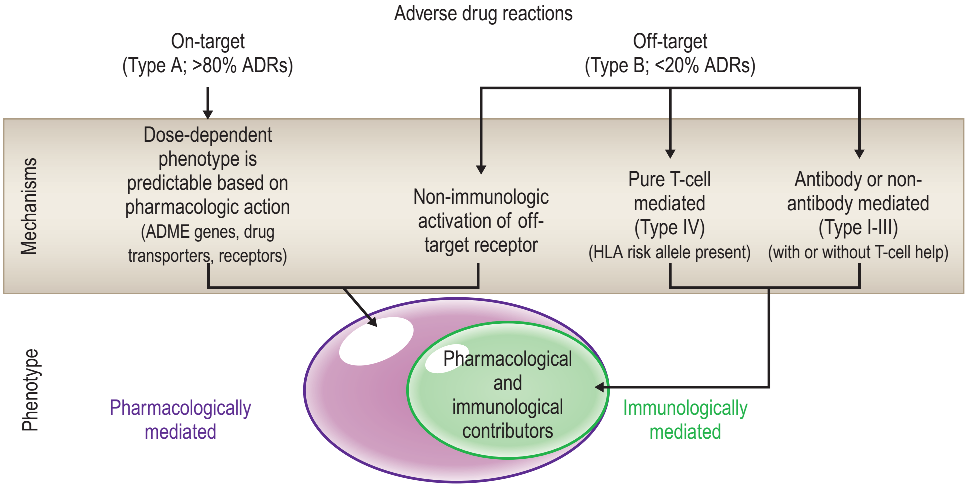

Adverse drug reactions
On-target
(Type A; >80% ADRs)
Off-target
(Type B; <20% ADRs)
Mechanisms
Dose-dependent phenotype is predictable based on pharmacologic action
(ADME genes, drug transporters, receptors)
Non-immunologic activation of off-target receptor
Pure T-cell mediated
(Type IV)
(HLA risk allele present)
Antibody or non-antibody mediated
(Type I-III)
(with or without T-cell help)
Phenotype
Pharmacologically mediated
Pharmacological and immunological contributors
Immunologically mediated

FIG 48.3 Adverse drug reactions can result from either on-target or off-target interactions between the drug and cellular components. (From White KD, et al. Evolving models of the immunopathogenesis of T cell-mediated drug allergy: the role of host, pathogens, and drug response. *J Allergy Clin Immunol.* 2015; 136: 219–34.)

the varied presentation and heterogeneity of T cell–mediated reactions, type IV reactions have been divided into four phenotypes: types IV a, b, c, and d.[27] Because the maturation of B cells and expansion to plasma cells producing Igs depend on the interaction of B cells with T cells in the presence of cytokines and costimulatory molecules, drugs can induce IgE and IgG antibodies and induce type I and type IV reactions in the same patient (Fig. 48.4). Type I reactions result from the release of inflammatory mediators from mast cells and basophils and often present with urticarial (see Fig. 48.4A) and systemic reactions, such as anaphylaxis.

ANTIBODY-MEDIATED DRUG HYPERSENSITIVITY REACTIONS

Type I IgE-Mediated Hypersensitivity

Drugs with hapten-like features bind to soluble and cell-bound proteins, inducing a humoral response, which can lead to IgE antibodies. Chimeric antibodies bearing foreign determinants can also elicit IgE responses. During the sensitization phase at the initial drug exposure, IgE antibodies are generated; these bind to FcεRIα on mucosal and connective tissue mast cells and blood basophils. B cells mature into IgE-secreting cells when activated by T cells in the presence of IL-4/IL13 and CD40/CD40L. The sensitization phase typically requires several days and patients do not develop reactions during the first exposure. Upon re-exposure, crosslinking of FcεRIα receptors on sensitized mast cells and basophils by polyvalent haptenated drug allergens causes cell activation and degranulation and the release of inflammatory mediators that are responsible for the clinical symptoms. Although IgE is produced in small amounts compared with IgG, most IgE is found on tissue mast cells in different organs, including blood vessel walls, accounting for the systemic nature of reactions when mast cells are activated. Small amounts of drug can induce severe systemic reactions, including anaphylaxis, depending on the patient's sensitization and the amount of specific IgE directed against the drug. The relationship between the amount of IgE and the dose of drug allergen can determine the induction of either cell activation or desensitization, activating inhibitory mechanisms.[28] In this type of reaction, symptoms occur within seconds of parenteral drug administration and within minutes of oral administration; both depend on the mediators released from mast cell and basophil granules. Preformed mediators, such as histamine, tryptase and tumor necrosis factor-α (TNF-α) are released first, prostaglandins (PGs) and leukotrienes (LTs) are generated from membrane arachidonic acid within minutes, and newly generated cytokines are released within hours of the initiation of the reaction. Pruritus, urticaria, and angioedema are followed by gastrointestinal (GI) symptoms, such as nausea, vomiting, and diarrhea; hoarseness; shortness of breath and wheezing; and dizziness, which can occur in rapid succession.[29] In highly sensitized individuals, systemic reactions can occur within minutes after intravenous exposure; hypotension, laryngeal edema, and death have occurred following administration of antibiotics and chemotherapeutic agents as a result of cardiovascular collapse or asphyxia. Although skin is the organ most commonly involved, and itching is a typical sign of the initial reaction, ≈60% of anaphylaxis-related deaths are due to asphyxia. Anaphylaxis is diagnosed when two or more organs are affected, when there is a sudden drop of blood pressure after exposure to a known drug allergen, or when there is laryngeal edema (Table 48.2).[30] Atypical symptoms include pain, which has been reported in patients with anaphylactic reactions to taxanes, oxaliplatin, and mAbs, such as rituximab.[31] The mediators responsible for pain have not been identified within mast cells or basophils. It has been suggested that activation of the kinin pathway generating bradykinins may be responsible.[31] The severity of the reactions is graded according to the extent of the symptoms. Reactions confined to the skin or one organ are considered mild and do not require epinephrine treatment. Reactions that involve two organs or are associated with vital sign changes are considered moderate and severe, respectively, and treatment with epinephrine is recommended.[32] Elevated tryptase levels in blood are found in patients with grade 2 and 3 reactions, indicating the involvement of mast cells. Depending on the magnitude of the reaction, tryptase levels return to baseline, usually within days, but sometimes remain elevated for a few weeks after the reaction.[8] Patients with severe anaphylactic reactions induced by drugs

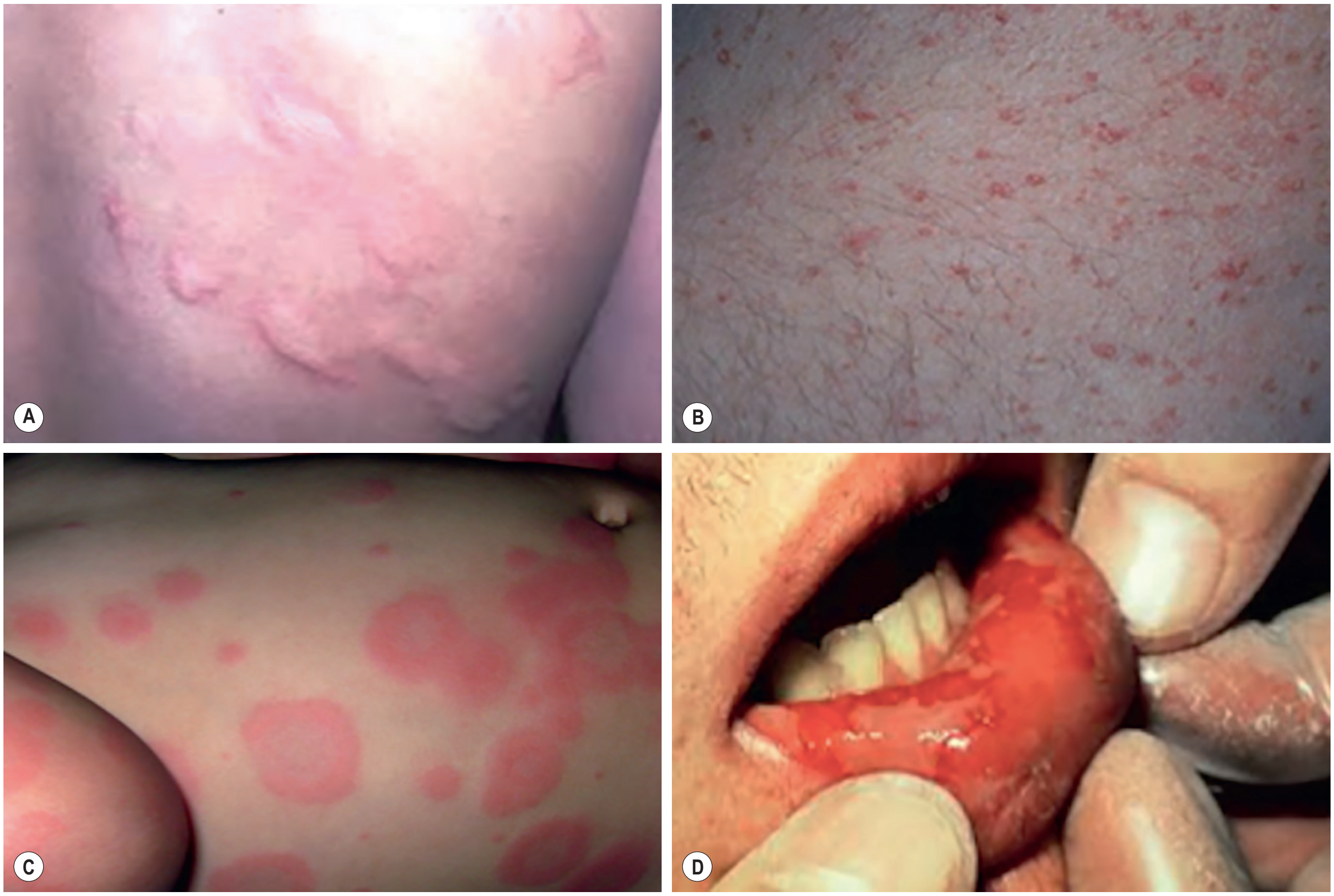


FIG 48.4 Examples of hypersensitivity reactions. (A) Type I: Urticaria. (B) Type IVc: maculopapular rash. (C) Type IVc: erythema multiforme. (D) Type IVc: mucosal involvement (Stevens-Johnson syndrome [SJS])

TABLE 48.2 Severity Grading System of Immediate Type I Hypersensitivity Reactions

Grade	Severity	Description
1	Mild	Symptoms are limited to the skin (*e.g.,* flushing) or involve a single organ/system and are mild (*e.g.,* mild back pain).
2	Moderate	Symptoms involve at least two organs/systems (*e.g.,* flushing and dyspnea), but there is no significant decrease in blood pressure or oxygen saturation.
3	Severe	Symptoms typically involve at least two organs/systems, and there is a significant decrease in blood pressure (systolic ≤90 mm Hg and/or syncope) and/or oxygen saturation (≤92%).

(Adapted from Brown SGA, et al. Clinical features and severity grading of anaphylaxis. J Allergy Clin Immunol. 2004; 114: 371–376.)

should be evaluated for mastocytosis if their tryptase levels do not normalize after the reaction. Common causes of IgE-mediated allergic drug reactions include: beta-lactam antibiotics (penicillins and cephalosporins); chemotherapeutic agents, such as carboplatin, cisplatin, and oxaliplatin; and both chimeric and nonchimeric mAbs.[33] Cross-reactivity occurs with beta-lactam medications; patients sensitized to penicillins can be subject to reactions on first exposure to cephalosporins because these antibiotics share epitopes that can bind beta-lactam–specific IgE. Patients with cancer who are exposed to one platin, such as carboplatin, can react on first exposure to another platin, likely as a result of shared cross-reactive epitopes.[34] Not all patients reacting with symptoms of type I hypersensitivity to taxanes, such as paclitaxel, have evidence of IgE sensitization. Reactions to cremophor, the diluent of paclitaxel, have been implicated in some reactions.[35]

Non–IgE-Mediated Type I Hypersensitivity Reactions

The definition of type I hypersensitivity reactions has been expanded to include reactions with similar symptoms without evidence of IgE. These are caused by the presence of multiple activating receptors on mast cells and basophils. Examples include muscle relaxants and quinolones, which can induce anaphylactic reactions with release of mast cell mediators, including tryptase, without evidence of IgE sensitization,[36] indicating that mechanisms other than IgE-mediated degranulation can trigger these effector cells. One candidate for the receptor involved is the G-coupled receptor MRGPRX2.[26] These reactions were previously called "pseudoallergic reactions." The symptoms are identical to IgE-mediated reactions and include hives, angioedema, respiratory distress, GI symptoms, and hypotension; epinephrine is required if the reactions are moderate or severe. Vancomycin can induce histamine release from mast cells without IgE involvement, in a dose-dependent fashion.

CLINICAL PEARLS

Vancomycin Red Man Syndrome

- Presents as erythema, flushing, rarely hypotension
- Histamine elevated in urine and blood
- Direct mast cell activation and release of histamine
- Can be potentiated by mast cell activation agents (*e.g.,* opioids)
- Treatment: slow infusion, dose reduction, and antihistamines

Type II Reactions: IgG-Mediated Cytotoxic Reactions

Type II and type III reactions result from the formation of complement-fixing IgG antibodies, typically of the IgG1 and IgG3 subclasses, which bind to Fcγ receptors (FcγRI, IIa, and IIIa) on macrophages, natural killer (NK) cells, granulocytes, and platelets, or form immune complexes and interact with complement receptors on these cells. In type II reactions, the antibody is directed against antigens on the cell membrane, or immune complex activation occurs on the cell surface. Both events can lead to cell destruction or sequestration; affected cells include erythrocytes, leukocytes, platelets, and bone marrow precursor cells. The development of type II reactions requires an IgG immune response to a drug hapten–carrier complex after high-dose and prolonged treatment, as is seen in penicillin and cephalosporin reactions, where the reaction is caused by complement fixation. IgM antibodies are sometimes implicated, as well as autoantibodies reactive to the carrier molecule. These immune reactions can persist after cessation of the drug. Another type of type II reaction occurs when the drug or its metabolites adhere to the erythrocyte or platelet surface, creating a new antigenic complex with the cell membrane. Quinine-induced thrombocytopenia occurs when IgG or IgM antibodies react with membrane epitopes, such as GPIIb/IIIa (fibrinogen receptor) or GPIb/IX (von Willebrand factor receptor); this only happens when the drug in present in soluble form.[37] This interaction results in platelet destruction and thrombocytopenia. Similar reactions occur with sulfonamides and quinidine. The antibody-coated cells can be sequestered in the liver and spleen by Fcγ or complement receptor binding. Rarely, intravascular destruction occurs via complement-mediated lysis.

Penicillin, cephalosporins, levodopa, methyldopa, quinidine, and some antiinflammatory drugs can induce hemolytic anemia, which presents with fatigue, pallor, shortness of breath, tachycardia, and dark urine.[38] Direct and indirect Coombs tests are positive, unconjugated levels of bilirubin are elevated, haptoglobin is decreased, and hemoglobinuria is present. Thrombocytopenia can occur with quinine, quinidine, sulfonamide antibiotics, and mAbs.[39] Drug-induced thrombocytopenia usually occurs after 5–8 days of exposure but may occur after a single exposure in a sensitized patient. The typical presentation is petechial hemorrhages in skin and mucosal bleeding, sometimes associated with GI and urinary bleeding. Intracranial bleeding is rare, and the platelet count returns to normal 3–5 days after the drug is discontinued. In heparin-induced thrombocytopenia, FcγRIIa receptors on platelets bind IgG and IgM immune complexes of heparin and platelet factor 4 (PF4), a CXC chemokine usually stored in platelet α-granules. About 50% of patients treated for >7 days with heparin develop antibodies against heparin and PF4. These immune complexes activate platelets, releasing PF4 and causing platelet destruction. Mild thrombocytopenia can occur in up to 5% of patients treated with heparin, and 10% develop paradoxical thrombosis, which can be life-threatening.

Type III Reactions

Immune complexes form during normal immune responses without any clinical consequences. Immune complexes can also be formed during treatment with drugs that form hapten–carrier complexes or with chimeric or humanized proteins, such as mAbs. These immune complexes are typically cleared by binding to the FcγRI or complement receptor CR1 on reticuloendothelial cells. Decreased clearance because of high levels complement defects or aberrant FcγRI function can lead to type III reactions. Recently a low copy number of FcγRIII was shown to be associated with glomerulonephritis. Type III reactions can present as small vessel vasculitis and/or serum sickness. Serum sickness was first described when heterologous serum or foreign serum was used for passive immunization. Antibodies appear in 4–10 days and react with drug antigens, forming soluble circulating immune complexes. These immune complexes capture complement and are deposited in postcapillary venules, attracting leukocytes by interacting with FcγRIII, eventually inducing the release of proteolytic enzymes and tissue damage. Nonprotein drugs, such as cefaclor, trimethoprim-sulfamethoxazole, cephalexin, amoxicillin, NSAIDs, and diuretics, are the most commonly implicated drugs, but these reactions can also occur with protein drugs, such as mAbs.

The incidence of hypersensitivity vasculitis is 10–30 cases per million per year. The presentation can include fever, skin rash, myalgia, arthralgia, and lymphadenopathy. Palpable purpura may be present in legs, with lesions that can coalesce and ulcerate. The GI tract, kidneys, and joints can be affected: the prognosis depends on the extent of systemic involvement. Patients taking systemic steroids may present with a milder form of reaction, with fewer manifestations. Renal biopsy reveals IgA-containing immune complexes, similar to those found in IgA nephropathy.

Type IV Reactions

Cell-mediated reactions are now subdivided on the basis of the cytokines produced by T cells and the type of effector immune cells induced by these cytokines, such as eosinophils and neutrophils.

Type IVa reactions correspond to classic T helper 1 (Th1)–type immune reactions, in which Th1 cells activate macrophages by secreting interferon-γ (IFN-γ) and other cytokines (TNF-α, IL-12). These Th1 cells drive the production of the complement-fixing antibodies involved in types II and III reactions and are costimulators of pro-inflammatory responses, and CD8 T-cell responses. Examples of type IVa reactions include tuberculin reactions, contact dermatitis (with activated CD8 cells) and sarcoidosis (with granuloma formation and monocyte activation).

Type IVb reactions correspond to type 2 immune responses with IL-4, IL-13, and IL-5 cytokines, which promote IgE- and IgG4-producing B cells, mast cells, and eosinophils. IL-5 production leads to eosinophilic inflammation, which is seen in many drug-induced hypersensitivity reactions, including DRESS. Type I reactions are linked to these reactions since IgE is produced as a result of IL-4 and IL-13 secretion by Th2 cells. Maculopapular

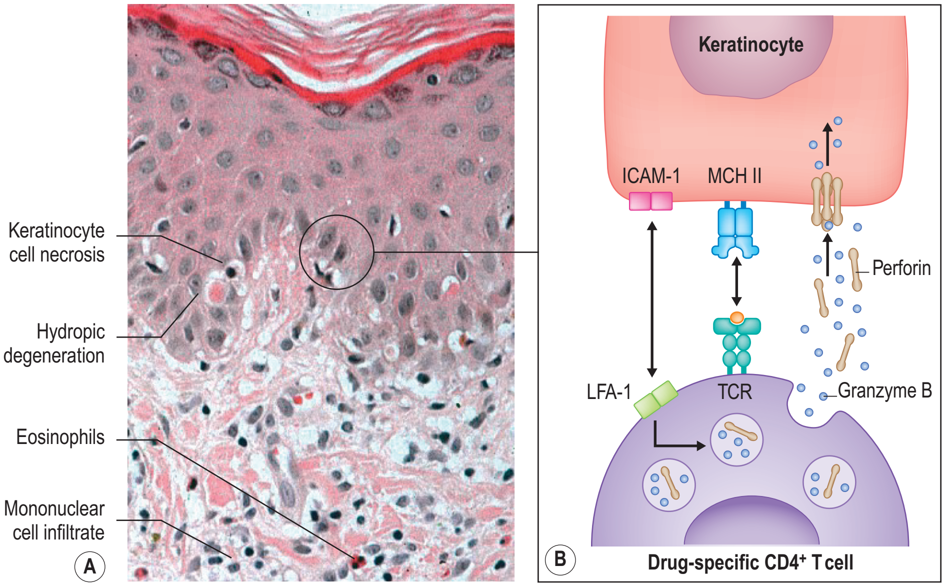


FIG 48.5 (A) Typical histology of a maculopapular exanthema (MPE). Note the focal keratinocyte necrosis, often in close apposition of T cells (which have cytotoxic potential and are perforin-positive). (B) Schematic representation of CD4-mediated killing of activated keratinocytes, which express MHC class II and intercellular adhesion molecule 1 (ICAM1). (Modified from Pichler WJ. Delayed drug hypersensitivity reactions. *Ann Intern Med.* 2003; 139: 683–693.)

exanthema (MPE) (Fig. 48.5), infections with nematodes, asthma, and rhinitis are considered part of this group of reactions.

KEY CONCEPT

Drug-Induced Reactions With Eosinophilia and Systemic Symptoms (DRESS) Syndrome

- Drug-induced
  - Anticonvulsants, antidepressants, non-antimicrobial sulfones/sulfonamides, nonsteroidal antiinflammatory drugs (NSAIDs), angiotensin-converting enzyme (ACE) inhibitors, beta-blockers, and antibiotics
- Systemic disorder
  - Hematological abnormalities (eosinophilia, atypical lymphocytes)
  - Multiorgan involvement
    - kidney, liver, heart, lung, thyroid
  - Hypogammaglobulinemia
  - Lymphadenopathy
  - Skin involvement

Type IVc reactions result from T-cell migration to tissues and direct damage or killing of tissue cells, such as hepatocytes, through perforin-, granzyme B– and Fas-ligand–dependent mechanisms, with the recruitment of other inflammatory cells, such as monocytes, eosinophils, and neutrophil polymorphonuclear leukocytes. These reactions present as delayed hypersensitivity reactions, such as SJS, TEN, and contact dermatitis (Fig. 48.4B, C and D).

Type IVd reactions occur when T cells activate sterile neutrophilic inflammation, as in acute generalized exanthematous pustulosis (AGEP). T cells producing CXCL8 and granulocyte macrophage–colony-stimulating factor (GM-CSF) recruit leukocytes, and the cytokines then prevent their apoptosis. Behçet disease, linear IgA bullous dermatosis, and pustular psoriasis are examples of type IVd reactions (Fig. 48.6).

KEY CONCEPT

Stevens-Johnson Syndrome (SJS) and Toxic Epidermal Necrolysis (TEN)

SJS
- Detachment <10% of body surface area (BSA), plus
- Widespread macules or flat atypical target lesions

Overlap SJS-TEN
- Detachment 10–30% of BSA, plus
- Widespread macules or flat atypical target lesions

TEN
- Detachment of >30% of BSA, plus
- Widespread macules or flat atypical target lesions
- Detachment >10% of BSA with large epidermal sheets and without any macules or target lesions

DIAGNOSIS OF DRUG HYPERSENSITIVITY: CLINICAL SYMPTOMS AND HISTORY

A detailed history is the most important element in the diagnosis of all drug reactions; it is necessary to determine the type of reaction and the test required to confirm the diagnosis and to establish management and treatment plans. Type I IgE- and non–IgE-mediated reactions occur with a rapid onset, within minutes of drug exposure, which can present as pure skin manifestations or as involvement of other organs. Sudden onset of hypotension can be the first manifestation during anesthesia or when infusing high doses of chemotherapy drugs, such as taxanes. Tryptase levels help determine the mechanism of the reaction and its extent. The basophil activation test (BAT) has been performed for several drugs yielding different results;

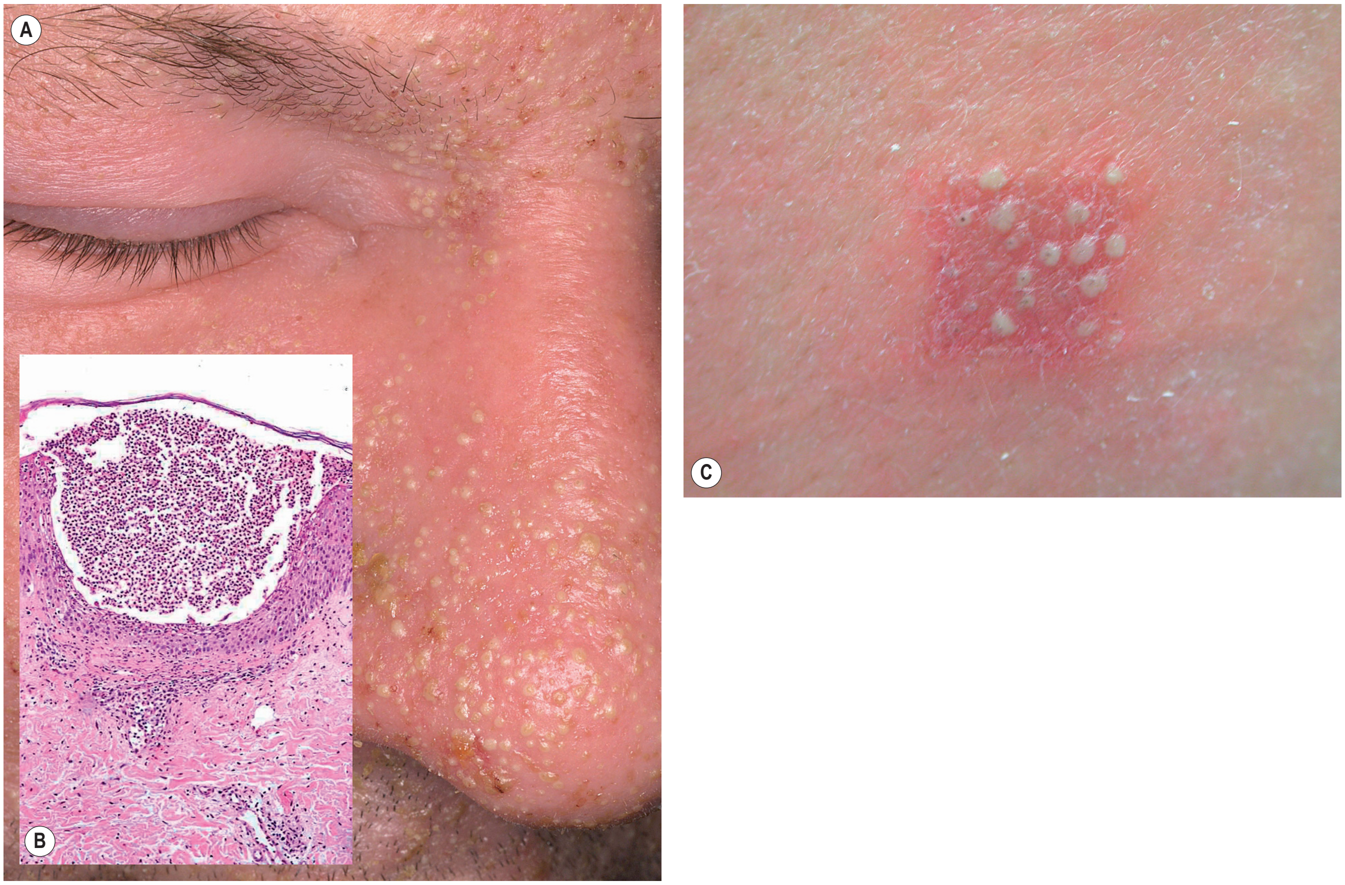


FIG 48.6 Clinical pictures and immunohistologies of drug-induced exanthema. (A) Pustular drug eruption (acute generalized exanthematous pustulosis [AGEP]). (B) Intraepidermal, nonfollicular pustules. (C) A patch test reaction leading to a pustular reaction. (From Schaerli P, Britschgi M, Keller M, et al. Characterization of human T cells that regulate neutrophilic skin inflammation. *J Immunol.* 2004; 173: 2151–8.)

the test is limited by the patient's number of basophils and premedications because basophils undergo apoptosis in the presence of steroids. BAT seems to be a promising test for diagnosis of platin hypersensitivity.

Delayed-type IV reactions are characterized by skin rashes. The differential diagnosis includes viral exanthema, other infections, food allergy, and graft-versus-host reactions. Severe reactions, such as DRESS, SJS, and TEN, are associated with blood eosinophilia, leukocytosis, and biochemical changes, such as elevated liver enzymes. Evidence of systemic involvement should be determined with whole-body skin evaluation, and the presence of fever, lymphadenopathy, and organomegaly should be noted. Nikolsky sign should be elicited because its presence can indicate a severe bullous condition. Danger signs without skin involvement, such as malaise, fever, photophobia, abdominal pain, and alterations in GI and urinary function, can be early signs of severe drug-induced syndromes. Type 1 reactions occur within an hour of exposure, but in patients receiving premedication with antihistamines and steroids, onset of urticaria or hypotension may take several hours. Type IV reactions are typically delayed and appear several hours after drug exposure. However, in highly sensitized individuals, symptoms can occur within 1–2 hours.

Diagnostic tools to identify the culprit drug are limited. They include skin testing and patch testing, but these tests are not standardized and have variable reliability.

Tryptase

Activation of mast cells during type I IgE and non-IgE reactions leads to release of granule mediators, including tryptase, which can be measured in serum. After secretion from tissue mast cells, tryptase requires 30 minutes to reach peripheral blood. Its levels peak in 2–4 hours after an anaphylactic event, but in patients with marked elevations above the normal range (11.4 ng/mL), the levels may remain elevated for >6–8 hours and in extreme cases for several days.

Skin Testing, Specific IgE, and Challenges

Skin testing for penicillin has been available for several years and its predictive value is >97%, meaning that patients with negative skin tests are not at risk for anaphylaxis. Patients sensitive to minor determinants may not be identified by current reagents in the United States because none of the tests for minor determinants has been approved. Skin testing for cephalosporin sensitivity has not been standardized but non-irritant concentrations are available for some of them (Table 48.3) as well as for other antibiotics. The sensitivity and specificity of these tests have geographical differences, possibly because of different rates of exposure. Skin testing to chemotherapy and mAbs can be done when an IgE-mediated reaction is suspected, ideally 2–4 weeks after the reaction to minimize false-negative results caused by natural desensitization. The

TABLE 48.3 Nonirritating Concentrations for 15 Commonly Used Antibiotics

Antimicrobial Drug	Full-Strength Concentration	NIC (as Dilution From Full-Strength Concentration)	No. of Patients Tested
Cefotaxime	100 mg/mL	10^{-1}	25
Cefuroxime	100 mg/mL	10^{-1}	25
Cefazolin	330 mg/mL	10^{-1}	25
Ceftazidime	100 mg/mL	10^{-1}	25
Ceftriaxone	100 mg/mL	10^{-1}	30
Tobramycin	40 mg/mL	10^{-1}	25
Ticarcillin	200 mg/mL	10^{-1}	25
Clindamycin	150 mg/mL	10^{-1}	25
Gentamicin	40 mg/mL	10^{-1}	30
Cotrimoxazole	80 mg/mL	10^{-2}	25
Levofloxacin	25 mg/mL	10^{-3}	25
Erythromycin	50 mg/mL	10^{-3}	25
Azithromycin	100 mg/mL	10^{-4}	30
Nafcillin	250 mg/mL	10^{-4}	25

(From Empedrad R, et al. Nonirritating intradermal skin test concentrations for commonly prescribed antibiotics. *J Allergy Clin Immunol.* 2003; 112: 629–630.)

negative predictive value (NPV) for carboplatin is very high with recent exposure, and patients with negative skin test results do not develop anaphylaxis,[24] but the NPV is not known for most mAbs. In a study of patients who needed to be desensitized to trastuzumab, infliximab, or rituximab, only 13 of 23 patients had positive skin test results.[40]

Drug challenges are considered the gold standard for type I reactions and should be performed to confirm tolerance in all patients with negative skin test results. Such challenges can trigger mild, acute, and/or delayed reactions but are considered extremely safe. Cofactors, such as viral infections, can influence these tests, and a patient with negative results of the skin test and challenge can present a recurrent skin rash if he or she takes the same drug during a viral infection. Patients with a history of mild-to-moderate reactions and a remote history of drug allergy, who have negative skin test results, can be safely challenged without an increased risk of anaphylaxis.[41]

Patch Testing, Delayed Intradermal Reading, and Lymphocyte Transformation Tests

Skin testing with percutaneous and intradermal drug injections can provoke a maculopapular reaction at the site of injection in 24–96 hours, mimicking the systemic reaction induced by the drug. These tests have been used to determine which drugs induce the various type IV reactions, such as MPE.[42] Patch testing, with application of the drug allergens to stripped skin after absorption of superficial T cells and other effector cells, is used to identify drugs responsible for contact dermatitis and other type IV reactions. Commercial reagents are available for common contact allergens and provide high specificity for metal allergies, such as nickel allergy. However, for some diseases, such as MPE, DRESS, and AGEP, the sensitivity is considered low. The lymphocyte transformation test relies on the activation and proliferation of T cells cultured in the presence of drug allergens. Reactivity is measured by ^{3}H-thymidine incorporation after several days in culture,[43] or by enzyme-linked immunospot (ELISPOT), 5(6)-carboxyfluorescein N-hydroxysuccinimidyl ester (CSFE) staining, or CD69 upregulation. These tests may prove valuable in the future but have not yet been standardized.

SPECIFIC DRUG HYPERSENSITIVITY

KEY CONCEPT

Drug Allergy

General Evaluation of Drug Allergy

- **History:** Indication for drug use, association with viral/bacterial infection
- **Physical examination**
- **Blood count/differential, Liver function tests**
- **Serum tryptase**
- **Skin testing**
- **Patch testing/Delayed reading of skin test**
- **Specific IgE, basophil activation test**
- **Genotyping**

Specific Evaluation of Drug Allergy

- **Name** of the drug, ingredients, preservatives
- **First exposure**
- **How long ago** did the reaction occur?
- **Re-exposure:** Has the patient been exposed to the drug or a related drug?
- **Other drugs** administered at the same time
- **Associated narcotics**
- **Symptoms and signs** of the reaction
- **Timing of symptoms** relative to the drug exposure

Underlying Condition for which the Drug was Prescribed

- Similar symptoms unrelated to the drug exposure (urticaria)
- Treatment and response to treatment (epinephrine)
- Timing of resolution

Beta-Lactams

Penicillins, aminopenicillins, and cephalosporins can induce type I, II, III, and IV reactions and are the medications most often implicated in drug hypersensitivity. These drugs are small chemical compounds with a common beta-lactam core and different side chains (Fig. 48.7). Drug antigens are created through hydrolysis when in contact with blood or other proteins. Penicilloyl is the dominant (or major) drug allergen responsible for most type I reactions. Minor components include penilloate and penicilloate, which have been implicated in severe anaphylactic reactions. Patients with a history of suspected penicillin allergy can be evaluated by skin testing with major and minor determinants. Over 90% of individuals with a history of penicillin allergy have no evidence of IgE sensitization and can safely be reexposed to beta-lactams with a very low rate of adverse reaction upon reexposure. This contrasts with chemotherapy drugs, such as carboplatin.[34] Once a patient has experienced a reaction, subsequent reexposure typically induces a reaction. Although both medications are haptens and probably trigger T-cell activation and B-cell expansion to produce IgE antibodies in a similar fashion, there are some important differences. The route of administration of chemotherapy drugs is typically intravenous, the doses are high, and the interval times are short, with repetitive doses administered over few weeks to months. In contrast, beta-lactam antibiotics are used for short periods during acute infections, so memory T and B cells may not be reactivated during reexposure. If taken orally, repeated exposure to beta-lactams may generate tolerance instead of allergy. Patients who have a history of penicillin allergy and are admitted to hospital have increased risks of complications from use of second-line antibiotics, with increased days of hospitalization and health care costs.[5]

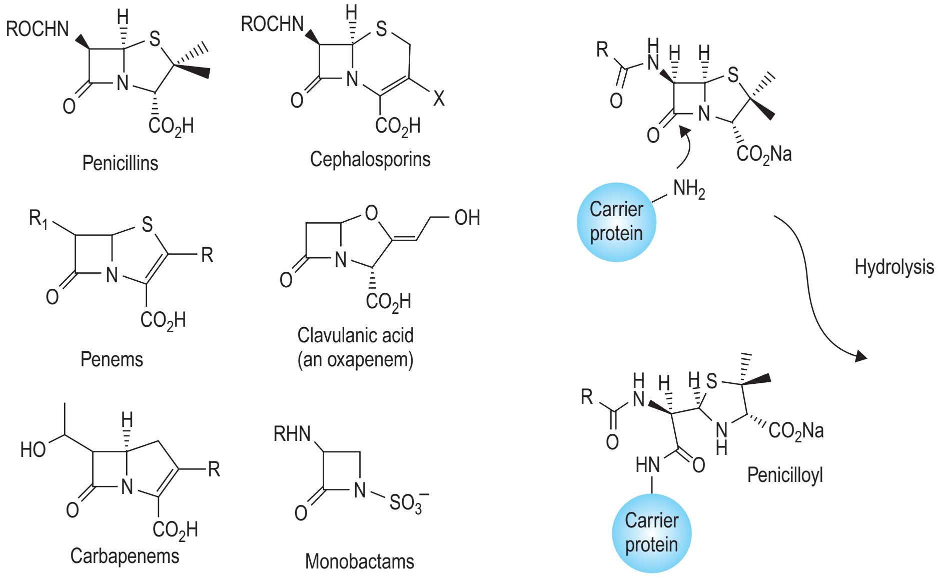


FIG 48.7 Drug antigens: beta-lactams.

The incidence of allergy to specific side-chain epitopes in beta-lactam compounds has increased in the last 20 years as a result of increased use of aminopenicillins and cephalosporins.[44]

Radio Contrast Media

Reactions to iodinated and noniodinated contrast media are relatively rare, and their mechanisms are not well understood, although premedication can protect affected patients. Some of these reactions are thought to be caused by activation of the complement pathway with generation of the anaphylatoxins C3a and C5a, which can bind to complement receptors on mast cells and induce mediator release. The newer nonionic, low-osmolality contrast media products can also be involved in type I reactions; in a few cases, positive skin test results suggest an IgE-dependent mechanism.[45]

Perioperative Anaphylaxis

Patients who have anaphylactic reactions during surgery are at risk for reactions during future operations. Evaluating the drug responsible for these reactions is complex because multiple drugs are used in patients undergoing surgery. Having an accurate account of the initial symptoms of the reaction and changes in vital signs and blood pressure in relation to the timing of each drug used is key to an accurate diagnosis. Neuromuscular-blocking agents, antibiotics, and latex are the most common causes of reactions, with cephalosporins often being implicated. In most cases, skin testing will identify the culprit drug, allowing safe drugs to be chosen for future procedures.[47]

Local Anesthetics

Reactions to local anesthetics are rare, and most are not IgE-mediated. Idiosyncratic reactions include the proarrhythmic effects of lidocaine and other amide drugs. There is no cross-reactivity between benzoic ester and amide drugs, so patients who have reacted to one group generally tolerate drugs from the other group (Table 48.4).

Sulfonamides

Delayed reactions to sulfonamide antibiotics are common, but all types of reactions have been seen (Fig. 48.8A). Delayed reactions range from benign maculopapular rashes to SJS/TEN. Patients with HIV[47] and cystic fibrosis are at high risk. Sulfamethoxazole metabolism and protein adduct formation stimulates T cells and induces delayed reactions. Nonantibiotic sulfonamides have a similar chemical structure, but their stoichiometry is different because of a shared arylamine group being in a different position.[48] In practice, patients with sulfonamide antibiotic hypersensitivity can be safely exposed to non-antibiotic sulfonamides.[49]

TABLE 48.4 Local Anesthetics

Benzoic Acid Esters	Amides and Other
Benzocaine	Bupivacaine
Butamben picrate	Dibucane
Chloroprocaine	Duclonine
Procaine	Etiodocaine
Proparacaine	Levobupivacaine
Tetracaine	Lidocaine
Cocaine	Mepivacaine
	Prilocaine
	Ropivacaine

Aspirin

Reactions to NSAIDs can induce respiratory and skin reactions as well as anaphylaxis. Up to 10% patients with asthma may have aspirin-exacerbated respiratory disease and react to aspirin and other NSAIDs. They typically have nasal polyps, anosmia, and severe asthma. The pathophysiology of these reactions involves decreased production of PGE_2 and increased production of LTs, which are found in urine and other secretions at baseline and after aspirin challenge. Aggregates of leukocytes and platelets are observed in the peripheral blood of these patients. It is possible that platelets contribute to the pathogenesis of the syndrome. Patients with aspirin-exacerbated respiratory disease (AERD) cannot tolerate any COX-1 inhibitor but can tolerate COX-2 inhibitors indicating a common mechanism of action for all NSAIDs (Fig. 48.9).[50] Urticaria and angioedema can be induced by aspirin and NSAIDs; patients can also present

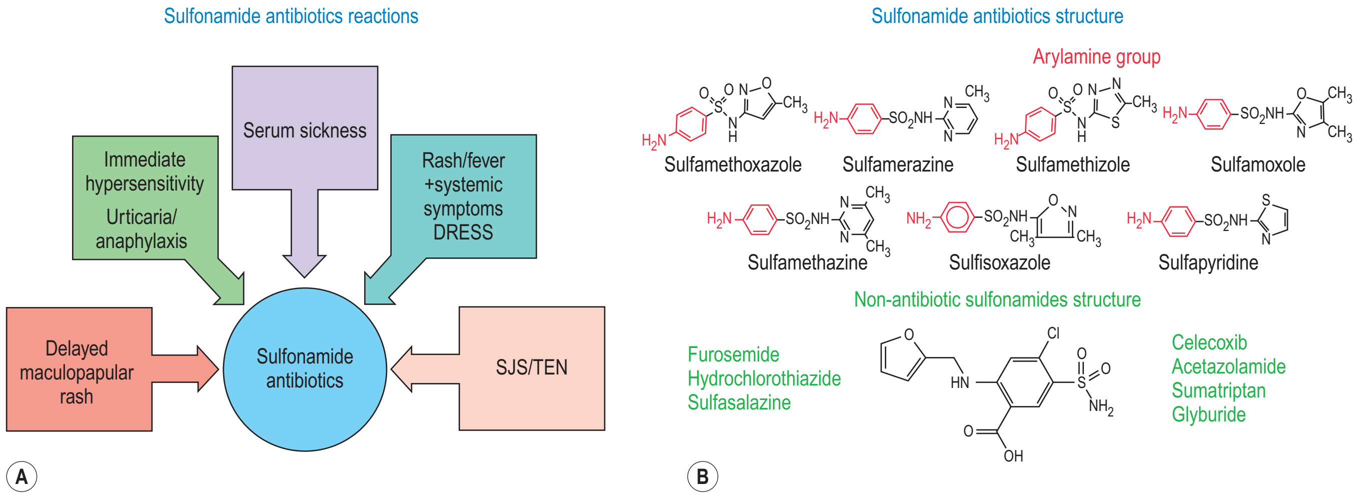

Sulfonamide antibiotics reactions
Serum sickness
Immediate hypersensitivity
Urticaria/ anaphylaxis
Rash/fever +systemic symptoms DRESS
Delayed maculopapular rash
Sulfonamide antibiotics
SJS/TEN
A
Sulfonamide antibiotics structure
Arylamine group
Sulfamethoxazole
Sulfamerazine
Sulfamethizole
Sulfamoxole
Sulfamethazine
Sulfisoxazole
Sulfapyridine
Non-antibiotic sulfonamides structure
Furosemide
Hydrochlorothiazide
Sulfasalazine
Celecoxib
Acetazolamide
Sumatriptan
Glyburide
B

FIG 48.8 Reactions to and structures of sulfonamide antibiotics (A). Structures of sulfonamide antibiotics (B). (Adapted from Tilles, SA, Slatore, CG. Hypersensitivity reactions to non-beta-lactam antibiotics. *Clin Rev Allergy Immunol.* 2003; 24: 221–8.)

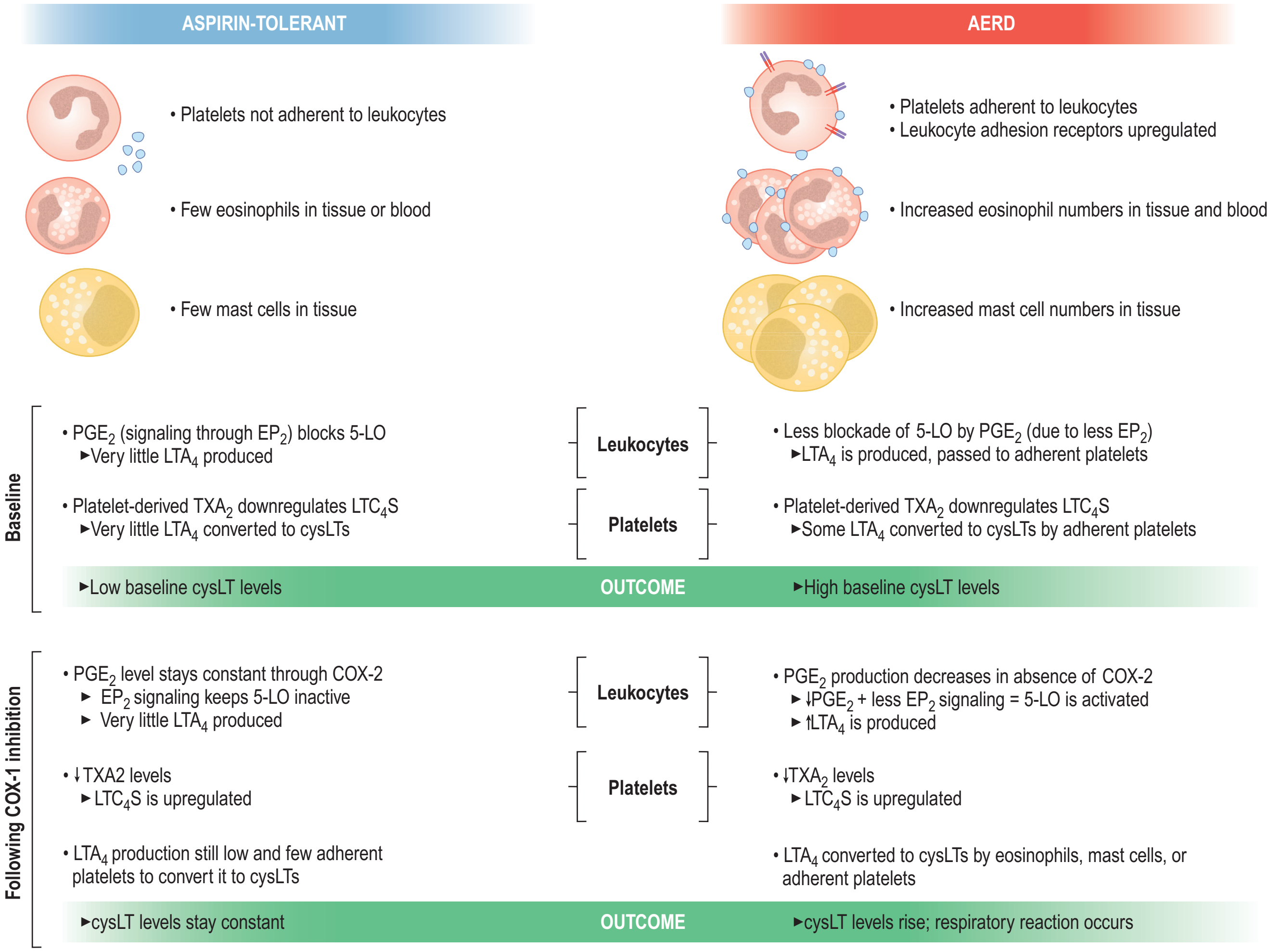

ASPIRIN-TOLERANT
AERD
• Platelets not adherent to leukocytes
• Platelets adherent to leukocytes
• Leukocyte adhesion receptors upregulated
• Few eosinophils in tissue or blood
• Increased eosinophil numbers in tissue and blood
• Few mast cells in tissue
• Increased mast cell numbers in tissue
Baseline
• PGE_2 (signaling through EP_2) blocks 5-LO
►Very little LTA_4 produced
Leukocytes
• Less blockade of 5-LO by PGE_2 (due to less EP_2)
►LTA_4 is produced, passed to adherent platelets
• Platelet-derived TXA_2 downregulates LTC_4S
►Very little LTA_4 converted to cysLTs
Platelets
• Platelet-derived TXA_2 downregulates LTC_4S
►Some LTA_4 converted to cysLTs by adherent platelets
►Low baseline cysLT levels
OUTCOME
►High baseline cysLT levels
Following COX-1 inhibition
• PGE_2 level stays constant through COX-2
► EP_2 signaling keeps 5-LO inactive
► Very little LTA_4 produced
Leukocytes
• PGE_2 production decreases in absence of COX-2
►↓PGE_2 + less EP_2 signaling = 5-LO is activated
►↑LTA_4 is produced
• ↓TXA2 levels
►LTC_4S is upregulated
Platelets
• ↓TXA_2 levels
►LTC_4S is upregulated
• LTA_4 production still low and few adherent platelets to convert it to cysLTs
• LTA_4 converted to cysLTs by eosinophils, mast cells, or adherent platelets
►cysLT levels stay constant
OUTCOME
►cysLT levels rise; respiratory reaction occurs

FIG 48.9 Summary of the pathogenesis of aspirin-exacerbated respiratory disease (AERD) at baseline and after cyclooxygenase 1 (COX-1) inhibition. PGE_2, prostaglandin E2; 5-LO, 5-lipoxygenase; LTA_4, leukotriene A4; LTC_4S, leukotriene C4 synthase; cysLTs, cysteinyl leukotrienes; COX-2, cyclo-oxygenase-2; TXA_2, thromboxane A2; EP_2, prostaglandin E2 receptor 2. (From Laidlaw TM, Boyce JA. Pathogenesis of aspirin-exacerbated respiratory disease and reactions. *Immunol Allergy Clin North Am.* 2013; 33: 195–210.)

with exacerbation of chronic urticaria. Anaphylaxis can occur to a single NSAID with tolerance to all others, but usually there is complete cross-reactivity. Desensitization protocols are available for AERD to ameliorate nasal polyps and increase the sense of smell. Desensitization has also been tried for urticarial reactions.

Biological Agents and Monoclonal Antibodies

The use of mAbs to target cancer and chronic inflammatory diseases has become widespread over the past decade. Inevitably this has led to adverse reactions, which sometimes prevent the use of first-line therapies. Some of the most frequently used mAbs are presented in Table 48.5, including their targets, incidence of overall injection/infusion site reactions, and rates of severe immediate hypersensitivity reactions (HSRs).[51] The immunogenicity of mAbs varies, depending on whether they are chimeric mouse–human, humanized, or fully human mAbs, such as adalimumab and ofatumumab. Reactions to mAbs can occur during the first exposure, as has been reported with cetuximab and trastuzumab or after multiple exposures, such as with rituximab. Infusion-related reactions to mAbs present as cytokine storm-like reactions, with such symptoms as nausea, chills, fever, and malaise. These are thought to be caused by release of proinflammatory cytokines (*e.g.*, IL-1, IL-6, and TNF-α) and respond to NSAIDs and steroids. Hypersensitivity reactions to cetuximab on first exposure have been attributed to sensitization to the galactose-α-1,3-galactose epitope caused by exposure to the lone star tick (*Amblyomma americanum*). The carbohydrate galactose-α-1,3-galactose is expressed on nonprimate mammalian proteins and also on the cetuximab heavy chain.

TABLE 48.5 Biological Agents: Actions, Incidence, and Hypersensitivity Drug Reactions

Drug	Target	Overall Reactions	HSR
Rituximab (Rituxan) IV	CD20	77% (first infusion)	5–10%*
Ofatumumab (Arzerra) IV	CD20	44 % (first infusion) 67% (combination therapy)	2%*
Obinutuzumab (Gazyva) IV	CD20	66%	*
Trastuzumab (Herceptin) IV	HER-2	40% (mild; first infusion)	0.6-5%*
Cetuximab (Erbitux) IV	EGFR	15–21%	1.1–5% 14–27% (Southern USA)*
Tocilizumab (Actemra) IV	IL-6 receptor	7–8%	0.1-0.7%*
Infliximab (Remicade) IV	TNF-α	5- 18%	1%*
Etanercept (Enbrel) SC	TNF-α	15- 37%	<2%*
Adalimumab (Humira) SC	TNF-α	20%	1%*
Golimumab (Simponi) SC	TNF-α	4–20%	Not reported
Certolizumab (Cimzia) SC	TNF-α	0.8–4.5%	Not reported
Brentuximab (Adcetris) IV	CD30	12%	*
Bevacizumab (Avastin) IV	VEGF-A	<3%	*
Omalizumab (Xolair) SC	IgE	45%	00.9–0.2%*

*Case reports of anaphylaxis.

(Modified from Galvão VR, et al. Hypersensitivity to biological agents-updated diagnosis, management, and treatment. *J Allergy Clin Immunol Pract.* 2015 Apr; 3(2): 175–185; quiz 186.)

HSRs to mAbs can also be caused by IgE against antibody moieties, as in some patients sensitized to rituximab. Both immediate and delayed hypersensitivity reactions (skin lesions with CD4 T cells and eosinophilic infiltrates) can occur with tocilizumab.

mAbs given subcutaneously can elicit injection-site reactions, with symptoms including local redness, warmth, burning, stinging, itching, urticaria, pain, and induration. Such reactions can start within 1 hour of the injection and usually resolve over a few days, but large and persistent reactions can lead to discontinuation of the mAb; desensitization protocols have been developed for patients who have no alternative medications (see Table 48.5).

Chemotherapy

Most cancer chemotherapy drugs can induce hypersensitivity reactions. Platins are the agents most able to induce IgE responses. Platin-induced reactions typically occur after several exposures; the incidence of carboplatin allergy is about 27% after seven lifetime exposures and up to 46% in patients who have had 15 or more doses. Patients carrying the *BRCA1* and *BRCA2* mutations are at increased risk for carboplatin hypersensitivity reactions. Most HSRs to platins occur during the infusion, with symptoms consistent with type I reactions, including anaphylaxis (*i.e.*, flushing, hives, wheezing, hypotension). They are generally amenable to desensitization. Positive immediate skin test results to platins have a positive predictive value (PPV) of up to 86% for patients exposed to ≥6 courses of carboplatin, but 50% patients with reactions to oxaliplatin may have negative skin test results, indicating that non-IgE mechanisms may be involved, such as direct activation of mast cells/basophils or a different cell target because some reactions are associated with cytokine storm–like reaction with fever and chills, suggesting IL-6 release. Repeated skin testing before each chemotherapy exposure has been suggested in patients with mild reactions to uncover sensitization before the patient reacts.[52] Adriamycin, cyclophosphamide, methotrexate, topotecan, and other chemotherapies can induce type I reactions, which are also amenable to rapid desensitization.

Taxanes Hypersensitivity

Taxanes include paclitaxel, docetaxel, cabazitaxel, and abraxane. Cremophor is used to solubilize paclitaxel molecules, and polysorbate 80 is used to solubilize docetaxel (Fig. 48.10). These solvents can cause complement activation, leading to generation of anaphylatoxins and mast cell activation. Initially, the side effects of paclitaxel were attributed to cremophor; these responded to premedication with antihistamines and high-dose corticosteroids. Despite using premedication, HSRs occur in up to 10% of patients treated with paclitaxel, and in 1% the reactions can be severe, necessitating discontinuation of the medication. Reactions

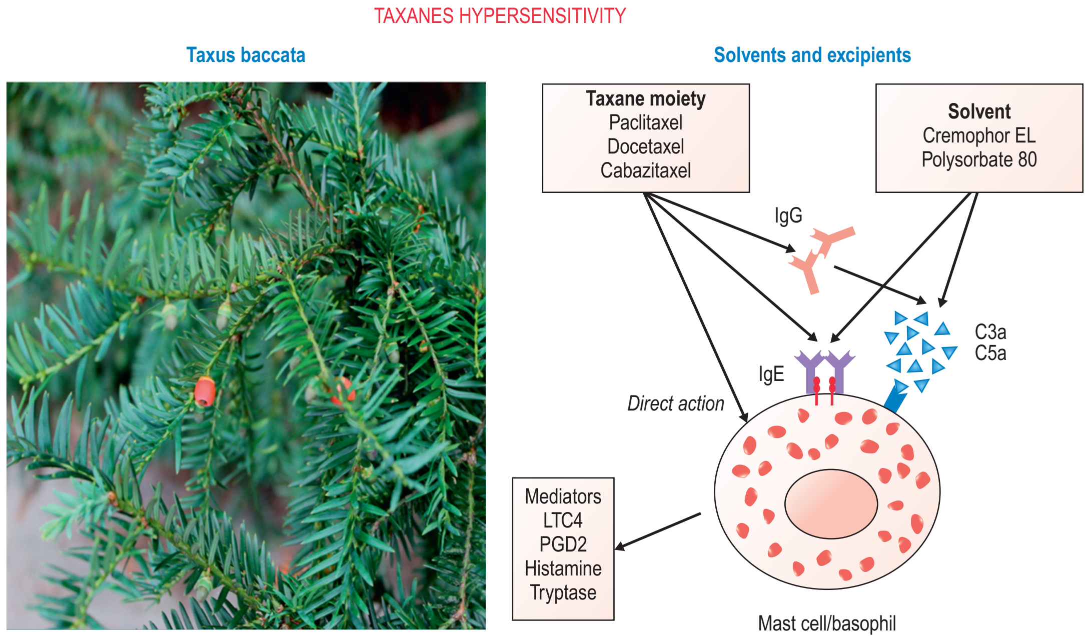


FIG 48.10 Taxanes: Plant origin, solvents, and excipients.

typically occur during the first or second lifetime exposure with such symptoms as throat tightness, flushing, hypotension, and dyspnea. However, some patients also report severe chest, back, and/or pelvic pain.

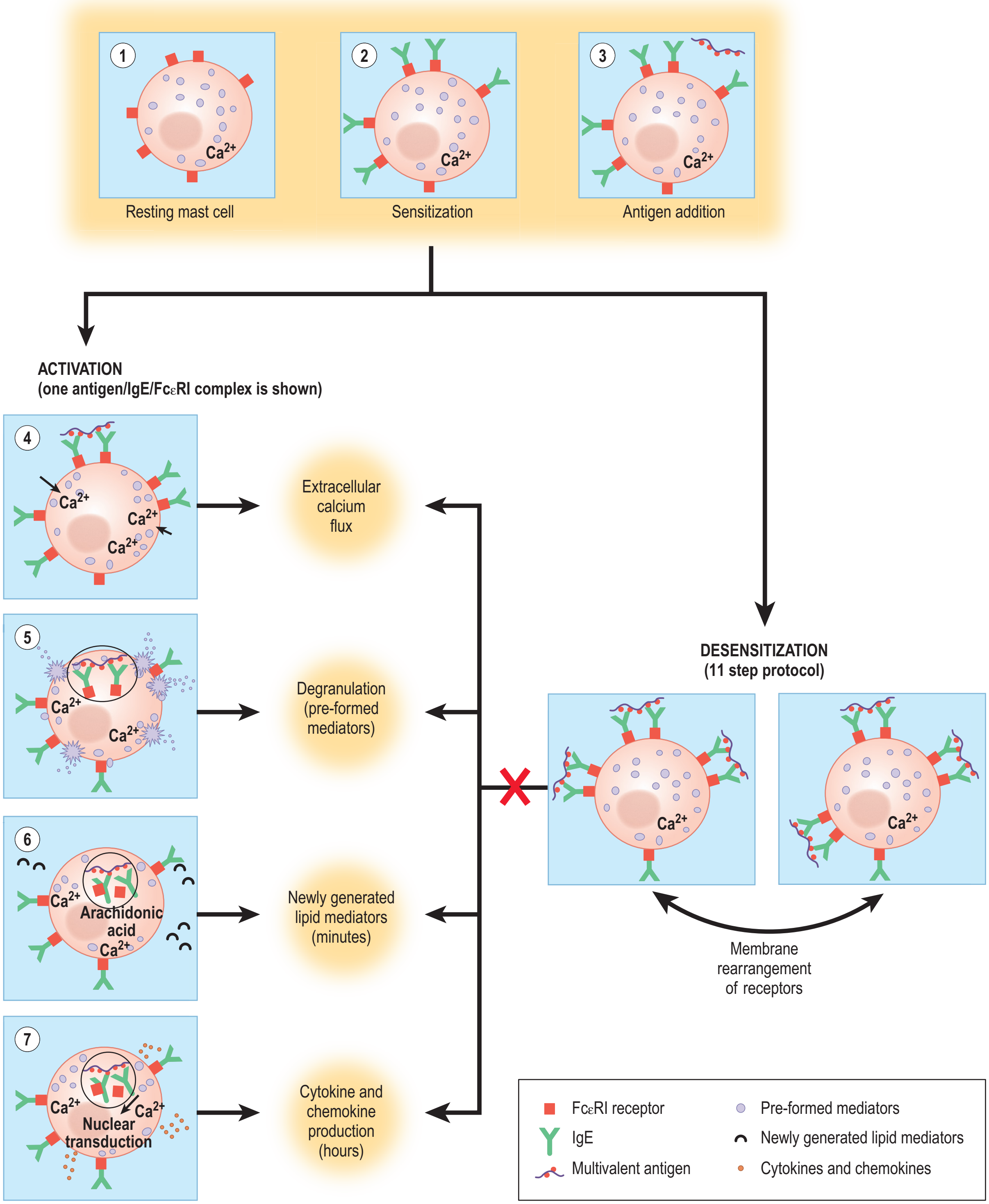
IgE-mediated HSRs to taxane molecules have been reported, leading to skin test evaluations for patients with taxane reactions. When skin testing was performed in 145 patients with taxane-related HSRs, 103 (71%) had positive results. Thirty-six (22%) patients with negative skin test results eventually resumed regular infusions.[53]

MANAGEMENT OF DRUG HYPERSENSITIVITY AND DESENSITIZATION

Stopping a drug suspected of inducing a reaction has therapeutic and diagnostic implications. If the symptoms are alleviated upon cessation, it is likely, but not certain, that the drug was responsible for the reaction. Although most type I symptoms are readily reversible, death has occurred in patients on concomitant beta-blockers or ACE inhibitors when epinephrine did not reverse cardiovascular collapse. Patients with severe anaphylactic drug reactions, as seen with chemotherapy and platins, may develop posttraumatic stress disorders. Severe reactions, such as DRESS, SJS, and TEN, do not have specific therapies and require intensive and aggressive treatments, including steroids and/or immunoglobulin infusion at high doses, to accelerate the resolution of inflammatory events and facilitate recovery. Recurrence of these reactions upon inadvertent exposure can be lethal, and those surviving can suffer life-debilitating consequences, such as depression and increased suicides. Blindness, permanent hair and nail loss, mucosal dryness, and fibrosis can dramatically decrease the quality of life in patients with SJS and TEN, and patients with DRESS can have persistent hepatitis and/or eosinophilic cardiac complications.

For type I IgE- and non–IgE-mediated reactions, desensitization is now the state-of-the-art management option in patients who still require the offending drug as first-line therapy. Desensitization has been successfully achieved for chemotherapy drugs, mAbs, antibiotics, and other medications in both adults and children. Although drug reactions are less common in children than in adults, perhaps because of fewer exposures, patients with cystic fibrosis have an average 30% sensitization rate to antibiotics; they can be successfully desensitized with first-line antibiotics, even during lung transplantation. The mechanism of desensitization for IgE-mediated reactions has been studied *in vitro* and in rodents. Although antigen induces membrane changes leading to internalization of the antigen, IgE and FcεRI, when cells are repeatedly exposed to increasing suboptimal drug concentrations, this prevents these membrane events and blocks calcium entry and the release of inflammatory mediators (Fig. 48.11). Other mechanisms that may be relevant to human desensitization have also been implicated (Fig. 48.12). Further research is needed to better understand these.

Application of the principles of desensitization has been remarkably successful, and thousands of patients have been safely desensitized, including patients suffering severe anaphylactic reactions. Indications for desensitization include type I IgE- and non–IgE-mediated reactions. The risk/benefit ratio needs careful consideration because desensitization can sometimes induce severe reactions. If there is a good alternative drug, desensitization should not be attempted. Desensitization protocols generally involve diluting the drug up to 1/1000 of its normal concentration, followed by doubling doses at short intervals. A 12-step protocol for carboplatin desensitization is shown in Table 48.6. For patients with high risk, four bags and 16 steps may be used as shown in Table 48.7 for ofatumumab. For patients with local and systemic reactions to subcutaneous injections, desensitization has been successfully achieved by the subcutaneous route.



① Resting mast cell
② Sensitization
③ Antigen addition
Ca2+
ACTIVATION
(one antigen/IgE/FcεRI complex is shown)
④ Ca2+
Extracellular calcium flux
⑤ Ca2+
Degranulation (pre-formed mediators)
⑥ Ca2+ Arachidonic acid
Newly generated lipid mediators (minutes)
⑦ Ca2+ Nuclear transduction
Cytokine and chemokine production (hours)
DESENSITIZATION
(11 step protocol)
Membrane rearrangement of receptors
FcεRI receptor
IgE
Multivalent antigen
Pre-formed mediators
Newly generated lipid mediators
Cytokines and chemokines

FIG 48.11 Simplified cartoon comparing activation and desensitization outcomes as well as a possible explanation of how rapid desensitization works and the rearrangement of the FcεRI receptors at the cell membrane. (From Castells et al. Hypersensitivity to antineoplastic agents: mechanisms and treatment with rapid desensitization. *Cancer Immunol Immunother.* 2012; 61: 1575–84.)

TABLE 48.6 Example of 12-Step Desensitization Protocol for Carboplatin

Target dose (mg)	300
Standard volume per bag (mL)	250
Final rate of infusion (mL/h)	80
Calculated target concentration (mg/mL)	1.2
Standard time of infusion (minutes)	187.5

	Volume		Concentration (mg/mL)		Total mg per Bag	Amount Infused (mL)
Solution 1	250	mL of	0.012	mg/mL	3	9.38
Solution 2	250	mL of	0.120	mg/mL	30	18.75
Solution 3	250	mL of	1.190	mg/mL	297.638	250
NB		The total volume and dose dispensed are more than the final dose given to patient because many of the solutions are not completely infused.				

Step	Solution	Rate (mL/h)	Time (min)	Volume Infused Per Step (mL)	Dose Administered With This Step (mg)	Cumulative Dose (mg)
1	1	2.5	15	0.63	0.0075	0.0075
2	1	5	15	1.25	0.015	0.0225
3	1	10	15	2.5	0.03	0.0525
4	1	20	15	5	0.06	0.1125
5	2	5	15	1.25	0.15	0.2625
6	2	10	15	2.5	0.3	0.5625
7	2	20	15	5	0.6	1.1625
8	2	40	15	10	1.2	2.3625
9	3	10	15	2.5	2.98	5.3389
10	3	20	15	5	5.95	11.2916
11	3	40	15	10	11.91	23.1971
12	3	80	174.375	232.5	276.8	300

Total time: 5.66 hours

TABLE 48.7 Example of a Four-Bag 16-Step High-Risk Desensitization to Ofatuzumab

Bag	Volume per Bag (mL)	Concentration (mg/mL)	Total Dose per Bag (mg)	Amount of Bag Infused (mL)
1	250	0.002	0.500	9.38
2	250	0.040	10.000	9.38
3	250	0.400	100.000	18.75
4	250	3.968	992.106	250.00
NB	The total volume and dose dispensed are more than the final dose given to patient because many of the solutions are not completely infused.			

Step	Bag	Rate (mL/h)	Time (min)	Cumulative Time(min)	Volume Infused per Step(mL)	Dose Administered With This Step (mg)	Cumulative Dose (mg)	Fold Increase per Step
1	1	2.5	15	15	0.625	0.001	0.001	
2	1	5	15	30	1.25	0.003	0.004	2
3	1	10	15	45	2.5	0.005	0.009	2
4	1	20	15	60	5	0.010	0.019	2
5	2	2.5	15	75	0.625	0.025	0.044	2.5
6	2	5	15	90	1.25	0.050	0.094	2
7	2	10	15	105	2.5	0.100	0.194	2
8	2	20	15	120	5	0.200	0.394	2
9	3	5	15	135	1.25	0.500	0.894	2.5
10	3	10	15	150	2.5	1.000	1.894	2
11	3	20	15	165	5	2.000	3.894	2
12	3	40	15	180	10	4.000	7.894	2
13	4	10	15	195	2.5	9.921	17.815	2.5
14	4	20	15	210	5	19.842	37.657	2
15	4	40	15	225	10	39.684	77.341	2
16	4	80	174.375	399.375	232,5	922.659	1,000.000	

Total infusion time: 6.6 hours
Standard volume per bag: 250 mL
Final rate of infusion: 80 mL/h (as patient now deemed tolerant)
Number of bags: 4; time per step: 15 minutes; total number of steps: 16; total dose: 1000 mg.

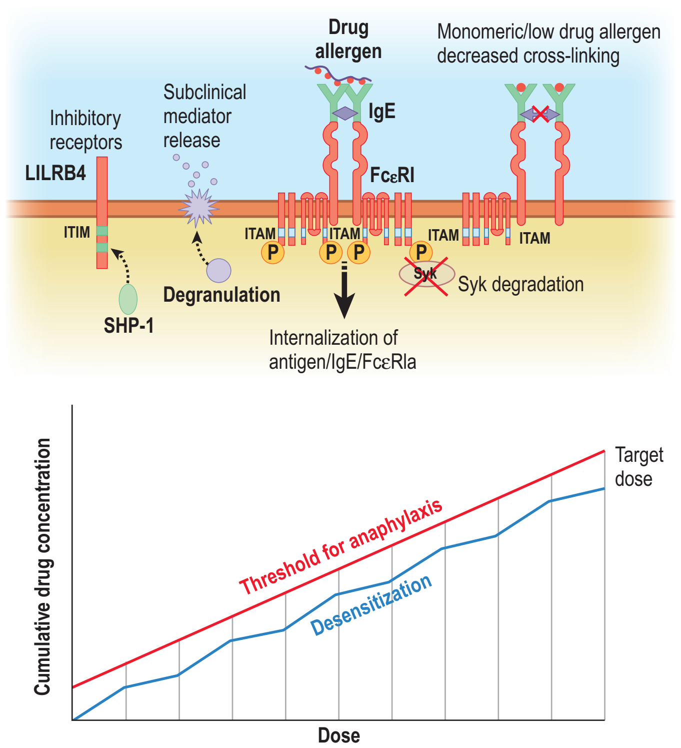

Drug allergen
Monomeric/low drug allergen decreased cross-linking
Subclinical mediator release
Inhibitory receptors
LILRB4
IgE
FcεRI
ITIM
ITAM
ITAM
ITAM
ITAM
P
P
P
P
Syk
Syk degradation
Degranulation
SHP-1
Internalization of antigen/IgE/FcεRIa
Cumulative drug concentration
Target dose
Threshold for anaphylaxis
Desensitization
Dose

FIG 48.12 Proposed mechanism for drug allergens/IgE desensitization. (Courtesy of Sancho-Serra MC and Bonamichi-Santos R.)

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

ON THE HORIZON

- There is a need for a new mechanistic approach and categorization of delayed hypersensitivity reactions, better diagnostic tools, and a standardized approach to desensitization protocols.
- Modulation and treatment of hypersensitivity reactions will include the use of anti–IgE and other mAbs as well as small molecules to inhibit signal transduction.
- Prevention of hypersensitivity reactions will include the identification of patients at risk by HLA genotyping.
- Drug desensitization should be more widely available. Patients with allergies can then be treated with first-line therapies, which will improve quality of life and reduce both costs and risks.

REFERENCES

1. Pirmohamed M, James S, Meakin S, et al. Adverse drug reactions as cause of admission to hospital: prospective analysis of 18,820 patients. BMJ 2004;329:15–19.
2. Hakkarainen KM, Hedna K, Petzold M, et al. Percentage of patients with preventable adverse drug reactions and preventability of adverse drug reactions—a meta-analysis. PLoS ONE 2012;7:e33236.
3. Kadoyama K, Kuwahara A, Yamamori M, et al. Hypersensitivity reactions to anticancer agents: data mining of the public version of the FDA adverse event reporting system, AERS. J Exp Clin Cancer Res 2011;30:93.
4. Castells M. Rapid desensitization for hypersensitivity reactions to medications. Immunol Allergy Clin North Am 2009;29:585–606.
5. Macy E, Contreras R. Health care use and serious infection prevalence associated with penicillin "allergy" in hospitalized patients: A cohort study. J Allergy Clin Immunol 2014;133:790–6.
6. Macy E, Mangat R, Burchette RJ. Penicillin skin testing in advance of need: multiyear follow-up in 568 test result-negative subjects exposed to oral penicillins. J Allergy Clin Immunol 2003;111:1111–15.
7. Mallal S, Phillips E, Carosi G, et al. HLA-B*5701 screening for hypersensitivity to abacavir. N Engl J Med 2008;358:568–79.
8. Schwartz LB, Metcalfe DD, Miller JS, et al. Tryptase levels as an indicator of mast-cell activation in systemic anaphylaxis and mastocytosis. NEJM 1987;316:1622–6.
9. Lazarou J, Pomeranz BH, Corey PN. Incidence of adverse drug reactions in hospitalized patients. JAMA 1998;279:1200–5.
10. Wheatley LM, Plaut M, Schwaninger JM, et al. Report from the National Institute of Allergy and Infectious Diseases workshop on drug allergy. J Allergy Clin Immunol 2015;136:262–271 e2.
11. Bigby M, Jick S, Jick H, et al. Drug-induced cutaneous reactions. A report from the Boston Collaborative Drug Surveillance Program on 15,438 consecutive inpatients, 1975 to 1982. JAMA 1986;256:3358–63.
12. Aun MV, Blanca M, Garro LS, et al. Nonsteroidal anti-inflammatory drugs are major causes of drug-induced anaphylaxis. J Allergy Clin Immunol Pract 2014;2:414–20.
13. Banerji A, Rudders S, Clark S, et al. Retrospective study of drug-induced anaphylaxis treated in the emergency department or hospital: patient characteristics, management, and 1-year follow-up. J Allergy Clin Immunol Pract 2014;2:46–51.
14. Oettgen HC. Regulation of the IgE isotype switch: new insights on cytokine signals and the functions of epsilon germline transcripts. Curr Opin Immunol 2000;12:618–23.
15. Gorevic PD, Levine BB. Desensitization of anaphylactic hypersensitivity specific for the penicilloate minor determinant of penicillin and carbenicillin. J Allergy Clin Immunol 1981;68:267–72.
16. White KD, Chung WH, Hung SI, et al. Evolving models of the immunopathogenesis of T cell-mediated drug allergy: the role of host, pathogens, and drug response. J Allergy Clin Immunol 2015;136:219–34.
17. Pirmohamed M, Ostrov DA, Park BK. New genetic findings lead the way to a better understanding of fundamental mechanisms of drug hypersensitivity. J Allergy Clin Immunol 2015;136:236–44.
18. Chung WH, Chang WC, Lee YS, et al. Genetic variants associated with phenytoin-related severe cutaneous adverse reactions. JAMA 2014;312:525–34.
19. Gueant JL, Romano A, Cornejo-Garcia JA, et al. HLA-DRA variants predict penicillin allergy in genome-wide fine-mapping genotyping. J Allergy Clin Immunol 2015;135:253–9.
20. Park MA, Matesic D, Markus PJ, et al. Female sex as a risk factor for penicillin allergy. Ann Allergy Asthma Immunol 2007;99:54–8.
21. Brown RH, Taenkhum K, Buckley TJ, et al. Different latex aeroallergen size distributions between powdered surgical and examination gloves: significance for environmental avoidance. J Allergy Clin Immunol 2004;114:358–63.
22. Parmar JS, Nasser S. Antibiotic allergy in cystic fibrosis. Thorax 2005;60:517–20.
23. Moon DH, Lee JM, Noonan AM, et al. Deleterious BRCA1/2 mutation is an independent risk factor for carboplatin hypersensitivity reactions. Br J Cancer 2013;109:1072–8.
24. Patil SU, Long AA, Ling M, et al. A protocol for risk stratification of patients with carboplatin-induced hypersensitivity reactions. J Allergy Clin Immunol 2012;129:443–7.
25. Foer D, Buchheit KM, Gargiulo AR, et al. Progestogen hypersensitivity in 24 cases: diagnosis, management, and proposed renaming and classification. J Allergy Clin Immunol Pract 2016;4:723–9.
26. McNeil BD, Pundir P, Meeker S, et al. Identification of a mast-cell-specific receptor crucial for pseudo-allergic drug reactions. Nature 2015; 519:237–41.
27. Pichler WJ, Yawalkar N, Britschgi M, et al. Cellular and molecular pathophysiology of cutaneous drug reactions. Am J Clin Dermatol 2002;3:229–38.

28. Sancho-Serra M, Simarro M, Castells M. Rapid IgE desensitization is antigen specific and impairs early and late mast cell responses targeting FcεRI internalization. Eur J Immunol 2011;41:1004–13.
29. Reber LL, Hernandez JD, Galli SJ. The pathophysiology of anaphylaxis. J Allergy Clin Immunol 2017;140:335–48.
30. Zweizig S, Roman LD, Muderspach LI. Death from anaphylaxis to cisplatin: a case report. Gynecol Oncol 1994;53:121–2.
31. Wong JT, Ling M, Patil S, et al. Oxaliplatin hypersensitivity: evaluation, implications of skin testing, and desensitization. J Allergy Clin Immunol Pract 2014;2:40–5.
32. Simons FE, Ebisawa M, Sanchez-Borges M, et al. 2015 update of the evidence base: World Allergy Organization anaphylaxis guidelines. World Allergy Organ J 2015;8:32.
33. Castells MC, Tennant NM, Sloane DE, et al. Hypersensitivity reactions to chemotherapy: outcomes and safety of rapid desensitization in 413 cases. J Allergy Clin Immunol 2008;122:574–80.
34. Caiado J, Venemalm L, Pereira-Santos MC, et al. Carboplatin-, oxaliplatin-, and cisplatin-specific IgE: cross-reactivity and value in the diagnosis of carboplatin and oxaliplatin allergy. J Allergy Clin Immunol Pract 2013;1:494–500.
35. Picard M, Pur L, Caiado J, et al. Risk stratification and skin testing to guide re-exposure in taxane-induced hypersensitivity reactions. J Allergy Clin Immunol 2016;137:1154–64.
36. Dybendal T, Guttormsen AB, Elsayed S, et al. Screening for mast cell tryptase and serum IgE antibodies in 18 patients with anaphylactic shock during general anaesthesia. Acta Anaesthesiol Scand 2003;47:1211–18.
37. Aster RH. Drug-induced immune thrombocytopenia: an overview of pathogenesis. Semin Hematol 1999;36:2–6.
38. Arndt PA, Garratty G. The changing spectrum of drug-induced immune hemolytic anemia. Semin Hematol 2005;42:137–44.
39. Greinacher A, Warkentin TE. Recognition, treatment, and prevention of heparin-induced thrombocytopenia: review and update. Thromb Res 2006;118:165–76.
40. Brennan PJ, Rodriguez Bouza T, Hsu FI, et al. Hypersensitivity reactions to mAbs: 105 desensitizations in 23 patients, from evaluation to treatment. J Allergy Clin Immunol 2009;124:1259–66.
41. Lieberman P, Nicklas RA, Oppenheimer J, et al. The diagnosis and management of anaphylaxis practice parameter: 2010 update. J Allergy Clin Immunol 2010;126:477–80, e1–42.
42. Barbaud A. Skin testing and patch testing in non-IgE-mediated drug allergy. Curr Allergy Asthma Rep 2014;14:442–8.
43. Pichler WJ, Tilch J. The lymphocyte transformation test in the diagnosis of drug hypersensitivity. Allergy 2004;59:809–20.
44. Miranda A, Blanca M, Vega JM, et al. Cross-reactivity between a penicillin and a cephalosporin with the same side chain. J Allergy Clin Immunol 1996;98:671–7.
45. Brockow K, Romano A, Aberer W, et al. Skin testing in patients with hypersensitivity reactions to iodinated contrast media—a European multicenter study. Allergy 2009;64:234–41.
46. Hepner DL, Castells MC. Anaphylaxis during the perioperative period. Anesth Analg 2003;97:1381–95.
47. Coopman SA, Johnson RA, Platt YR, et al. Cutaneous disease and drug reactions in HIV infection. N Engl J Med 1993;328:1670–4.
48. Naisbitt DJ, Farrell J, Gordon SF, et al. Covalent binding of the nitroso metabolite of sulfamethoxazole leads to toxicity and major histocompatibility complex-restricted antigen presentation. Mol Pharmacol 2002;62:628–37.
49. Strom BL, Schinnar R, Apter AJ, et al. Absence of cross-reactivity between sulfonamide antibiotics and sulfonamide nonantibiotics. NEJM 2003;349:1628–35.
50. Laidlaw TM, Kidder MS, Bhattacharyya N, et al. Cysteinyl leukotriene overproduction in aspirin-exacerbated respiratory disease is driven by platelet-adherent leukocytes. Blood 2012;119:3790–8.
51. Galvão VR, Castells MC. Hypersensitivity to biological agents—updated diagnosis, management, and treatment. J Allergy Clin Immunol Pract 2015;3:175–86.
52. Markman M, Zanotti K, Peterson G, et al. Expanded experience with an intradermal skin test to predict for the presence or absence of carboplatin hypersensitivity. J Clin Oncol 2003;21:4611–14.
53. Picard M, Pur L, Caiado J, et al. Risk stratification and skin testing to guide re-exposure in taxane-induced hypersensitivity reactions. J Allergy Clin Immunol 2016;137:1154–1164 e12.

MULTIPLE-CHOICE QUESTIONS

1. Human herpesvirus 6/7 reactivation during an adverse drug reaction is frequently seen in:
   A. Stevens Johnson syndrome
   B. Toxic epidermal necrolysis
   C. Maculopapular exanthema
   D. Drug reaction with eosinophilia and systemic symptoms

2. Screening for which human leukocyte antigen (HLA) alleles should be done in a patients with carbamazepine-induced Stevens-Johnson syndrome/toxic epidermal necrosis (SJS/TEN)?
   A. HLA-B* 57:01
   B. HLA-B* 15:02
   C. HLA-B* 13:01
   D. HLA-B* 58:01

3. The *most* important risk factor associated with allergic reactions to carboplatin in women with ovarian cancer include:
   A. High cancer burden with metastatic disease
   B. *BRCA1* and *BRCA2* mutations
   C. Food allergy
   D. Asthma

4. Patients with sulfonamide antibiotic allergy:
   A. Typically present with IgE-mediated reactions
   B. Can use nonantibiotic sulfonamides
   C. Are not at risk for reactions to other antibiotics
   D. Cannot be desensitized

5. Drug desensitizations:
   A. Are limited to type I, IgE-mediated reactions
   B. Are not performed in patients with anaphylaxis
   C. Can induce drug tolerance
   D. Require specific protocols based on the dose, initial reaction, and patient risk

6. Serum markers of drug hypersensitivity and mast cell activation include:
   A. Histamine
   B. Complement CH50
   C. Tryptase
   D. Bradykinins

49

Occupational Respiratory Allergies

Catherine Lemière, James G. Martin

Occupational asthma (OA) is a unique phenotype of asthma, which provides us with an opportunity to understand the origins not only of OA but also of asthma in general. In the vast majority of cases, the onset of OA occurs in adulthood, in response to a specific exposure at work. As such, it opens a unique window on the natural history of asthma and the risk factors associated with the risk of development of the disease. This chapter will review the mechanisms of occupational sensitization, the evaluation of the disease from exposure to the manifestation of the disease clinically, and the main steps in diagnosis and management of sensitizer-induced OA.

KEY CONCEPTS

- Occupational sensitizers induce sensitization through different mechanisms, depending on their nature (proteins vs chemicals).
- Atopy, rhinitis, and genetic markers are host risk factors predisposing to develop occupational asthma (OA).
- The level of exposure to the causal agent is the main environmental risk factor predisposing to development of OA.
- Experimental models support roles for a diverse range of lymphocytes in airway hyperresponsiveness (AHR) and inflammation induced by allergens and nonallergenic stimuli.
- Objective confirmation of a diagnosis of OA is crucial because of socioeconomic implications for affected workers.

MECHANISMS OF OCCUPATIONAL SENSITIZATION

More than 400 distinct agents have been documented as causing OA.[1] The type of agents to which the subjects are exposed drives different pathophysiological mechanisms. The agents responsible for OA are usually considered according to their molecular weights. High-molecular-weight (HMW) agents (>10 kilodaltons [kDa]) include proteins of animal and vegetable origin and microorganisms, whereas low-molecular-weight (LMW) agents include wood dust, drugs, metals, and chemicals. Exposure to HMW agents usually induces the production of allergen-specific immunoglobulin E (IgE), whereas only a few LMW agents have been associated with the production of specific IgE.[2]

IgE-Mediated

HMW agents are proteins and in general bear similarities to the aeroallergens triggering usual allergic asthma. They act as complete antigens provoking the production of specific IgE antibodies. Sensitization involves the uptake of antigen by immature dendritic cells (DCs) residing in the subepithelial space within the airway walls.[3] DCs extend cellular projections between the epithelial cells and are believed to sample the airway microenvironment. Once antigen has been taken up, the DC migrates to regional lymph nodes, undergoes maturation, and expresses antigenic peptide on its surface in conjunction with major histocompatibility complex (MHC) type II molecules. In this location, they encounter circulating CD4 T lymphocytes that stochastically sample the DC surface.[4] The DC eventually interacts with CD4 T cells expressing high affinity receptors for the antigen–MHC II complex. The T cells expand and exit the lymph node, and T-helper-2 (Th2) cells will eventually resettle in the airways, attracted by chemokine gradients, in particular, the macrophage-derived chemokine (MDC) CCL22 and the thymus and activation-regulated chemokine (TARC) CCL17, which are secreted by epithelial cells and act on their cognate receptor CCR4 on the T cells. In the airways, CD4 T cells may react to antigen presentation locally on future exposures,[5] synthesizing interleukin (IL)-4, IL-5, and IL-13. Certain characteristics of antigens, such as protease activity or endogenous nicotinamide adenine dinucleotide phosphate (NADPH) oxidase activity, have been associated with allergenicity, likely through the effects of these stimuli on epithelial cells, which, in turn, secrete cytokines that not only favor the development of Th2 cells but also activate novel innate lymphoid cells that themselves produce type 2 cytokines. Th2 cells provide key help to B cells to undergo isotype switching and produce IgE. The binding of IgE to both high- and low-affinity receptors on a number of cells and its subsequent interaction with antigen not only result in mediator release from mast cells and basophils but also have effects on the functions of other cells, such as neutrophils and smooth muscle cells.[6]

A few LMW occupational agents also induce specific IgE antibodies, but they do so by acting as haptens and binding with proteins to form functional antigens. Specific IgE to isocyanates and plicatic acid (western red cedar, *Thuja plicata*) have been identified. However, the significance of finding specific IgE to isocyanates or plicatic acid is not so straightforward to interpret. In contrast, OA caused by platinum salts or acid anhydrides is clearly IgE mediated, and so the presence of specific IgE to those agents is a useful additional tool for making a diagnosis of OA to those agents.

Non–IgE-Mediated

Cell-mediated reactions are likely to play an important role in OA because of LMW agents. Although the predominant immune response to chemical respiratory allergens may be of the Th2 type, other cells may play important support or regulatory roles. CD4 as well as CD8 T cells and different cytokines, such as IL-1, IL-4, IL-5, IL-6, and IL-15, have been found in the biopsy material, bronchoalveolar lavage (BAL), and sputum of patients with

TABLE 49.1 Summary of the Most Frequently Encountered Occupational Agents Causing Occupational Asthma (OA)

Agents	Occupation/Industry
High-Molecular-Weight Agents	
Immunoglobulin E (IgE)–Dependent Causes	
Cereals, flour	Flour mills, bakers, pastry makers
Animal-derived allergens: laboratory animals	Laboratory workers, farmers
Enzymes: α-amylase, maxatase, alcalase, papain, bromelain, pancreatin	Baking product production, bakers, detergent production, pharmaceutical industry, food industry
Seafood	Seafood processing
Latex	Health care workers, laboratory technicians
Low-Molecular-Weight Agents	
IgE-Dependent Causes	
Acid anhydrides, phthalic, trimellitic, tetrachlorophthalic anhydrides	Epoxy resin workers
Platinum salts	Platinum refining
Other Potential Immunological Mechanisms	
Diisocyanates: toluene diisocyanate (TDI), methylene diphenyl-diisocyanate (MDI), hexamethylene diisocyanate (HDI)	Polyurethane production, plastic industry, insulation, molding, spray painting
Wood dusts: western red cedar, iroko, obeche, oak, and others	Sawmill workers, carpenters, cabinet and furniture makers
Amines and epoxy resins	Rubber industry, cosmetics, hair dye manufacturing, fur industry, rubber additives
Colophony	Electronic industry (flux)
Acrylates: cyanoacrylates, methacrylates, di- and triacrylates	Adhesives, dental and orthopedic materials, sculptured fingernails, printing inks, paints and coatings
Pharmaceutical products	Pharmaceutical industry
Formaldehyde, glutaraldehyde, biocides	Health care workers, cleaners
Persulfate salts: hair bleach	Hairdressers
Reactive dyes: reactive black 5, pyrazolone derivatives, vinyl sulfones, carmine	Textile workers, food industry workers
Metals: chromium, nickel, cobalt	Metal refinery, metal alloy production, electroplating, welding

isocyanate-induced asthma. Murine models of toluene diisocyanate "asthma" have shown a dependence of airway hyperresponsiveness (AHR) and Th2-type inflammation on both CD4 and CD8 cells, but primarily the former. IL-4 and IL-13 are major determinants of the pathobiology. Recent studies have suggested a role for innate lymphoid (lineage negative) cells to contribute to type 2 cytokine production in the airways in response to nonallergenic stimuli. Although few in number, these type 2 innate lymphoid cells can produce significant quantities of IL-5 and IL-13. IL-13 is a pleiotropic cytokine affecting the epithelium, smooth muscle, and immune cells and possibly drives asthma that is induced by allergens or irritants. The potent oxidant ozone induces AHR in the mouse through activation of inducible natural killer (NK) cells, so similar mechanisms may apply in human subjects exposed to molecules that cause oxidative stress.

Neutrophils are also likely to be involved in isocyanate-induced asthma, as shown by an increase in myeloperoxidase and IL-8 after exposure to toluene diisocyanate (TDI). A mixed Th1/Th2 pattern of cytokine production has been observed in subjects with red cedar–induced asthma. Furthermore, a specific inhalational challenge test induced a mixed Th1/Th2 response in which CD8 T cells were the main producers of interferon (IFN)-γ. However, histopathological changes in the airways in isocyanate-induced asthma are similar to those found in atopic asthma that is not associated with occupational sensitization.

There is evidence that isocyanates can stimulate human innate immune responses by upregulating immune pattern recognition receptors on monocytes and increasing the chemokines that regulate monocyte/macrophage trafficking (macrophage migration inhibitory factor [MIF], monocyte chemoattractant protein-1 [MCP-1]). Furthermore, repetitive antigenic stimulation of peripheral blood mononuclear cells from patients with diisocyanate asthma induced the synthesis of tumor necrosis factor (TNF)-α and MCP-1, but not IL-4 or IL-5. The bioactive lipid lysophosphatidic acid is released from lung epithelial cells by allergen or isocyanate and has the potential to contribute to the pathogenesis of asthma through a number of biological effects.

Table 49.1 summarizes the most frequent agents responsible for occupational asthma.

NATURAL HISTORY OF OCCUPATIONAL ASTHMA AND RISK FACTORS

Natural History of the Development of the Disease

Studies performed in apprentices shed light on the natural history of occupational asthma and risk factors for developing occupational asthma. Epidemiological studies give some indication of the course of the disease. Atopy and exposure levels have been shown to predict sensitization to laboratory animals; sensitization has been shown to precede the occurrence of rhinitis and asthma at the workplace.[7]

Risk Factors

The occurrence of OA is the result of an interaction between environmental and host factors (Fig. 49.1). Several host factors have been shown to increase the risk of developing OA (Table 49.2).

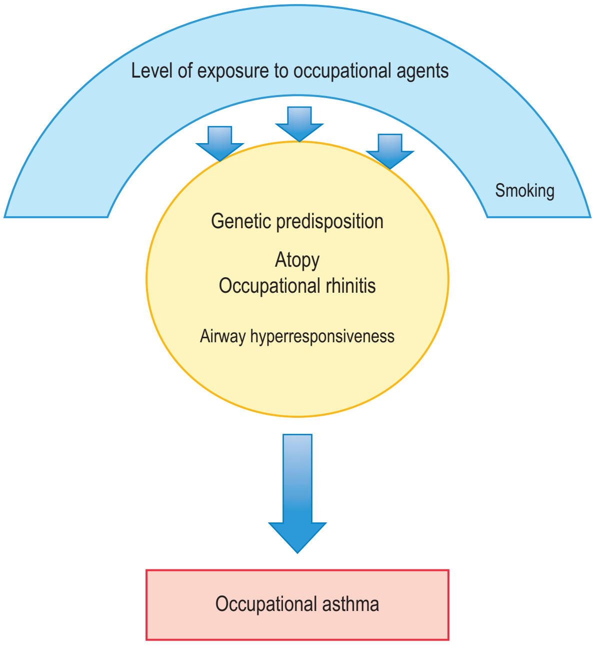


FIG 49.1 Environmental and host factor favoring the occurrence of occupational asthma.

TABLE 49.2 Risk Factors Associated With the Occurrence of Occupational Asthma (OA)

Risk Factor	Mechanisms
Environmental Factors	
High level of exposure Cigarette smoking	Activation of airway epithelium and sensory nerves through Toll-like receptors (TLRs) and transient receptor potential (TRP) channels.
Host-Related Factors	
Atopy	Promotion of a T-helper-2 (Th2) phenotype by multiple environmental stimuli when interacting with dendritic cells (DCs), including interleukin-4 (IL-4), thymic stromal lymphopoietin (TSLP), eosinophil-derived neurotoxin (EDN), and cysteinyl leukotrienes. Facilitation of antigen presentation by immunoglobulin E (IgE) binding to B cells and DCs
Genetic markers	The genetic polymorphisms identified do not yet lead to clear explanations of the known pathophysiological processes.
Preexisting nonspecific bronchial hyperresponsiveness and work-related rhinitis	Airway remodeling may be necessary before asthma becomes manifest in subjects with rhinitis.

Host Factors

Atopy

Clinical evidence. Atopy is clearly associated with increased risk of sensitization to HMW agents. For example, atopy has been identified as one of the determinants for sensitization to rodents in apprentices working in animal facilities.[7] Furthermore, the vast majority of apprentices who become sensitized to HMW agents have atopy, in contrast to those who do not get sensitized.[7] Therefore, and not surprisingly, atopy is also a risk factor for developing OA to HMW agents, since subjects who develop OA need to be sensitized to a specific agent to develop an asthmatic reaction on exposure. Atopy has been associated with OA to several different HMW agents, including OA caused by flour, laboratory animals, latex, snow crab, detergent enzymes, and α-amylase. It seems that the sensitization to specific allergen occurs mainly in the first 1–2 years after the beginning of exposure.[8] However, atopy does not seem to influence the timing of onset of OA symptoms.[9]

Pathophysiology. The predisposition to atopy has increased markedly in recent years. The explanation for the increase is considered to be largely environmental but has not been fully identified, although airborne particulates have been shown to have adjuvant effects and may, therefore, play a part. The hygiene hypothesis has led to evidence of protective effects of certain environments, such as the farm. A fundamental question remains as to why T-cell differentiation into Th2 cells occurs in the airways. Multiple stimuli within the airway environment have been shown to promote a Th2 phenotype when interacting with DCs, including IL-4, thymic stromal lymphopoietin (TSLP), eosinophil-derived neurotoxin, and cysteinyl leukotrienes. Interaction of antigens directly with airway epithelial cells releases alarmins, such as IL-33, and TSLP, which condition DCs such that their interactions with CD4 T cells direct them to produce Th2 type cytokines. These epithelial-derived cytokines also direct innate lymphoid cells to produce type II cytokines.

Certain characteristics of antigens, such as protease activity or endogenous NADPH oxidase activity, have been associated with allergenicity.[10,11] Binding of IgE to B cells and DCs facilitates the process of antigen presentation and provides a potential explanation for the predisposition, once sensitization to one allergen has occurred, to develop allergy to unrelated antigens.[6]

Rhinitis and airway hyperresponsiveness

Clinical evidence. The presence of AHR and rhinitis before entering a workplace is an independent risk factor for subsequent IgE sensitization to allergens present at the workplace. The development of occupational rhinitis during exposure often precedes the occurrence of OA. However, the predictive value of work-related nasal symptoms is only 11.4% for the subsequent development of OA in workers exposed to laboratory animals, over a follow-up period of 30–42 months.[2]

Pathophysiology. The onset of allergic rhinitis before asthma suggests that additional pathophysiological factors are required for the expression of lower airway disease. The inflammation affecting the upper airway is generally felt to be representative of that present concurrently in the lower airway. However, the clinical expression of lower airway disease requires that the airways become reactive to the mediators of bronchoconstriction and mucus secretion, such as cysteinyl leukotrienes. The mechanism(s) of the increase in airway responsiveness remains uncertain but may involve growth of airway smooth muscle and/or changes in airway smooth muscle contractility through increase in expression of smooth muscle proteins. Other possible contributory changes include altered matrix deposition, an increase in mucous glands, and goblet cell hyperplasia, all of which are forms of airway remodeling that may be required before asthma becomes clinically manifest.

Genetic predisposition

Clinical evidence. There has been a tremendous effort to identify genetic predisposition for subjects developing occupational asthma. Certain human leukocyte antigen (HLA) class II molecules (*i.e.*, HLA-DR, HLA-DQ, and HLA-DP alleles) were found to be factors that either placed subjects at risk of OA or were protective against OA caused by various LMW and HMW agents. Genes associated with Th2-cell differentiation may also play a role in the development of OA. Increased risks of developing isocyanate-induced asthma may be associated with genetic variations in transforming growth factor (TNF)-α, TGF-β1, PTGS1(cyclooxygenase-1), and PTGS2 (cyclooxygenase-2) genes, as well as genes involved in the protection against oxidative stress, such as glutathione-S-transferase and N-acetyltransferase. Genome-wide association studies have identified genetic polymorphisms in catenin α-3, α-T catenin (CTNNA3) associated with isocyanate-induced asthma. Although some genetic markers seem to indicate susceptibility or a protective effect for developing OA, we are not yet ready to use these in clinical practice.[12]

Pathophysiology. As is frequently the case, the genetic polymorphisms identified do not lead to clear explanations of the known pathophysiological processes but often appear to be peripheral to the issue of excessive airway narrowing. Although genes associated with inflammatory processes and oxidative stress are expressed in association with OA driven by HMW or LMW agents, the precise mechanistic links await elucidation. Catenins are linked to epithelial cell adhesion processes and differentiation and, as such, may affect epithelial responses to injury or inflammation. Allergens that trigger neutrophilic inflammation will necessarily cause oxidative stress, as neutrophils are a rich source of reactive oxygen species and produce hypochlorite in high concentrations within the airway lumen. Prostaglandin synthesis may also result in proinflammatory effects, vasodilation, and airway smooth muscle constriction.

Environmental Factors

Level of exposure. The level of exposure to sensitizers has been well established as a strong predictor of the occurrence of OA and appears as the most important environmental risk factor for developing OA. A dose–response relationship has been identified between the level of exposure to HMW agents and the occurrence of IgE-mediated sensitization and OA. Such a dose–response relationship has also been documented for some LMW agents, such as platinum salts, acid anhydrides, and isocyanates.[2]

Tobacco smoke and irritants. Cigarette smoking may increase the risk of IgE-mediated sensitization to some HMW and LMW agents, although the association between smoking and the occurrence of OA is weak. The role of other environmental cofactors, such as nonrespiratory routes of exposure and concomitant exposure to endotoxin and pollutants at work, remains largely uncertain. Exposure to pollutants augments the response to allergens.[13] It is also possible that exposure to irritants in the workplace modifies the immunological response to sensitizers.

Pathophysiology. Exposures in the workplace may be complex and do not only involve sensitizing substances. It is increasingly recognized that nonimmunological stimuli, such as endotoxin, cold, and irritants, activate the airway epithelium and sensory nerves through Toll-like receptors (TLRs) and transient receptor potential (TRP) channels. The release of proinflammatory mediators may be predicted to evoke neutrophilic inflammation directly but also favors eosinophilic inflammation by the release of mediators favoring Th2-cell differentiation or type II innate lymphoid cell stimulation. Experimental models have shown a dependence of allergic sensitization on concomitant TLR4 activation by endotoxin, although in high concentrations Th2-cell inflammation is inhibited. In addition, TRPA1 is activated by stimuli, such as cigarette smoke, chlorine, aldehydes, scents, and endogenous products of oxidative stress, and has been shown to be essential for allergic "asthma" in a murine model.[14] Confirmation of such mechanisms in humans is required.

DIAGNOSIS AND MANAGEMENT OF OCCUPATIONAL ASTHMA

Diagnosis of Occupational Asthma

Sensitizer-induced OA should be suspected in every worker exposed to potential sensitizers with new-onset asthma or with asthma that has become difficult to control. The diagnosis of OA should not be discarded only on the basis of the time between onset of asthma and the beginning of the occupational exposure, since approximately 20% of subjects with OA report having childhood asthma or asthma onset before entering the workforce. Although some of them had a remission of their asthma before entering the workforce, many of them still experience asthma symptoms when entering the workforce.[15,16]

The respiratory symptoms (*e.g.*, dyspnea, chest tightness, wheezing, cough, and sputum production) are similar to those encountered in non–work-related asthma, but in OA, their occurrence is usually modulated by the work exposure. The symptoms can start at the beginning of the work shift, toward its end, or even after working hours, with remission or improvement during weekends and holidays. Rhinitis often precedes the occurrence of the respiratory symptoms, especially with HMW agents. A thorough clinical and occupational history must be carefully recorded, but the diagnosis of OA cannot be made only on the basis of a compatible history, which has a low positive predictive value. A comprehensive investigation should be performed to accurately diagnose OA.[1] Fig. 49.2 summarizes the different steps leading to a diagnosis of OA.

The investigation of OA should start as soon as the diagnosis is suspected, as a long duration of exposure after the occurrence of respiratory symptoms is associated with a poor prognosis. Immunological assessment is particularly important for the diagnosis of HMW agents. Skin prick tests or serological assessment of specific IgE have a high sensitivity but a relatively poor specificity, since they indicate sensitization, which is necessary but not sufficient for the disease. Unfortunately, immunological tests are limited by the lack of standardized commercially available reagents for skin and *in vitro* tests for allergens implicated in OA.

In any case, the diagnosis of asthma needs to be confirmed in a worker who has a clinical and occupational history compatible with OA, by documenting reversible airflow limitation and/or AHR.[17] Although the lack of AHR does not exclude the diagnosis of OA in subjects who have been removed from exposure, in subjects that are still working a negative methacholine challenge has a 95% negative predictive value for excluding the diagnosis of OA.[18] The work-relatedness of asthma should be assessed through serial measurements of peak expiratory flow (PEF) and/or AHR at work and off work and/or specific inhalation challenges in the laboratory or at the workplace.

Specific inhalation challenge tests involve exposing the subject to the putative cause of OA in the laboratory and/or at the workplace.[19] Although these tests are considered reference tests for diagnosing OA, they are only available in a small number of

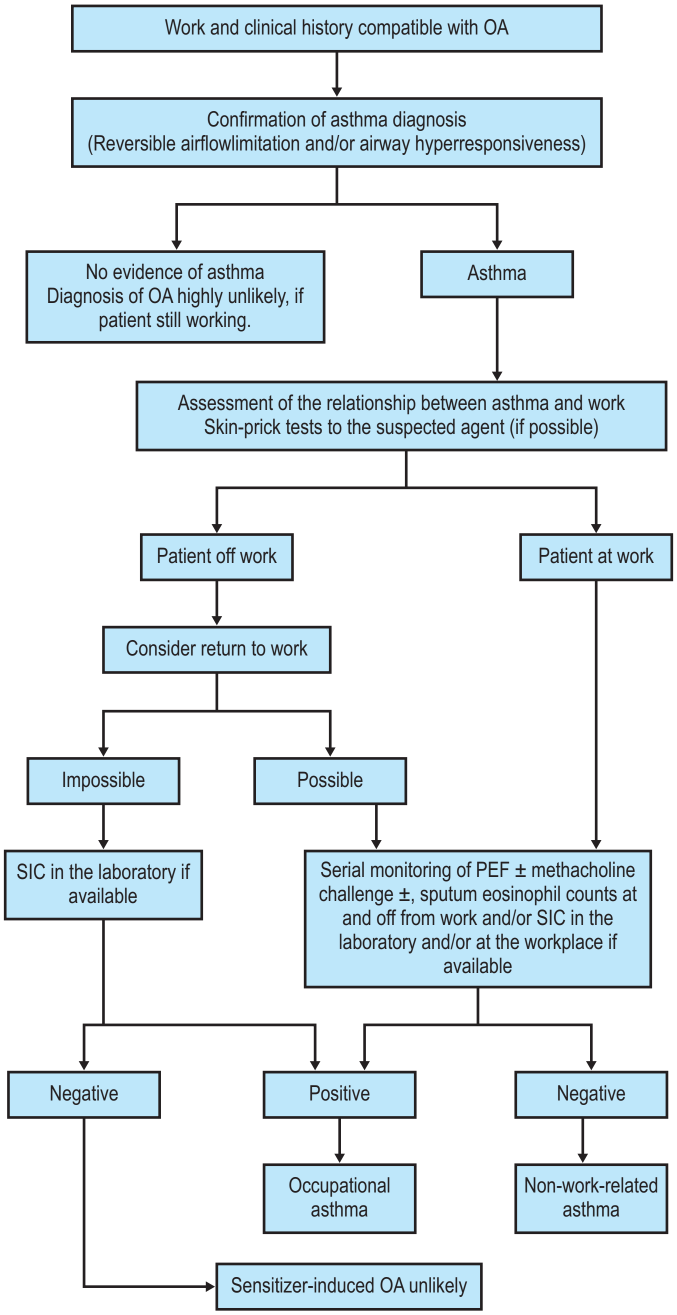

Work and clinical history compatible with OA
Confirmation of asthma diagnosis (Reversible airflowlimitation and/or airway hyperresponsiveness)
No evidence of asthma Diagnosis of OA highly unlikely, if patient still working.
Asthma
Assessment of the relationship between asthma and work Skin-prick tests to the suspected agent (if possible)
Patient off work
Patient at work
Consider return to work
Impossible
Possible
SIC in the laboratory if available
Serial monitoring of PEF ± methacholine challenge ±, sputum eosinophil counts at and off from work and/or SIC in the laboratory and/or at the workplace if available
Negative
Positive
Negative
Occupational asthma
Non-work-related asthma
Sensitizer-induced OA unlikely

FIG 49.2 Stepwise approach for diagnosing occupational asthma.

centers worldwide. Specific challenge tests are especially useful when *(i)* the diagnosis of OA remains uncertain after measurement of serial PEF or monitoring of airway responsiveness; *(ii)* a patient clearly has OA, but the causal agent needs to be identified; *(iii)* a new agent is suspected of causing OA; and *(iv)* the patient cannot be returned to the incriminated workplace. A false-negative response can occur if the wrong agent is used or if the exposure conditions are not comparable with those in the workplace.

Eosinophilic inflammation has been shown to increase after exposure to the causal agent(s) in the majority of subjects with OA. An increase in sputum cell count >3% after exposure to the suspected occupational agent has a specificity of 97% for a diagnosis of OA.[20] Although fractional concentration of expired nitric oxide (FeNO) levels are correlated with eosinophilic airway inflammation, the sensitivity of FeNO measured 24 hours after exposure appears to be too low (37%) to be useful in clinical practice.[21]

MANAGEMENT OF OCCUPATIONAL ASTHMA

Workers with sensitizer-induced OA who continue to be exposed to their causal agent are at a high risk of deterioration of asthma symptoms, airway obstruction, and nonspecific AHR. Complete and definitive avoidance of exposure to the causal agent is the optimal treatment for immunological OA. However, complete elimination of exposure or reassignment of affected workers to jobs that avoid exposure is not always possible. Reduction of exposure to the agent causing OA has been assessed as an alternative option to total avoidance. Unfortunately, this option is associated with a lower likelihood of asthma improvement and a higher risk of worsening. Therefore this management approach should be restricted to selected patients and requires careful medical monitoring to ensure early identification of asthma worsening.

Small uncontrolled studies suggest that immunotherapy may be effective in some cases of OA caused by HMW agents. However, some safety concerns remain. Further studies need to be conducted before immunotherapy can be recommended for OA to HMW agents.

Anti-IgE (omalizumab) may improve asthma control in subjects with flour-induced OA who remain exposed to the work environment, although further prospective investigations are required in subjects who choose to remain in an environment of exposure.

Even in spite of complete removal from exposure to the offending agent, asthma symptoms and AHR to methacholine persist in approximately 70% of the patients with OA several years after removal from the offending environment. Besides environmental interventions, pharmacological treatment of OA should follow standard clinical practice guidelines for asthma.

ON THE HORIZON

- Identification of new genetic markers may allow the identification of subgroups with high or low risks of developing occupational asthma (OA).
- Identification of new biomarkers may improve the diagnosis of OA.
- Identification of different OA phenotypes and endotypes that may improve the management of the disease.

Our understanding of how and why OA develops in a given individual has improved within the last 20 years, but many aspects remain obscure. Awareness that the level of exposure plays a pivotal role in the occurrence of sensitization has helped improve primary prevention of the disease. Although several markers of genetic predisposition have been identified, these findings have not yet been translated into biomarkers that can be used in clinical practice. Once OA is present, there is no reliable and effective medication to cure it. Instead, a thorough and timely investigation needs to be performed to diagnose it with certainty and then remove the patient from exposure to ensure the best possible prognosis.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Tarlo SM, Lemiere C. Occupational asthma. N Engl J Med 2014;370:640–9.
2. Lemiere C, Vandenplas O. Asthma in the workplace. In: Broaddus VC, editor. Murray & Nadel's textbook of respiratory medicine. 6th ed. St. Louis, MO: Elsevier; 2016. p. 1295–306.
3. Lambrecht BN, Hammad H. Biology of lung dendritic cells at the origin of asthma. Immunity 2009;31:412–24.
4. Miller MJ, Hejazi AS, Wei SH, et al. T cell repertoire scanning is promoted by dynamic dendritic cell behavior and random T cell motility in the lymph node. Proc Natl Acad Sci USA 2004;101:998–1003.
5. Huh JC, Strickland DH, Jahnsen FL, et al. Bidirectional interactions between antigen-bearing respiratory tract dendritic cells (DCs) and T cells precede the late phase reaction in experimental asthma: DC activation occurs in the airway mucosa but not in the lung parenchyma. J Exp Med 2003;198:19–30.
6. Galli SJ, Tsai M. IgE and mast cells in allergic disease. Nat Med 2012;18:693–704.
7. Moscato G, Pala G, Boillat MA, et al. EAACI position paper: prevention of work-related respiratory allergies among pre-apprentices or apprentices and young workers. Allergy 2011;66:1164–73.
8. Gautrin D, Ghezzo H, Infante-Rivard C, et al. Natural history of sensitization, symptoms and diseases in apprentices exposed to laboratory animals. Eur Respir J 2001;17:904–8.
9. Remen T, Acouetey DS, Paris C, et al. Early incidence of occupational asthma is not accelerated by atopy in the bakery/pastry and hairdressing sectors. Int J Tuberc Lung Dis 2013;17:973–81.
10. Romberger DJ, Heires AJ, Nordgren TM, et al. Proteases in agricultural dust induce lung inflammation through PAR-1 and PAR-2 activation. Am J Physiol Lung Cell Mol Physiol 2015;309:L388–99.
11. Boldogh I, Bacsi A, Choudhury BK, et al. ROS generated by pollen NADPH oxidase provide a signal that augments antigen-induced allergic airway inflammation. J Clin Invest 2005;115:2169–79.
12. Hur GY, Park HS. Biological and genetic markers in occupational asthma. Curr Allergy Asthma Rep 2015;15:488.
13. Carlsten C, Blomberg A, Pui M, et al. Diesel exhaust augments allergen-induced lower airway inflammation in allergic individuals: a controlled human exposure study. Thorax 2016;71:35–44.
14. Caceres AI, Brackmann M, Elia MD, et al. A sensory neuronal ion channel essential for airway inflammation and hyperreactivity in asthma. Proc Natl Acad Sci USA 2009;106:9099–104.
15. Larbanois A, Jamart J, Delwiche J, et al. Socioeconomic outcome of subjects experiencing asthma symptoms at work. Eur Respir J 2002;19:1107–13.
16. Lemiere C, Boulet LP, Chaboillez S, et al. Work-exacerbated asthma and occupational asthma: do they really differ? J Allergy Clin Immunol 2013;131:704–10.
17. Lougheed MD, Lemiere C, Ducharme FM, et al. Canadian Thoracic Society 2012 guideline update: diagnosis and management of asthma in preschoolers, children and adults. Can Respir J 2012;19:127–64.
18. Pralong JA, Lemiere C, Rochat T, et al. Predictive value of nonspecific bronchial responsiveness in occupational asthma. J Allergy Clin Immunol 2016;137:412–16.
19. Vandenplas O, Suojalehto H, Aasen TB, et al. Specific inhalation challenge in the diagnosis of occupational asthma: consensus statement. Eur Respir J 2014;43:1573–87.
20. Vandenplas O, D'Alpaos V, Heymans J, et al. Sputum eosinophilia: an early marker of bronchial response to occupational agents. Allergy 2009;64:754–61.
21. Lemiere C, D'Alpaos V, Chaboillez S, et al. Investigation of occupational asthma: sputum cell counts or exhaled nitric oxide? Chest 2010;137:617–22.

MULTIPLE-CHOICE QUESTIONS

1. Which characteristics have been identified as risk factors for developing occupational asthma (OA)?
   A. Atopy
   B. Obesity
   C. Smoking
   D. Genetic markers
   E. None of the above

2. In a subject with a history of childhood asthma who decided to quit his work because of recurrence of asthma, which factors can exclude the diagnosis of occupational asthma (OA)?
   A. Childhood asthma
   B. Absence of airway hyperresponsiveness (AHR)
   C. Normal spirometry
   D. Absence of eosinophilic inflammation
   E. None of the above

3. In a subject with diagnosed sensitizer-induced occupational asthma (OA) what is the most effective intervention?
   A. Complete removal from exposure
   B. Reduction of exposure
   C. Treatment with anti–immunoglobulin E (IgE) therapy
   D. Immunotherapy to the occupational agent.

50

Mechanisms of Autoimmunity

Brendan Antiochos, Antony Rosen

Human autoimmune diseases occur frequently (affecting in aggregate >5% of the population worldwide) and impose a significant burden of morbidity and mortality on the human population.[1] Autoimmune diseases are defined as diseases in which immune responses to specific self antigens contribute to the ongoing tissue damage occurring in the disease. Both the specificity of the immune response and its role in tissue damage are central components of the definition. Autoimmune diseases can be either tissue specific (*e.g.*, thyroid, β-cells of the pancreas), where unique tissue-specific antigens are targeted, or can be more systemic, in which multiple tissues are affected and a variety of apparently ubiquitously expressed autoantigens are targeted.[2] Although the definition appears relatively simple in concept, the complexity of this spectrum of disorders is enormous and has greatly challenged elucidation of simple shared mechanisms. This complexity affects almost every domain, including genetics, phenotypic expression, and kinetics. In the last case, there is frequently a prolonged period (weeks to months) between initial onset of symptoms and development of the diagnostic phenotype, and the expression of disease may vary within the same individual over time. However, despite this enormous complexity, there is a striking association of the clinical phenotype with the targets of the autoimmune response. This association is, in fact, so strong that autoantibodies have been used for diagnosis and prognosis in the human autoimmune diseases.[2] For example, autoantibodies recognizing thyroid peroxidase are found in patients with autoimmune thyroiditis, autoantibodies to the Sm splicing ribonucleoprotein (RNP) complex are diagnostic of systemic lupus erythematosus (SLE), and autoantibodies recognizing topoisomerase-1 are found in patients with the diffuse form of scleroderma. The immune response in autoimmune diseases has features of an adaptive immune response (usually directed against exogenous antigens), but its targets are autoantigens. Since the adaptive immune response is initiated when suprathreshold concentrations of molecules with structures not previously tolerized by the host are encountered in a proimmune context, the association of specific autoantibodies with distinct clinical phenotypes provides critical clues to understanding the initiation and propagation of autoimmune diseases.

This chapter highlights some of the mechanistic principles that underlie autoimmune diseases. The extraordinary breadth and complexity of this disease spectrum means that the areas included cannot nearly encompass everything relevant.

KEY CONCEPTS

Autoantibodies in Autoimmune Diseases

- Some autoantibodies precede the development of any symptoms by years (*e.g.*, antinuclear antibodies and antiphospholipid antibodies in systemic lupus erythematosus [SLE], anti– cyclic citrullinated peptide [CCP] in rheumatoid arthritis [RA]).
- Some autoantibodies only occur at around the time of onset of disease manifestations (*e.g.*, anti-Sm and anti-ribonucleic protein [RNP] in SLE).
- There is a striking association of specific autoantibodies with distinct clinical phenotypes (*e.g.*, antitopoisomerase-1 with diffuse scleroderma and interstitial lung disease).

THE DISTINCT PHASES IN THE DEVELOPMENT OF AUTOIMMUNITY

A major barrier to understanding mechanisms of autoimmunity comes from difficulty in defining early events in these diseases. Since diseases are only recognizable after development of the diagnostic phenotype, there has been a tendency to interpret findings at the time of disease diagnosis with events present at disease initiation. Significant recent data from studies of the development of autoantibodies over time in patients who subsequently manifest an autoimmune disease have demonstrated that the onset of an autoimmune response and the development of clinical symptoms are generally separated in time. In the case of type I diabetes, development of islet cell autoantibodies frequently precedes diabetes, and additional islet cell autoantibodies accrue over time.[3] Similarly, autoantibodies recognizing citrullinated proteins (rheumatoid arthritis [RA]–specific autoantibodies; see below) frequently precede the development of RA.[4] These findings indicate that either a threshold needs to be exceeded in terms of tissue damage before symptoms manifest, or that there are two distinct phases in disease development, one marked by production of a group of autoantibodies and the second by autoamplifying tissue damage. In a landmark study in SLE, Arbuckle et al. have provided important insights into this issue.[5] They analyzed sera collected from patients from the US military, who subsequently developed SLE. Strikingly, autoantibodies in SLE could be divided into two groups: *(i)* those that precede the diagnosis of SLE by several years—these included antinuclear and antiphospholipid antibodies; and *(ii)* those that occurred around the time of onset of symptoms—these included anti-Sm, anti-RNP, and to a lesser extent anti-DNA. The observation that one group of autoantibodies precedes symptoms in SLE and that another group appears coincident with the phenotype strongly suggests that the groups mark distinct events in the development of autoimmune disease. Members of the first group are likely markers of disease initiation; members of the second group are likely markers of disease propagation. The antigens targeted by the immune system in this latter phase (*i.e.,* associated with clinical disease) are more likely to have some function in disease propagation, possibly through their possession of proinflammatory or adjuvant functions (see below).[6]

KEY CONCEPTS

Barriers to Defining Mechanisms of Human Autoimmune Disease

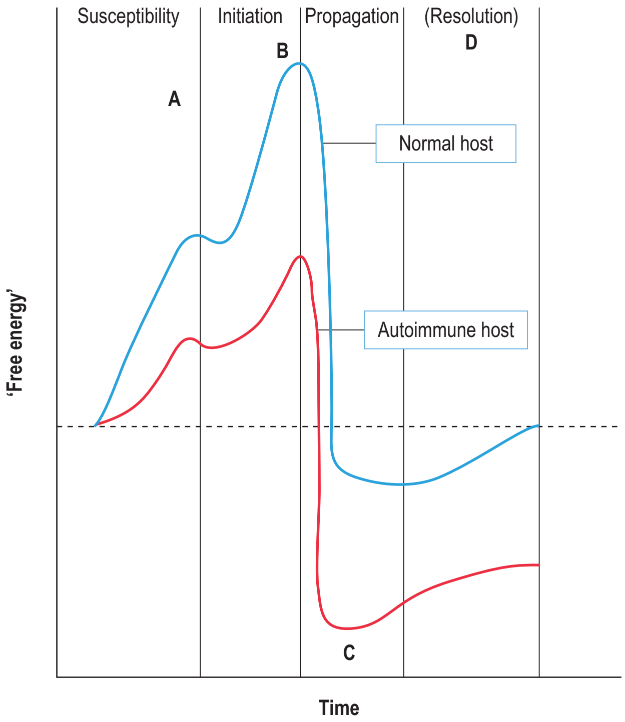
- Genetic and phenotypic complexity
- Interval between initiating events and development of diagnostic phenotype
- Challenges in quantifying human immune responses

It is therefore useful to examine the development of autoimmune diseases in four phases (Fig. 50.1):

1. *Susceptibility phase*—before disease, but where one or several preconditions for later initiation are satisfied. This would include impaired tolerance induction or altered immune signaling thresholds. The susceptibility phase could either be inherited or acquired and permanent or transient.
2. *Initiation phase*—before onset of clinical disease, but marked by the presence of an autoimmune response (*e.g.,* in the case of SLE—antiphospholipid antibodies).
3. *Propagation phase*—this corresponds with the onset of clinical disease, marked by propagation-specific immune responses (*e.g.,* in the case of SLE—anti-Sm antibodies).
4. *Regulation/resolution phase*—it should also be noted that in many cases during disease propagation, immunoregulatory pathways are also activated, which may result in natural inhibition of clinical disease over time. In rare cases, these inhibitory pathways can lead to permanent resolution. This resolution phase will not be discussed further here, but its existence provides important evidence that homeostasis can be reestablished even after the amplified phenotype develops.

PHASE I: SUSCEPTIBILITY

Although autoimmune diseases in humans are genetically complex, significant advances in understanding have occurred over the past several years. In some cases, advances have come from the study of autoimmunity with mendelian patterns of inheritance (*e.g.,* autoimmune polyendocrinopathy with candidiasis and ectodermal dysplasia [APECED], immune dysfunction/polyendocrinopathy/enteropathy/X-linked [IPEX] syndrome, C1q deficiency). Advances have also come from genetic association studies of various autoimmune phenotypes (*e.g.,* SLE, type I diabetes mellitus [DM]). There have also been important advances in the genetics of autoimmunity in several mouse models. These studies highlight a critical role for pathways of tolerance induction, immunoregulation, and setpoints/thresholds for immune signaling in avoiding emergence of autoimmunity.



A	**Susceptibility**
	• Impaired tolerance induction • Impaired production of regulatory T cells • Altered immune signaling thresholds
B	**Initiation**
	• Suprathreshold concentration of autoantigens • Non-tolerized structure • Pro-immune context – infection, malignancy, exposure to adjuvants
C	**Propagation**
	• Acquisition of adjuvant properties by disease specific autoantigens • Increased autoantigen expression in the target tissue • Immune effector pathways generate/expose autoantigen, which further drives the immune response

FIG 50.1 Mechanisms of Autoimmunity. Autoimmune diseases result from a complex interplay of pathways and events that allow autoreactivity to manifest, and cause self-sustaining tissue damage. Mechanistically, it is useful to divide the process into three phases: (1) susceptibility phase—this is present before disease and is the phase in which one or several preconditions for later initiation are satisfied; (2) initiation phase—this is marked by the presence of autoimmunity, but precedes the diagnostic clinical phenotype; (3) propagation phase—this is marked by autoimmunity and tissue damage, in which immune effector pathways cause damage and provide antigen to drive the ongoing immune response.

Incomplete Thymic Tolerance Induction Predisposes to Autoimmunity

Significant insights into basic mechanisms can derive from the study of rare human phenotypes. This has been true for autoimmunity, where several monogenic disorders have defined important pathogenic principles. Autoimmune polyendocrinopathy syndrome type 1 (APS-1;also called APECED) is a

KEY CONCEPTS

Mechanisms Underlying Susceptibility to Autoimmunity

- Incomplete induction of tolerance in the thymus to peripherally expressed autoantigens (autoimmune regulator [AIRE] deficiency causing autoimmune polyendocrinopathy with candidiasis and ectodermal dysplasia [APECED])
- Impaired clearance and tolerance induction by apoptotic cells (*e.g.,* deficiency of C1q, C4, milk fat globule–epidermal growth factor (EGF) 8 [MFG-E8], Mer)
- Defective production of regulatory T cells (FOXP3 deficiency causing immune dysfunction/polyendocrinopathy/enteropathy/X-linked [IPEX] syndrome)
- Altered immune signaling thresholds (*e.g.,* cytotoxic T lymphocyte antigen-4 [CTLA-4] polymorphisms, protein tyrosine phosphatase, non–receptor type 22 [PTPN22] polymorphisms)

rare disease in which patients develop multiple autoimmune diseases, often beginning in childhood.[7] Although candidiasis and ectodermal dystrophy (including involvement of enamel and nails, as well as keratopathy) are features of the disease, the syndrome is characterized by striking autoimmunity directed against multiple different target tissues. Autoimmune processes include autoimmune hypoparathyroidism, Addison disease, autoimmune gastritis with pernicious anemia, type I DM, thyroid disease, autoimmune hepatitis, celiac disease, and gonadal failure. Numerous autoantigens have been defined as targets of autoimmunity in APS-1 and include enzymes specifically expressed in various endocrine tissues (*e.g.,* steroid 21-hydroxylase, specific for adrenal cortex; steroid 17α-hydroxylase, found in adrenal cortex and gonads; GAD65, found in pancreatic islets; and thyroid peroxidase). The genetic basis of APS-1 was mapped to a gene on chromosome 21q22.3, subsequently termed *AIRE* (for autoimmune regulator). *AIRE* expression is highest in the thymus, where it is expressed in medullary thymic epithelial cells. Several predicted structural features of the AIRE protein and its localization in nuclear dots suggested that the protein might be a transcriptional regulator, and significant evidence for this proposal was obtained *in vitro.* Several *AIRE*-deficient mouse models were subsequently generated, which allowed for the definition of important pathogenic pathways in APS-1 that may be broadly relevant to the mechanisms of autoimmunity in general. Thus mice deficient in *AIRE* developed various autoimmune phenotypes, resembling those found in human APS-1. These included multiorgan lymphocytic infiltration and autoantibodies, as well as autoimmune eye disease. In an elegant series of experiments, Anderson et al. demonstrated that AIRE regulates expression in thymic epithelial cells of various peripheral autoantigens normally expressed exclusively in endocrine target tissues.[8] Thus AIRE appears to regulate the ectopic expression in the thymus of tissue-restricted autoantigens and to provide an antigen source against which to establish central tolerance.[8]

These data demonstrate that expression of peripheral autoantigens in the thymus constitutes a major barrier to the subsequent development of autoimmunity against these peripheral sites. Although it is possible that similar principles apply to ubiquitously expressed autoantigens targeted in systemic autoimmune diseases, there are currently few data directly addressing this issue.

Impaired Clearance and Tolerance Induction by Apoptotic Cells: Susceptibility Defect in Systemic Autoimmunity

Although little is known in humans about the thymic pathways of tolerance induction to ubiquitously expressed autoantigens, there is accumulating evidence to suggest that in the periphery, apoptotic cells play an important role in providing a source of autoantigens against which the organism becomes tolerant.[9] Apoptotic cells are generally very efficiently cleared by phagocytic cells; these events are normally associated with the production of antiinflammatory cytokines and result in tolerance induction.[9] Interestingly, early components of the classical complement pathway (*e.g.,* C1q and C4) and cross-reactive protein (CRP) are required for efficient apoptotic cell clearance, with production of interleukin-10 (IL-10) and transforming growth factor-β (TGF-β). Therefore it is of particular note that homozygous C1q deficiency is associated with a striking susceptibility to SLE in humans, and this suggests that rapid, efficient, tolerance-inducing clearance of apoptotic cells may play a role similar to AIRE expression in the thymus in preventing subsequent emergence of autoimmunity to ubiquitously expressed autoantigens.[10] Additional support for this model comes from recent studies of milk fat globule–epidermal growth factor (EGF) 8 (MFG-E8), a glycoprotein secreted from macrophages that is required for the efficient attachment and clearance of apoptotic cells by macrophages and immature dendritic cells (DCs). MFG-E8 is also expressed in tingible-body macrophages in the germinal centers of secondary lymphoid tissues. Interestingly, many unengulfed apoptotic cells are present in the germinal centers of the spleen in *MFG-E8*–deficient mice, which develop a striking lupus-like phenotype (reviewed in Rai and Wakeland[11]). Together, the data strongly suggest that efficient, antiinflammatory clearance of apoptotic cells plays a central role in tolerance induction and prevention of autoimmunity.

Defective Production of Regulatory T Cells

Although there are pathways that *(i)* regulate autoantigen expression at sites of tolerance induction and *(ii)* guide autoantigens toward tolerance-inducing outcomes, these pathways alone are clearly insufficient to prevent the emergence of autoimmune disease. This fact is highlighted by the emergence of autoimmunity when regulatory T-cell (Treg) differentiation is abnormal in humans with IPEX syndrome, the human equivalent of the *scurfy* mouse. IPEX is a rare, X-linked recessive disorder characterized by type I DM, thyroiditis, atopic dermatitis, and inflammatory bowel disease (IBD) and is caused by mutations in the *FOXP3* gene.[12] FOXP3 is a member of the forkhead family of transcription factors and is essential for the development of $CD4^+$ Tregs, which have been shown to regulate the activation and differentiation of effector T cells at many different levels (Chapter 18).

Preliminary investigations into the role of Tregs in human autoimmune diseases have produced mixed results, but these cells likely play important roles in regulating disease onset and amplitude. As additional cell subsets and details about function emerge, this area will be clarified significantly.

Signaling Thresholds and Susceptibility to Autoimmunity

Several modulators of T-cell signaling have been defined as important susceptibility determinants in autoimmunity.[11] For example, cytotoxic T lymphocyte antigen-4 (CTLA-4) polymorphisms are associated with increased risk of a variety of autoimmune diseases, including type I DM, Graves disease, and RA. Similarly, a functional polymorphism in protein tyrosine phosphatase, non–receptor type 22 (PTPN22), has been identified as a major risk factor for several human autoimmune diseases, including SLE, RA, and type I DM. Although the exact mechanisms underlying susceptibility to autoimmunity remain unclear, in both cases, the polymorphisms appear to regulate the balance of stimulatory and inhibitory signaling in effector and Tregs, favoring effector T-cell activation.

Recent genetic studies have also suggested a potential role for innate immune sensors in autoimmunity, focusing attention on the critical balance between activation of the immune response to mitigate infectious damage and limiting the magnitude of the response to avoid immunopathology. Gain-of-function (GOF) variants of *IFIH1* (encoding melanoma differentiation-associated protein 5 [MDA5]) are associated with susceptibility to SLE.[13] This RNA helicase is essential for the detection of cytoplasmic viral RNA and activation of type I interferon (IFN) secretion by infected cells. The increased susceptibility to SLE with increased function of this antiviral pathway suggests that excessive IFN signaling may facilitate the development of autoimmunity. Key immune signaling pathways that protect the host from deleterious infectious and malignant challenges but potentially also enhance damage of self tissues in the process may, therefore, be important susceptibility factors in autoimmune diseases.

Thus there are many barriers to the development of autoimmunity, including effective tolerance induction in the thymus and the periphery, tightly regulated immune signaling, and homeostatic pathways of immunoregulation to limit anti-self responses should they occur. It is likely that the genetic susceptibility to autoimmunity in outbred humans represents an integrated threshold involving genes that regulate these various pathways—upon which environmental and stochastic events act to accomplish disease initiation and propagation.

PHASE 2: INITIATION

Initiation of an adaptive immune response requires presentation to T cells of suprathreshold concentrations of molecules with structure not previously tolerized by the host. One of the more persuasive models proposed to explain the persistence of potentially autoreactive T cells within the repertoire of the host is that of immunodominance of T-cell epitopes. This model provides major insights into the pathogenesis of autoimmunity.[11,14]

KEY CONCEPTS

Potential Mechanisms That Can Alter Antigen Processing to Reveal Potentially Cryptic Epitopes

- Modification of autoantigen processing through high-affinity binding to ligands or antibodies
- Distinct proteolytic machinery in the thymus and periphery—or differential modification of proteolytic activity
- Modification of autoantigen structure that modifies its processing by endogenous antigen-presenting cell (APC) machinery, generally through posttranslational modifications
- Novel proteolytic events not present in the normal APC pathways (*e.g.,* novel cleavage during cell death or damage or inflammation)
- Novel forms of autoantigens generated by mutation, truncation, or splicing

Dominance and Crypticity

Studies by Sercarz et al.[14] have stressed that although antigens contain numerous potential determinants that could be presented on major histocompatibility complex (MHC) class II molecules during antigen processing, not all determinants in a particular molecule are equally likely to be efficiently presented. Those determinants that are most efficiently presented are termed "dominant"; those that are not loaded onto MHC class II to a significant degree are termed "cryptic." For self antigens, it is likely that a constant set of dominant determinants are generated during antigen processing under most circumstances, with similar outcomes in the thymus and the periphery. Antigens processed by the "standard" pathway are, therefore, fully tolerized, with the T-cell repertoire purged of reactivity to the dominant self. However, the balance of dominant and cryptic epitopes presented during antigen processing is influenced significantly by changes of protein structure, which occur during various relevant physiological states[15] (Chapter 6). Several potential mechanisms that may alter antigen processing to reveal potentially cryptic epitopes are summarized below.

High-Affinity Binding of Antigen to Ligands or Antibodies

Several studies have demonstrated that antigen processing can be dramatically altered when the antigen binds with high affinity to a ligand or antibody. The study by Simitsek and colleagues (reviewed in Lanzavecchia[15]) demonstrated that presentation of T-cell determinants in tetanus toxin can be either enhanced or suppressed as a direct consequence of antibody modulation of antigen processing in human B-lymphoblastoid cells. Remarkably, a single bound antibody can simultaneously enhance the presentation of one T-cell determinant by more than 10-fold while strongly suppressing the presentation of a different T-cell determinant. Biochemical analyses have shown that both the suppressed and boosted determinants fall within an extended domain of antigen stabilized by this antibody during proteolysis. Thus ligand-induced changes in processing can destroy dominant determinants or reveal cryptic self determinants. Similar observations have also been made with numerous other antigen–antibody partners.[15]

Tissue-Specific Protease Expression

The study by Watts and associates (reviewed in Darrah and Rosen[16]) showed that a principal human leukocyte antigen–D related type 2 (HLA-DR2)–restricted epitope in myelin basic protein amino acids 85–99 (MBP85–99) contains a processing site for asparagine endopeptidase (AEP), with cleavage by AEP abolishing the epitope. AEP activity is, therefore, a critical factor in presentation of this epitope.[16] In human antigen-presenting cells (APCs), presentation of MBP85–99 is inversely proportional to the amount of cellular AEP activity, and inhibition of AEP greatly enhances presentation of the MBP85–99 epitope. Interestingly, both MBP and AEP are expressed in the thymus, AEP at abundant levels. These data suggest that this major epitope in neurological autoimmunity may not be presented under normal circumstances in the thymus as a result of destruction by AEP, therefore raising the potential for later presentation in the periphery in the setting of decreased AEP activity.

Posttranslational Modification of Autoantigen Structure

Autoantigens undergo a variety of posttranslational modifications, including phosphorylation, proteolytic cleavage, ubiquitination, transglutamination, citrullination, and isoaspartyl modification.[17] In several cases, autoantibodies recognize exclusively the modified form of the antigen (*e.g.,* RNA polymerase-II large subunit, serine/arginine-rich [SR] proteins, citrullinated vimentin, and other RA autoantigens), indicating that the modified forms of the molecules are important in driving the immune response. In addition, Doyle and Mamula[17] have demonstrated that post-translational modification of autoantigen structure may be more broadly relevant than can be appreciated by studying autoantibody specificity alone. They showed that although mouse immunization with a murine cytochrome c peptide (amino acids 90–104) resulted in no T- or B-cell response, immunization with the isoaspartyl form of this peptide resulted in strong T- and B-cell responses. The autoantibodies that were elicited recognized both the modified and the native forms of the antigen, but T cells only recognized the isoaspartyl form. Similar observations have also been made for several SLE autoantigens. The difficulty detecting and quantifying antigen-specific T cells in various autoimmune diseases may reflect their preferential recognition of subtly modified forms of autoantigen. This is an important area for future study—currently there is no systematic way to generate the relevant autoantigen forms.

Novel Antigen Cleavage During Cell Damage, Cell Death, or Inflammation

Recent studies have provided evidence that single proteolytic events early in antigen processing can play critical roles in defining the epitopes generated. For example, Watts and colleagues (reviewed in Darrah and Rosen[16]) have conclusively demonstrated that early cleavage by AEP determines subsequent proteolytic events. Modifications of the antigen that affect this early cleavage dramatically change the epitopes loaded onto MHC class II molecules.[16]

Inflammatory microenvironments can create significant potential to load distinct epitopes because unique proteolytic activities are present. Activated inflammatory cells constitute a major source of proteases, including various cytotoxic lymphocyte granule proteases (granzymes), as well as numerous neutrophil and monocyte granule proteases. It is of interest that numerous autoantigens targeted in systemic autoimmune diseases are substrates for these inflammatory proteases and that unique autoantigen fragments are generated through the activity of granzyme B and potentially other similar proteases.[16] Such autoantigen forms are not generated during other forms of cell damage or death, and similar activity is not observed against nonautoantigens. Thus novel proteolytic cleavage of intracellular autoantigens during activity of cytotoxic immune effector pathways may provide a source of cryptic epitopes not generated during homeostatic "tolerance-inducing" tissue turnover. Direct evidence that inflammatory protease-mediated revelation of cryptic epitopes is relevant to initiation of autoimmunity *in vivo* is still needed.

Autoantigen Alteration Caused by Mutation, Truncation, or Splicing

Since the final epitopes generated and loaded onto MHC class II molecules can be profoundly influenced by single, early cleavage events during antigen processing, relatively minor but critically placed changes in the primary structure of autoantigens can have the capacity to influence peptide selection. A study of the melanoma and vitiligo-associated autoantigens, tyrosinase-related proteins (TRPs) 1 and 2 has strikingly demonstrated that this mechanism may play an important role in generating immune responses to self and tumor antigens.[18] In this study, Engelhorn et al. examined whether mutated self gene products are more likely to initiate immunity and used a systematic approach to define some of the principles that determine this. They generated complementary DNA (cDNA) libraries encoding large numbers of random mutations in syngeneic TRPs. They then used an approach of DNA immunization of black mice to test the immunogenicity of the altered proteins encoded by the pools of mutated cDNA. Immunization with nonmutated proteins induced no detectable immune responses, consistent with establishment of tolerance to the full-length molecules. In contrast, the mutated cDNA pools elicited both autoimmune depigmentation and the ability to reject melanoma tumors. Additional analysis showed that autoimmunity resulted from mutations that altered autoantigen cell biology, particularly degradation rates and pathways. Mutations also created new T-helper (Th) cell epitopes and induced recognition of nonmutated but previously cryptic epitopes. Interestingly, mutations themselves did not form part of CD8 epitopes that drive the anti-self and antitumor immune responses. Mutated molecules that were immunogenic were frequently truncated, leading the authors to propose that inappropriately truncated self proteins can provoke autoimmunity when present in a proinflammatory environment. This study has provided an important mechanistic underpinning for the proposal that accumulated mutations have a role in the initiation of autoimmunity and that "autoimmunity" might target the cancer mutanome.

Recent work in the autoimmune rheumatic diseases (scleroderma and dermatomyositis) highlights the relationship between cancer and the autoimmune process. In these diseases, patients demonstrate an increased risk of cancer compared with controls. They also show a temporal clustering of cancer diagnosis around the time of dermatomyositis or scleroderma onset (reviewed in Shah[19]). Similarly, there is evidence for an association between SLE and cancer, particularly lymphoma, clustered within the first 2 years of SLE diagnosis.[20] These associations, both with timing of diagnosis and preferentially with specific tumor types, are strongly indicative of a nonrandom clustering of autoimmune processes and cancer, which is likely of mechanistic significance.

The mechanistic link between scleroderma and cancer was confirmed in a study that tested the hypothesis that an anticancer immune response may target a mutated autoantigen in the patient's cancer, which spreads to the wild-type version of the antigen to trigger a self-sustaining autoimmune disease (reviewed in Shah[19]). In this study of patients with scleroderma and cancer, genetic changes (either mutations or loss of heterozygosity) were identified in *POLR3A* (encoding RNA polymerase III large subunit) in the tumors of six of eight patients with anti-RNA polymerase III antibodies, whereas no such changes were identified in eight patients who lacked these antibodies. Anti-RNA polymerase III autoantibodies in patients with cancer demonstrated no specificity for the mutant form of the protein over the wild-type version, and analyses of peripheral blood lymphocytes identified T-cell reactivity against the mutant protein. These data suggest that somatic mutations arising in the context of cancer may prompt immune responses that mediate immunoediting as well as damage to nontumor host tissues.

It is, therefore, likely that somatic mutations acquired with age and their association with malignancy are important in the genesis of some forms of human autoimmunity. Additional studies to confirm this and elucidate the underlying mechanisms remain a high priority. However, the barriers to such studies in humans are very significant, as effective anticancer immunity may be phenotypically silent, and convenient technologies to quantify somatic mutation and specific immune responses in normal individuals will be needed to draw conclusions of causality.

Antigen Mimicry

The process of antigen mimicry (see below) has frequently been proposed as a potential initiator of autoimmune diseases.[21] This mechanism, particularly when isolated, is only likely relevant to those autoimmune processes clearly associated with antecedent infections, particularly infections that resolve spontaneously and subsequently recur upon reexposure to the offending agent. The mechanism may, however, also play a role in initiation of the autoimmune response in self-sustaining autoimmune processes but, in this case, requires that T-cell responses to the cross-reacting self antigen are initiated.

Foreign antigens, which often differ from their homologous self antigens in some areas, may, nevertheless, bear significant structural similarity to self antigens in other regions. Initiation of an immune response to the foreign antigen may generate a cross-reactive antibody response that also recognizes the self protein (antigen mimicry). When the antigen is a cell surface molecule, antibody-mediated effector pathways can lead to host tissue damage. Although the antibody response is cross-reactive with self molecules, the T cells that drive this response are generally directed at the foreign antigen, at least initially (see below). Diseases involving this sort of "antigen mimicry," therefore, tend to be self-limiting, although they can recur upon reexposure to the offending antigen. It is important to realize that antigen mimicry alone cannot explain self-sustaining autoimmune diseases, which are driven by self antigens and autoreactive T cells. In these cases, there is a requirement for overcoming T-cell tolerance to the self protein. The central issues in this regard are *(i)* how T-cell tolerance to self antigens might initially be broken, and *(ii)* once this has occurred, why these antigens continue to drive the immune response to self. The simultaneous liberation of self antigen in the presence of the cross-reactive antibody response has been proposed to play critical roles in this regard. For example, several studies have suggested that when a humoral response to a foreign protein is induced that cross-reacts with the self antigen, a strong Th-cell response specific for the self antigen can occur. The simultaneous liberation of significant amounts of self antigen in the setting of a cross-reactive antibody response may allow for effective presentation of cryptic epitopes in the self antigen to autoreactive T cells by activated cross-reactive B cells. If continued release of self antigen occurs, a specific, adaptive immune response to self will be sustained. Antigen release from tissues likely plays a critical role in driving this autoimmune process. Understanding the mechanisms of ongoing antigen release at sites of tissue damage in autoimmune disease (*e.g.*, unique pathways of cell injury and death) is a high priority for future work, as it provides a novel target for therapy (see below).

It is clear from the above discussion that extraordinary complexity is operative in initiation of the human autoimmune diseases. The patient population is genetically heterogeneous, the human immune system is complex and extremely plastic, and it interacts with a plethora of environmental stimuli and stochastic events. The simultaneous confluence of susceptibility factors and initiation forces to set off the self-sustained and autoamplifying process is, therefore, a rare occurrence. In contrast, once activation of autoreactive T cells has occurred, the ability of the immune system to vigorously respond to low concentrations of antigen, to amplify the specific effector response to those antigens, and to spread the response to additional antigens in that tissue, greatly reduces the stringency that must be met to keep the process going.

PHASE III: PROPAGATION

Principles of Amplification

One of the central features of human autoimmunity is the tendency of the process to amplify progressively with the accumulation of significant immune-mediated tissue damage. Furthermore, in the vast majority of cases, once such amplification begins, the process is very unlikely to resolve spontaneously. Autoantigens themselves can be very important in this phase, in terms of both acquisition of adjuvant properties and regulation of levels of expression. The essential features of amplification are a substrate cycle, in which antigen expression and adjuvant properties induce an immune response, which, in turn, induces increased antigen expression and tissue damage—and further drives the immune response. The importance of tissue-specific autoantigen expression in focusing such immune responses is only beginning to be recognized.

Acquisition of Adjuvant Properties by Disease-Specific Autoantigens

In spite of the fact that tens of thousands of molecules could be targeted by the immune system in autoimmunity, the number of molecules that are frequently targeted in the different phenotypes are markedly restricted—limited perhaps to 100 or so. This has led to the proposal that frequently targeted autoantigens may themselves have properties that make them proimmune. Work by Howard and associates (reviewed in Plotz[22]) provided important initial support for this proposal. They observed that the autoantigenic histidyl aminoacyl tRNA synthetase (HRS), which is targeted in autoimmune myositis (but not non–autoantigenic lysyl- and aspartyl-aminoacyl transfer RNA [tRNA] synthetases), is a chemoattractant to immature DCs and other leukocytes. The authors suggested that the selection of a self molecule as a target for an autoantibody response may be a consequence of proinflammatory properties of the molecule itself. They further suggested that modification of autoantigen structure during processes of cell damage or death may be critical in recruiting these additional functions of autoantigens.

Role of Innate Immune Receptors in Amplification

Several innate immune receptor systems that sense and transduce the signals from pathogen-associated molecular patterns (PAMPs) (*e.g.*, receptors of the Toll-like receptor [TLR] and retinoic acid-inducible gene I [RIG-I]–like receptor [RLR] families) may also play roles in transducing the proinflammatory properties of autoantigens.[6] Ligands for TLRs include both microbial and endogenous molecules, the latter group being particularly relevant to autoimmunity (see below). Microbial ligands include components of gram-positive bacteria, gram-negative bacteria, yeasts, and protozoans. For example, lipoteichoic acid and fungal products (*e.g.*, zymosan) signal through TLR2, lipopolysaccharide activates TLR4 signaling, dsRNA signals through TLR3, flagellin through TLR5, single-strand RNA (ssRNA) through TLR7 and TLR8, bacterial or viral DNA through TLR9, and *Toxoplasma* profilin through TLR11.[23] Although viral and bacterial nucleic acids are the most likely ligands for TLRs, accumulating data demonstrate that complexes containing endogenous nucleic acids are able to signal through TLRs. The exact nature and source of endogenous ligands for TLRs *in vivo* remains unclear; however, recent studies have demonstrated that components from stressed, injured, and dying cells may play critical roles in this regard.[6]

Working in several models, numerous investigators have now provided evidence that the frequent targeting of nucleoprotein autoantigens (which contain DNA or RNA) results from the ability of these nucleic acid components to ligate TLRs (reviewed in Kawasaki et al.[6]). For example, antichromatin immune complexes can activate DCs and B cells through ligation of both FcγR and TLR9. B-cell activation via TLR7 ligation by RNA-containing immune complexes has also been demonstrated. *In vivo* studies have linked these *in vitro* findings to the development of autoimmunity. For example, when *TLR9* deficiency is bred onto MRL-lpr mice—which are an excellent model of SLE—animals no longer develop autoantibody responses to chromatin. Interestingly, these animals, nevertheless, manifest the SLE phenotype—in some cases, more severely than the TLR9-sufficient animals. Similarly, when mice are rendered TLR7 deficient, the autoantibody response to Sm is markedly inhibited, and severity of the SLE phenotype is reduced. Reciprocally, overexpression of TLR7 is sufficient to induce an SLE-like disease (reviewed in Kawasaki et al.[6]). These data confirm that autoantigens frequently selected in different autoimmune phenotypes likely have the dual property of being able to simultaneously activate the innate and adaptive immune systems. The source, form, and uptake of nucleic acids clearly influence adjuvant activity. For example, bacterial and viral DNA and oligonucleotides with CpG motifs have significant adjuvant activity, whereas mammalian genomic DNA, in which CpG is usually methylated, has very poor adjuvant activity. In contrast, human DNA present within immune complexes in SLE serum activates plasmacytoid DCs effectively in a DNA-dependent way. Several potential explanations have been advanced to explain these observations.[6] First, FcγR-mediated uptake effectively captures self DNA bound by anti-DNA antibodies and directs it to the correct endosomal compartment for TLR signaling. Second, coligation of TLR9 and either B-cell receptor or FcγR alters the signaling threshold of immune complexes. Last, the difference lies in the nucleic acid itself, with additional modifications of DNA and RNA structure occurring in cells under different physiological circumstances (*e.g.,* cell death) and regulating nucleic acid binding to TLRs.[24] Recent studies have suggested that other sensors of PAMPs (*e.g.,* nucleotide-binding oligomerization domain [NOD]–like receptors [NLRs], RIG-like receptors) may play similar roles to TLRs in amplifying adaptive responses to specific intracellular autoantigens (*e.g.,* MDA5, interferon-γ–inducible protein 16 [IFI16]), but definitive evidence is not yet available.

The TLR–IFN interface has been recognized as critical in the propagation phase of systemic autoimmune diseases (reviewed in Hall and Rosen,[25] and Swiecki and Colonna[26]). Much attention has been focused on plasmacytoid dendritic cells (pDCs), a relatively rare class of immature DCs that can secrete large amounts of type I IFNs upon TLR ligation and that express TLR7 and TLR9 at high levels.[26] Recent studies in mouse models have demonstrated the ability of pDCs to alternately promote wound healing or chronic inflammation, depending on the predisposing genetic background. For example, nonspecific skin injury in wild-type mice (by tape stripping) leads to pDC recruitment, TLR7 and TLR9 recognition of nucleic acids, and transient expression of type I IFNs with subsequent wound healing.[27] In contrast, the same skin damage in a lupus-prone mouse strain results in chronic inflammation mediated by sustained type I IFN expression, which can be ameliorated with pDC depletion or TLR7/9 inhibition.[28] Extending these observations to human autoimmunity, IFN-α–inducing activity is abrogated by chloroquine or bafilomycin, agents that interfere with endosome acidification and TLR7 and TLR9 signaling. In addition, Rönnblom and colleagues (reviewed in Hall and Rosen[25]) demonstrated that when added to material from apoptotic or necrotic cells, autoantibodies from patients with SLE and Sjögren syndrome with specificity for DNA or RNA autoantigens induce striking IFN secretion. Type I IFNs have a broad set of functions that likely contribute to the feed forward, propagation phase of systemic autoimmune diseases.[6] For example, they *(i)* promote the differentiation of monocytes into mature DCs, which drive autoreactive T- and B-cell responses; *(ii)* increase target cell sensitivity to killing pathways; *(iii)* upregulate cytotoxic effector pathways; and *(iv)* upregulate expression of autoantigens, such as Ro52.

This general ability of autoantigens, particularly in the context of immune complexes, to stimulate secretion of IFN and other cytokines is likely an important principle in the initiation and propagation of autoimmunity.

Enhanced Autoantigen Expression in the Target Tissue

The striking association of specific autoantibody responses with distinct phenotypes suggests that autoantigen expression or form in specific target tissues may play an important role in both focusing the immune response and generating tissue damage. Unfortunately, very little is currently known about such parameters *in vivo* in relevant target tissues, both in normal and pathological circumstances.

Recent studies on human autoimmune myopathies have begun to provide important insights into this problem. Myositis-specific autoantigens are expressed at very low levels in control muscle but at high levels in myositis tissue, where antigen expression is at highest levels in regenerating muscle cells.[29] Interestingly, histidyl tRNA synthetase (HRS) expression is also found at high levels in lungs, and anti-HRS antibodies are associated with autoimmune myopathies with interstitial lung disease. Recently, 3-hydroxy-3-methylglutaryl-coenzyme A reductase (HMGCR) was identified as an autoantigen in patients with statin-associated autoimmune myopathy.[30] In addition to the induction of HMGCR protein expression by statin treatment of cultured cells, this group also demonstrated increased expression of this protein in regenerating muscle fibers of anti–HMGCR-positive patients. These data suggest that enhanced autoantigen expression in the target tissue may be a feature of disease propagation and that antigen expression during tissue repair may provide an ongoing antigen source to sustain and amplify tissue damage.

Defining whether similar principles are operating in other autoimmune diseases and elucidating the pathways of antigen expression and modification in relevant target tissues are an important areas of focus in future studies and may have important therapeutic potential.

ON THE HORIZON

- Precise clinical and molecular phenotyping of patients at various disease stages is critical for improving diagnosis, monitoring, and treatment of autoimmune diseases.
- Understanding mechanisms of disease amplification, propagation, and regulation will enable the development of effective targeted therapies.
- Understanding the mechanisms of human autoimmunity will likely provide important insights into the normal functioning of the immune system, particularly with regard to natural cancer immunity.

TRANSLATIONAL RESEARCH

The challenge in the next decade is to continue elucidating the underlying mechanisms of autoimmunity to improve diagnosis and treatment and enable more precise monitoring of these diseases. The existence of distinct phases in the development of autoimmune disease provides an opportunity to identify patients at risk of developing tissue damage before this occurs and potentially to interdict the process before the amplification cycle is established. Understanding natural mechanisms of regulation of disease amplitude, or in some circumstances, even disease resolution, will provide important opportunities for therapy. The role of the target tissue as an active participant, rather than a passive bystander, in the ongoing destruction will likely be understood in molecular detail and provide novel approaches to decreasing the amplitude of amplifying cycles. The identification of the precise biomarkers of all these events requires exceptional clinical phenotyping of patients early in the disease course, longitudinal analysis by centers of excellence that follow up large numbers of patients with similar phenotypes over time, and effective coupling to basic laboratory enterprise. Critical areas of future investigation relevant to diagnosis and therapy include clarification of the roles of genetic and epigenetic factors, interactions between the innate and adaptive immune responses, Tregs, and focus on the target tissue and antigen contributions to ongoing amplification.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Davidson A, Diamond B. General features of autoimmune disease. In: Rose NR, Mackay IR, editors. The autoimmune diseases. St Louis: Elsevier Academic Press; 2006.
2. Bizzaro N. Autoantibodies as predictors of disease: the clinical and experimental evidence. Autoimmun Rev 2007;6(6):325–33.
3. Insel RA, Dunne JL, Atkinson MA, et al. Staging presymptomatic type 1 diabetes: a scientific statement of JDRF, the Endocrine Society, and the American Diabetes Association. Diabetes Care 2015;38(10):1964–74.
4. Dekkers J, Toes RE, Huizinga TW, et al. The role of anticitrullinated protein antibodies in the early stages of rheumatoid arthritis. Curr Opin Rheumatol 2016;28(3):275–81.
5. Arbuckle MR, McClain MT, Rubertone MV, et al. Development of autoantibodies before the clinical onset of systemic lupus erythematosus. N Engl J Med 2003;349(16):1526–33.
6. Kawasaki T, Kawai T, Akira S. Recognition of nucleic acids by pattern-recognition receptors and its relevance in autoimmunity. Immunol Rev 2011;243(1):61–73.
7. Cutolo M. Autoimmune polyendocrine syndromes. Autoimmun Rev 2014;13(2):85–9.
8. Anderson MS, Venanzi ES, Klein L, et al. Projection of an immunological self shadow within the thymus by the AIRE protein. Science 2002;298(5597):1395–401.
9. Elliott MR, Ravichandran KS. Clearance of apoptotic cells: implications in health and disease. J Cell Biol 2010;189(7):1059–70.
10. Macedo AC, Isaac L. Systemic Lupus Erythematosus and Deficiencies of Early Components of the Complement Classical Pathway. Front Immunol 2016;7:55.
11. Rai E, Wakeland EK. Genetic predisposition to autoimmunity–what have we learned? Semin Immunol 2011;23(2):67–83.
12. Bennett CL, Christie J, Ramsdell F, et al. The immune dysregulation, polyendocrinopathy, enteropathy, X-linked syndrome (IPEX) is caused by mutations of FOXP3. Nat Genet 2001;27(1):20–1.
13. Oliveira L, Sinicato NA, Postal M, et al. Dysregulation of antiviral helicase pathways in systemic lupus erythematosus. Front Genet 2014;5:418.
14. Sercarz EE, Lehmann PV, Ametani A, et al. Dominance and crypticity of T cell antigenic determinants. Annu Rev Immunol 1993;11:729–66.
15. Lanzavecchia A. How can cryptic epitopes trigger autoimmunity? J Exp Med 1995;181(6):1945–8.
16. Darrah E, Rosen A. Granzyme B cleavage of autoantigens in autoimmunity. Cell Death Differ 2010;17(4):624–32.
17. Doyle HA, Mamula MJ. Posttranslational modifications of self-antigens. Ann N Y Acad Sci 2005;1050:1–9.
18. Engelhorn ME, Guevara-Patino JA, Noffz G, et al. Autoimmunity and tumor immunity induced by immune responses to mutations in self. Nat Med 2006;12(2):198–206.
19. Shah AA, Casciola-Rosen L, Rosen A. Review: cancer-induced autoimmunity in the rheumatic diseases. Arthritis Rheumatol 2015;67(2):317–26.
20. Bernatsky S, Boivin JF, Joseph L, et al. An international cohort study of cancer in systemic lupus erythematosus. Arthritis Rheum 2005;52(5):1481–90.
21. Cusick MF, Libbey JE, Fujinami RS. Molecular mimicry as a mechanism of autoimmune disease. Clin Rev Allergy Immunol 2012;42(1):102–11.
22. Plotz PH. The autoantibody repertoire: searching for order. Nat Rev Immunol 2003;3(1):73–8.
23. Kawai T, Akira S. Toll-like receptors and their crosstalk with other innate receptors in infection and immunity. Immunity 2011;34(5):637–50.
24. Green NM, Marshak-Rothstein A. Toll-like receptor driven B cell activation in the induction of systemic autoimmunity. Semin Immunol 2011;23(2):106–12.
25. Hall JC, Rosen A. Type I interferons: crucial participants in disease amplification in autoimmunity. Nat Rev Rheumatol 2010;6(1):40–9.
26. Swiecki M, Colonna M. Unraveling the functions of plasmacytoid dendritic cells during viral infections, autoimmunity, and tolerance. Immunol Rev 2010;234(1):142–62.
27. Gregorio J, Meller S, Conrad C, et al. Plasmacytoid dendritic cells sense skin injury and promote wound healing through type I interferons. J Exp Med 2010;207(13):2921–30.
28. Guiducci C, Tripodo C, Gong M, et al. Autoimmune skin inflammation is dependent on plasmacytoid dendritic cell activation by nucleic acids via TLR7 and TLR9. J Exp Med 2010;207(13):2931–42.
29. Casciola-Rosen L, Nagaraju K, Plotz P, et al. Enhanced autoantigen expression in regenerating muscle cells in idiopathic inflammatory myopathy. J Exp Med 2005;201(4):591–601.
30. Mammen AL, Chung T, Christopher-Stine L, et al. Autoantibodies against 3-hydroxy-3-methylglutaryl-coenzyme A reductase in patients with statin-associated autoimmune myopathy. Arthritis Rheum 2011;63(3):713–21.

MULTIPLE-CHOICE QUESTIONS

1. Antibodies directed against which of the following antigens has been shown to be associated with an increased risk of cancer in scleroderma?
   A. Topoisomerase
   B. RNA polymerase III
   C. Centromere
   D. Fibrillarin
   E. NXP-2

2. Which of the following autoantibodies is likely associated with systemic lupus erythematosus (SLE) propagation, based on the observation that it arises closest to the time of symptoms onset?
   A. Antinuclear
   B. Anti-La
   C. Antiphospholipid
   D. Anti-Sm

51

Systemic Lupus Erythematosus

Cynthia Aranow, Betty Diamond, Meggan Mackay

Systemic lupus erythematosus (SLE) is a systemic autoimmune disease characterized by the production of autoantibodies and a broad spectrum of clinical manifestations. It most commonly presents in women during their childbearing years. Although the etiology of SLE is unknown, both genetic and environmental factors contribute to loss of self-tolerance. Current therapeutic modalities are antiinflammatory and immunosuppressive.

EPIDEMIOLOGY

The American College of Rheumatology (ACR) classification (Table 51.1), created initially to classify this disease with multiple diverse manifestations, has undergone several revisions in efforts to improve sensitivity and specificity of the criteria.[1] The revised 1997 classification scheme is the most widely accepted instrument for "diagnosing" lupus; however, these criteria do not represent the full spectrum of disease. Patients fulfilling 4 out of the 11 criteria are classified with SLE with approximately 95% certainty, although many individuals who meet only two or three criteria may also be considered as having SLE. More recently, the Systemic Lupus International Collaborating Clinics (SLICCs) have proposed and validated a revised classification scheme that seeks to improve sensitivity and specificity by inclusion of a broader range of immunological and clinical criteria. A patient will satisfy the SLICCs classification of SLE if there is biopsy-proven lupus nephritis with an autoantibody (antinuclear antibody [ANA] or anti-dsDNA) or if a minimum of 4 predefined criteria are met (1 of these 4 criteria must be clinical, and one must be immunological).[2] During the childbearing years, the ratio of women to men with lupus is approximately 9:1. This ratio is less in younger and older populations, supporting a role for hormonal factors in disease induction. The majority of lupus presents during adulthood; approximately 20% of cases are in the pediatric population. Recent studies suggest that age at diagnosis may be increasing in some populations; mean ages at diagnosis reported since 2002 range from 31 years in Martinique and Brazil to 51.7 years in Wisconsin (United States) and 47 years in Sweden.

Lupus occurs throughout the world; susceptibility is linked to race and ethnicity as well as environmental exposures. Worldwide incidence rates vary from 1.9–8.7/100 000, and prevalence rates vary from 19.3–207/100 000;[3] however, incidence rates in African Americans, African Caribbeans, Hispanics, and Asians are approximately three times greater than in Caucasians. Although epidemiological data are scant, it appears that the incidence of SLE in Africa is lower. Data from animal models suggest this may be a consequence of a protective effect of malaria infection. Clinical manifestations, disease activity, damage accrual, mortality, and response to therapy are also modulated by ethnicity, and all studies support observations of increased disease activity, damage, and mortality in nonwhite populations (Table 51.2).[4]

Mortality

Increased awareness leading to earlier diagnoses, availability and use of immunosuppressive agents, and improved treatment for comorbid diseases have contributed to a dramatic decrease in mortality over the last half-century, with a 5-year survival rate of 50% in the 1950s to 96–99% in the past 5 years.[5] Despite this increase in survival, lupus patients continue to suffer significant morbidity and mortality related to disease and medication effects. Compared with the general population, risk of death is reported as 2–5 times higher in lupus patients, with standardized mortality rates as high as 6.7 and 7.8 in SLE cohorts with lupus nephritis. Overall, mortality is modulated by gender, ethnicity, race, geographical location, disease comorbidity, age at diagnosis, and the presence of lupus nephritis or end-stage renal disease. Generally, highest mortality rates are associated with male sex, black race, and active lupus nephritis. The leading causes of death include cardiovascular disease, infection, and active SLE; the order of frequency depends on the population reported. A bimodal distribution of death has been well recognized; early deaths often result from infections or active disease, whereas deaths occurring later in the course of disease are frequently attributed to end-stage organ damage and cardiovascular disease, usually in the setting of inactive disease. Both cardiovascular disease and organ failure are associated with damage accrued from recurrent SLE disease activity and medication effects.

DAMAGE

With improved survival, the impact of comorbid conditions has become increasingly important. Patients may accrue irreversible damage over time resulting from active disease and/or toxicities of medications. The SLICCs/ACR Damage Index (SDI) is a validated instrument to measure damage. Damage accumulates over time and is associated with poorer quality of life as well as increased morbidity and mortality. Renal, musculoskeletal, and cardiac domains are important contributors to damage. There is increased awareness of cardiovascular disease in SLE. As discussed below (cardiac involvement: coronary artery disease), events attributable to atherosclerotic disease occur significantly more frequently in lupus cohorts compared with their age-matched controls. Traditional risk factors, *e.g.*, hypertension, hyperlipidemia, physical inactivity, and impaired glucose control, are enriched in lupus patients. However, the risk of cardiovascular events continues to

TABLE 51.1 American College of Rheumatology Criteria for Systemic Lupus Erythematosus

Criteria	Description
Malar rash	Fixed malar erythema, flat or raised
Discoid rash	Erythematous raised patches with adherent keratotic scaling and follicular plugging; atrophic scarring may occur in older lesions
Photosensitivity	Skin rash as an unusual reaction to sunlight, by patient history or physician observation
Oral ulcers	Oral and nasopharyngeal ulcers, usually painless, observed by physician
Arthritis	Nonerosive arthritis involving two or more peripheral joints, characterized by tenderness, swelling, or effusion
Serositis	Pleuritis (convincing history of pleuritic pain or rub heard by physician or evidence of pleural effusion) *or* Pericarditis (documented by electrocardiogram or rub or evidence of pericardial effusion)
Renal disorder	Persistent proteinuria > 0.5 grams per day or > than 3+ if quantification not performed *or* Cellular casts may be red cell, hemoglobin, granular, tubular, or mixed
Neurological disorder	Seizures—in the absence of offending drugs or known metabolic derangements, *e.g.,* uremia, ketoacidosis, or electrolyte imbalance *or* Psychosis—in the absence of offending drugs or known metabolic derangements, *e.g.,* uremia, ketoacidosis, or electrolyte imbalance
Hematological disorder	Hemolytic anemia—with reticulocytosis *or* Leukopenia (<4000/mm^3 total on two or more occasions) *or* Lymphopenia (<1500/mm^3 on two or more occasions) *or* Thrombocytopenia (<100 000/mm^3 in the absence of offending drugs)
Immunological disorder	Anti-ds DNA: antibody to native DNA in abnormal titer *or* Anti-Sm: presence of antibody to Sm nuclear antigen *or* Positive finding of antiphospholipid antibodies based on (1) an abnormal serum level of immunoglobulin G (IgG) or IgM anticardiolipin antibodies; (2) a positive test for lupus coagulant using a standard method; or (3) a false-positive serological test for syphilis known to be positive for at least 6 months and confirmed by *Treponema pallidum* immobilization or fluorescent treponemal antibody absorption test
Antinuclear antibodies	An abnormal titer of antinuclear antibody by immunofluorescence or an equivalent assay at any point in time and in the absence of drugs known to be associated with "drug-induced lupus" syndrome

be greater in lupus patients compared with their healthy peers even after controlling for traditional risk factors, suggesting that lupus *per se* confers a risk for atherosclerotic events.

IMMUNOPATHOGENESIS

The immune system is designed to protect the host against foreign pathogens and to remove cellular debris without damaging self. With their universal production of autoantibodies and characteristic pathological findings of inflammation, vasculitis, vasculopathy, and immune complex deposition, SLE patients display a failure to maintain immune tolerance and immune homeostasis. The heterogeneity of disease manifestations reflects the multiplicity of genetic, hormonal, and immune abnormalities and the diversity of environmental triggers or modifiers contributing to clinical disease (Table 51.3). Progression from initial autoreactivity to clinical disease occurs over time (Fig. 51.1), with the first detectable immune abnormalities of either the presence of serum autoantibodies or the increased expression of interferon (IFN)-inducible gene in blood cells (the IFN signature).

Autoantibodies

ANAs are present in over 98% of patients diagnosed with SLE. Their presence is not specific to SLE, as they are observed in patients with other autoimmune diseases, malignancies, and viral (hepatitis) and parasitic (malaria) infections as well as in response to environmental triggers such as therapeutic agents (see section on drug-induced lupus, below). Furthermore, ANAs are found in low titer in 5% of the general population, with prevalence increasing with age. Common ANA specificities found in lupus patients include double-stranded (ds) DNA, single-stranded (ss) DNA, extractable nuclear antigens (Sm, RNP, Ro, and La), histones, and chromatin. Specific ANAs are associated with disease subsets such as anti-Ro antibodies with subacute cutaneous and neonatal SLE, anti-dsDNA, and anti-C1q antibodies with renal disease.[6] Most autoantibody titers do not correlate with disease activity; anti-ds DNA antibodies are a notable exception. Their fluctuation with disease activity suggests a pathogenic role for this autoantibody, and monitoring their titer helps to predict impending disease flare in some patients.

The Predisposed Host: Genetic Contributions

SLE is a multigenic disease. Most disease-associated alleles are present in healthy individuals. Only when multiple alleles are present, along with an appropriate environmental trigger, will a lupus-like phenotype arise. Familial disease clustering and a higher disease concordance in monozygotic than dizygotic twins suggest both an underlying genetic susceptibility and the importance of environmental or epigenetic factors. Cell types with highest expression of genes for which there are SLE susceptibility alleles are B cells and dendritic cells (DCs), suggesting that these cell lineages may be drivers of disease. This is an intriguing hypothesis, as it suggests that alterations in B-cell tolerance mechanisms and alterations in antigen presentation to T- effector and T-regulatory cells and altered cytokine production are central to disease pathogenesis.

Presumed susceptibility genes identified in humans include those affecting lymphocyte activation, proliferation and apoptosis,

TABLE 51.2 **Prevalence (%) of American College of Rheumatology Criteria in Different Ethnic Cohorts**

	†Hispanic PROFILE cohort[a] (*n* = 78)	†Hispanic Puerto Rican[b] (*n* = 134)	‡Caucasian Spain[c] (*n* = 239)	‡Caucasian USA[d] (*n* = 46)	†Caucasian PROFILE cohort[a] (*n* = 260)	†Caucasian Norwegian[e] (*n* = 346)	†Caucasian Danish[f] (*n* = 513)	†African American PROFILE cohort[a] (*n* = 216)	‡Chinese[d] (*n* = 175)
Malar rash	64	72	NA	24	67	40	48	45	58
Discoid rash	6	10	27	24	12	13	14	33	6
Photosensitivity	59	77	29	46	72	41	43	46	31
Oral/nasal ulcers	58	30	18	7	57	1	11	46	15
Arthritis	91	67	71	54	87	83	67	89	54
Seizure/psychosis	12	9	6	4	9	8	13	16	9
Renal	59	30	23	54	23	17	45	54	29
Serositis	64	28	33	26	42	34	39	60	11
Cytopenias	85	77	55	83	62	36	67	82	58
Antinuclear antibody	97	93	100	83	97	99	98	97	95
Immunological	83	NA	NA	57	65	57	98	79	81

†Cumulative data.

‡Inception data.

[a] Alarcón GS, McGwin Jr G, et al. Baseline characteristics of a multiethnic lupus cohort: PROFILE. Lupus 2002;11(2):95–101.

[b] Vilá LM, Mayor AM, et al. Clinical and immunological manifestations in 134 Puerto Rican patients with systemic lupus erythematosus. Lupus 1999;8(4):279–86.

[c] Buján S, Ordi-Ros J, et al. Contribution of the initial features of systemic lupus erythematosus to the clinical evolution and survival of a cohort of Mediterranean patients. Ann Rheum Dis 2003;62(9):859–65.

[d] Thumboo J, Uramoto K, et al. A comparative study of the clinical manifestations of systemic lupus erythematosus in Caucasians in Rochester, Minnesota, and Chinese in Singapore, from 1980 to 1992. Arthritis Rheum 2001;45(6):494–500.

[e] Gilboe IM, Husby G. Application of the 1982 revised criteria for the classification of systemic lupus erythematosus on a cohort of 346 Norwegian patients with connective tissue disease. Scand J Rheumatol 1999;28(2):81–7.

[f] Jacobsen S, Petersen J, et al. A multicentre study of 513 Danish patients with systemic lupus erythematosus. I. Disease manifestations and analyses of clinical subsets. Clin Rheumatol 1998;17(6):468–77.

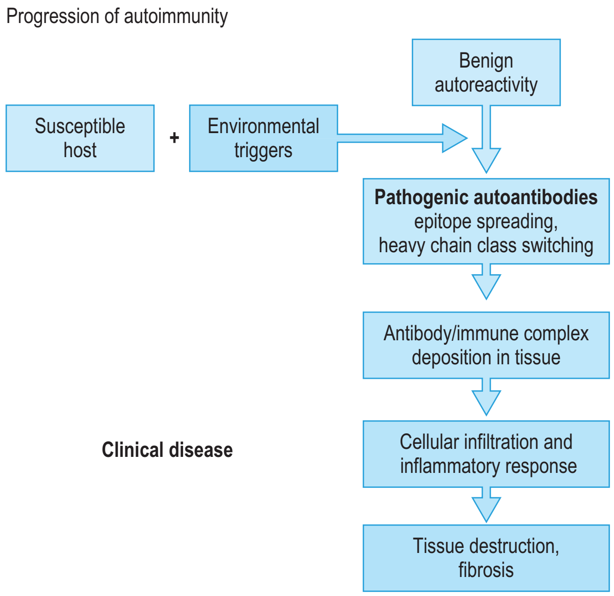

Progression of autoimmunity
Susceptible host
+
Environmental triggers
Benign autoreactivity
Pathogenic autoantibodies
epitope spreading,
heavy chain class switching
Antibody/immune complex deposition in tissue
Clinical disease
Cellular infiltration and inflammatory response
Tissue destruction, fibrosis

FIG 51.1 Spectrum of Autoimmunity.

TABLE 51.3 Factors Contributing to Autoimmunity

- Genetic factors
- Loss of peripheral tolerance
  - B-cell abnormalities
  - T-cell abnormalities
  - Dendritic cell abnormalities
- Cytokine milieu
- Hormonal influences
- Environmental triggers

cytokine production, antigen presentation, and clearance of apoptotic debris. Many of these genes are also implicated in susceptibility to other autoimmune diseases (*e.g., CTLA4* in Graves disease and type 1 diabetes, *PTPN22* polymorphisms in rheumatoid arthritis and type 1 diabetes).[7,8] Genome-wide association studies (GWASs) and genome-wide linkage analyses have utilized high-throughput techniques to study hundreds of thousands of single nucleotide polymorphisms (SNPs) in individual patients with SLE and have yielded approximately 50 SLE-associated susceptibility loci,[9] some of which are of particular relevance in ethnic and racial groups (Chapter 33).

Genes Associated With Antigen Presentation

Polymorphisms of major histocompatibility complex (MHC) genes determine the peptides of self and foreign antigens presented within MHC molecules that select the naïve T-cell repertoire. Human leukocyte antigen (HLA)-DR2 haplotypes in African American, African, Taiwanese, and Korean populations and HLA-DR3 haplotypes in Caucasian populations have been associated with a two- to threefold increased risk for developing SLE.[10] Associations between anti-Ro antibodies and HLA-DR3 and anti-La antibodies and HLA-DR25 are consistent with the concept of antigen-driven processes involving T-cell recognition. The association of HLA with SLE is not as dominant as with other autoimmune disease, suggesting that the T-cell response may dictate autospecificity (as above) but is less important in determining risk of disease.

Genes Associated With Impaired Clearance of Apoptotic Debris

Patients with severe deficiencies of C2, C4, and C1q display disease risks of 10%, 75%, and 90%, respectively, for SLE.[11] Reduced uptake of apoptotic cells has been implicated in disease initiation in murine models of SLE and is seen histopathologically in lymph nodes of lupus patients.[12] Though rare, individuals homozygous for null alleles of C2, C4, and C1q are at significantly increased risk. Polymorphisms of mannose-binding lectin (MBL) and C-reactive protein (CRP)—acute-phase reactants that facilitate opsonization and phagocytosis of immune complexes, apoptotic debris, and microbes—also associate with SLE susceptibility; MBL haplotypes appear to be more important in Chinese and Spanish populations than in Caucasians. *TREX1* encodes a 3′ repair exonuclease that monitors DNA synthesis; TREX1 deficiency leads to accumulation of endogenous DNA and is associated with increased expression of IFN and autoimmunity. *ITGAM* encodes the α-chain of an integrin adhesion molecule that binds C3b fragments and other proinflammatory molecules. GWASs have identified associations of *TREX1* and *ITGAM* variants with and SLE susceptibility in European populations. *ITGAM* has also been reported as a susceptibility gene in African American and Hispanic populations but not in Asian populations.[9]

The Fcγ receptors—FcγR1 (CD 64), FcγRII (CD32), and FcγRIII (CD16)—have different binding affinities for immunoglobulin (IgG) and immune complexes as well as different cell-specific expression and function. FcγRI, FcγRIIa, and FcγRIIIa and IIIb are all activating receptors. Cross-linking these receptors by immune complexes results in degranulation, phagocytosis, antibody-dependent cellular cytotoxicity, cytokine gene transcription, and release of inflammatory mediators by myeloid cells. Substitutions of one or more amino acids in the activating FcγR genes—arginine (R) for histidine (H) at position 131 in γFcγRIIa and phenylalanine (F) for valine (V) at position 158 in γFcγRIIIa—results in decreased affinity for IgG immune complexes. Because apoptotic debris is opsonized by antibody to promote clearance, such deficiencies may lead to immune activation by apoptotic debris. Associations between the γFcγRIIa R131H allele and disease susceptibility or nephritis occur in Brazilian, Thai, Korean, German, and African American populations, and FcR polymorphisms may also predict a therapeutic response to immunobiological agents such as rituximab.[13] γFcγRIIIa F158V and γFcγRIIIb NA2/NA2 polymorphisms are reported to associate with disease susceptibility in Dutch, Korean, Thai, and Caucasian populations (reviewed in Mackay et al., 2006).[14] In contrast, FcγRIIb is an inhibitory receptor. Engagement of FcγRIIb on DCs delivers an inhibitory signal. Cross-linking of FcγRIIb and the B-cell receptor (BCR) results in decreased intracellular calcium flux with decreased B-cell activation and proliferation. The FcγRIIb I232T allele leads to the inability of the receptor to enter lipid rafts and perform its inhibitory function. This allele is associated with SLE in Asian populations,[15] and dysregulation of FcγRIIB expression on activated memory B cells has been reported in SLE patients.[14]

Genes Associated With Lymphocyte Activation, Proliferation, and Function

Several SLE risk genes have been implicated in the regulation and activation of lymphocytes. *BLK, LYN,* and *BANK1* encode the tyrosine-kinase proteins Blk and Lyn and B-cell scaffold protein with ankarin repeats; all are associated with intracellular signaling pathways. *ETS1* and *IKZF1* encode transcription factors and are believed to play a role in B-cell differentiation and self-tolerance. Cytotoxic T lymphocyte antigen-4 (CTLA-4) is upregulated on T cells after activation and dampens inflammatory responses. It has a higher affinity than CD28 for B7.1 (CD80) and B7.2 (CD86), thereby competitively inhibiting engagement of CD28 and blocking the costimulatory signal required for T-cell activation. CTLA-4 ligation of B7 also activates indoleamine dioxygenase (IDO) expression, an enzyme involved in tryptophan metabolism, and diminishes T-cell proliferation. Finally, CTLA-4 is critical for activation of regulatory T cells. In CTLA-$4^{-/-}$ mice, an uncontrolled, lethal inflammatory response occurs. CTLA-4 alleles with decreased production of soluble CTLA-4 are implicated in the pathogenesis of several autoimmune diseases, including Sjögren disease, ulcerative colitis, psoriasis, type 1 diabetes, and multiple sclerosis as well as SLE.[16]

The *PTPN22* gene encodes a tyrosine phosphatase that has been suggested to downregulate T-cell and B-cell receptor (BCR) activation, although the precise role of polymorphic PTPN22 I disease pathogenesis remains unresolved. *A PTPN22* polymorphism resulting in diminished ability to control T-cell antigen receptor activation has been reported in SLE and other autoimmune diseases, suggesting a common mechanism of immune dysregulation. The PTPN22 risk allele has also been associated with reduced type I IFN production by myeloid cells after activation of Toll-like receptor (TLR) 7.

Genes Encoding Cytokines and Chemokines

IFN regulatory factor 5 gene (*IRF5*) encodes a critical transcription factor in the type I IFN pathway; the *IRF5* locus has the strongest association with SLE outside of the MHC region. Four allelic variants have been identified in multiple ethnically diverse populations. Variants of the signal transducer and activator of transcription factor 4 protein (*STAT4*) gene have also been associated with SLE susceptibility in GWASs of Caucasian and Asian populations. There are data suggesting interaction between IRF5 and STAT4, and the presence of one or more risk alleles of the two genes may confer increased susceptibility. IL-1 receptor-associated kinase 1 (*IRAK1*) and methyl-CPG-binding protein 2 (*MECP2*) genes are both found on the X chromosome. IRAK1 regulates multiple pathways in innate and adaptive immune responses, including the link between immune complexes, TLR signaling, T-cell receptor (TCR) signaling, and IFN production. Monocyte chemoattractant protein (MCP-1) is a potent chemoattractant for monocytes, memory T cells, and natural killer T cells. MCP-1 expression is upregulated in renal tubular cells, and glomeruli in lupus nephritis and urine levels of MCP-1 are increased in patients with active lupus nephritis. An MCP-1 polymorphism resulting in increased MCP-1 production has been associated with SLE nephritis.[17] Polymorphisms in the tyrosine kinase-2 (*TYK2*) gene are associated with increased expression of type I IFNs (IFN-α, IFN-β) in SLE. IFN-α-regulated genes are highly expressed in peripheral blood cells from SLE patients compared with healthy controls (the IFN signature).[18] IFN-α mediates maturation of DCs and monocytes, increasing the capacity for T-cell activation, and promotes B-cell differentiation and Ig class switching. However, in two murine models of lupus, decreases in type I IFNs unexpectedly led to worsening disease, consistent with numerous studies showing that type I IFNs have both pro- and antiinflammatory effects on innate and adaptive immune responses and suggesting that the effect of IFN-α on autoimmunity is complex.

Multiple polymorphisms in the *IL10* gene have been reported, with conflicting results with respect to SLE susceptibility. A meta-analysis of 15 studies concluded that some *IL10* polymorphisms do associate with SLE, but their importance is modulated by ethnic background.[19]

In the NZB/W murine model, a tumor necrosis factor (TNF) allele associated with low production is linked with disease, and treatment with TNF decreases autoantibody production. Consistent with this observation, TNF blockade for rheumatoid arthritis or inflammatory bowel disease can lead to autoantibody production and infrequently to frank lupus. Several polymorphisms for genes encoding TNF-α and -β (lymphotoxin-α) have been associated with SLE; these associations are also influenced by ethnicity.[20]

Genes Associated With Cell Survival

Fas ligand (expressed on activated T cells) binding to Fas (CD95) stimulates a signaling pathway resulting in apoptotic death of the Fas-expressing cell. Fas-induced apoptosis of activated cells contributes to the elimination of autoreactive B and T lymphocytes. Lymphopenia in SLE has been associated with increased Fas expression on lymphocytes, and Fas and Fas ligand alleles have been linked to disease susceptibility.[21] Whereas lack of Fas expression in mice leads to an SLE-like phenotype, lack of Fas expression in humans is associated with a lymphoproliferative disease that does not share autoantigen specificities and target organs with SLE.

Bcl-2 family genes encode intracellular proteins that are either pro- or antiapoptotic. Increased expression of Bcl-2, an antiapoptotic molecule, leads to a lupus-like serology and nephritis in mice with certain genetic backgrounds. Increased intracellular levels of Bcl-2 have been reported in SLE, particularly in Chinese and Mexican populations. The combination of a Bcl-2 susceptibility allele and interleukin-10 (IL-10) susceptibility allele confers a 40-fold increased risk of SLE, demonstrating that *infelicitous* combinations of risk alleles potentiate risk.[22]

Genes Regulating Target Organ Damage

Only a few genes are known to regulate the vulnerability of target organs, the kidney in particular, to autoimmune attack. In mice, genes encoding kallikreins, which upregulate bradykinins, have SLE susceptibility alleles. In human studies a risk allele for ABIN1, which regulates nuclear factor (NF)-κB activation, can lead to greater kidney disease. Two risk alleles of APDOL1 are particularly common in the African American population and are thought to contribute to the increased severity of renal disease in this population.

Epigenetic Contributions

Epigenetic regulation plays a determining role in gene activation. Major epigenetic influences in SLE involve DNA methylation at cytosine-guanine nucleotides (CpG methylation) and histone posttranslational modifications (lysine acetylation or methylation, phosphorylation of serine or threonine, arginine methylation).[23] SLE susceptibility and autoantibody production is associated with DNA hypomethylation.[24] Several drugs known to induce a lupus-like disease (procainamide, hydralazine) also cause decreased DNA methylation. A landmark study of high-throughput analysis of DNA methylation in discordant twins demonstrated greater DNA hypomethylation in the affected siblings.[25] It is not surprising that genes encoding integrins, NGAL, CD40 ligand, IFN-γ receptor, and IL-6 are among the hypomethylated genes in SLE. Mechanisms for hypomethylation remain unclear; decreased efficacy of DNA methyltransferases (DNMTs) and overexpression of microRNAs (miRNA) that interfere with DNMT activity have been proposed.[26]

Vitamin D has been shown to contribute to the regulation of the epigenome. Vitamin D levels are low in many lupus patients, but the relationship of vitamin D levels to disease risk or disease severity remains controversial.[27]

Inflammation itself may alter the epigenome; some studies suggest that the metabolic modulator metformin can reverse some of the inflammation-induced alterations. Recent studies show that metformin can reduce murine lupus, although the mechanism for this therapeutic effect is not completely clear.

B Cells

B-Cell Selection

The process of immunoglobulin variable region gene rearrangement produces large numbers of self-reactive B cells. Most B cells displaying self-reactive immunoglobulin are deleted centrally in the bone marrow and at subsequent checkpoints in the periphery (Fig. 51.2), so the frequency of autoreactive cells decreases from ~75% in immature B cells in the bone marrow to ~20% in the mature naïve B-cell population in healthy individuals. Many of the autoreactive B cells in the naïve population are normally suppressed by anergy induction. Several of these checkpoints appear deficient in SLE. The IgM cross-reactive autoantibodies made by these remaining autoreactive mature naïve B cells are thought to facilitate clearance of apoptotic cells, decreasing the development of potentially pathogenic T- and B-cell responses to self-antigens. When IgM autoreactive antibodies are missing in mice, a lupus-like phenotype develops.

Another important peripheral checkpoint is entry into the T-cell-dependent, long-lived memory compartment. B cells with the 9G4 idiotype express antibodies encoded by the *VH4-34* gene, reactive with *N*-acetyllactosamine (NAL) determinants of glycoproteins on blood group antigens targeted by cold agglutinins, gangliosides, gastrointestinal (GI) mucins, glycolipids, and CD45 on B lymphocytes.[28] 9G4 B cells are reported to be present in 5–10% of the naïve B-cell population in healthy donors as well as in the IgM memory compartment. However, 9G4 B cells are excluded from the T-cell–dependent IgG memory and plasma cell populations, suggesting that these autoreactive cells fail to cross a developmental checkpoint after activation in normal individuals. Evaluation of tonsillar biopsies and spleens from healthy donors shows that the frequency of germinal centers with $9G4^+$ cells is less than 1%, implying that negative selection of autoreactive cells occurs at the transition of naïve to germinal center B cells. In contrast, tonsillar biopsies from SLE patients demonstrate that 15–20% of germinal centers are positive for autoreactive 9G4 B cells. This study, as well as others, suggests that checkpoints exist for entry into and exodus from the germinal center response and that these may be altered in SLE.

For example, it is clear that some pathogenic autoantibodies are not derived from natural autoantibodies. Back mutation of several anti-DNA antibodies to their germline-encoded precursors has identified nonautoreactive precursors. The failure of censoring mechanisms in germinal centers to prevent the maturation of autoreactive cells may reflect intrinsic B-cell abnormalities or abnormalities in costimulatory molecules, cytokines (such as B cell–activating factor [BAFF]; see below), follicular DCs, or T-cell, B-cell interactions.

B-Cell Signaling

Hyperactive B-cell responses to immunological stimulation are implicated in the production of pathogenic antibodies. SLE B cells have increased intracellular Ca^{2+} flux in response to BCR signaling,[29] partially due to FcγRIIb dysfunction (Ile 232 Thr).[15]

The intracellular protein tyrosine kinase Lyn has both positive and negative effects on BCR signaling. Decreased expression of Lyn results in increased intracellular Ca^{2+} flux and B-cell hyperactivity. Correspondingly, Lyn expression is decreased in resting and activated B cells in one-half to two-thirds of SLE patients, suggesting that Lyn may modulate negative regulation of BCR signaling in human disease. In contrast, a Lyn mutation in mice leading to increased expression results in autoantibody production and severe glomerulonephritis, suggesting that balanced Lyn expression is critical to tolerance pathways.

B-Cell Rescue

B lymphocyte stimulator (BAFF; also known as BLyS) is a member of the TNF family and participates in B-cell maturation and survival. BAFF enhances survival of B cells through engagement of several receptors–BCMA, BAFF-R, and TACI. High levels of this cytokine allow survival of autoreactive B cells in mice, resulting in a lupus-like disease (reviewed in Liu and Davidson).[30] Evidence supporting a role for BAFF in autoreactive B-cell rescue and human SLE includes the elevated levels seen in lupus patients, associations of BAFF levels with autoantibody titers, and, in some reports, correlations with disease activity. Recently belimumab, a monoclonal antibody directed against soluble BAFF,

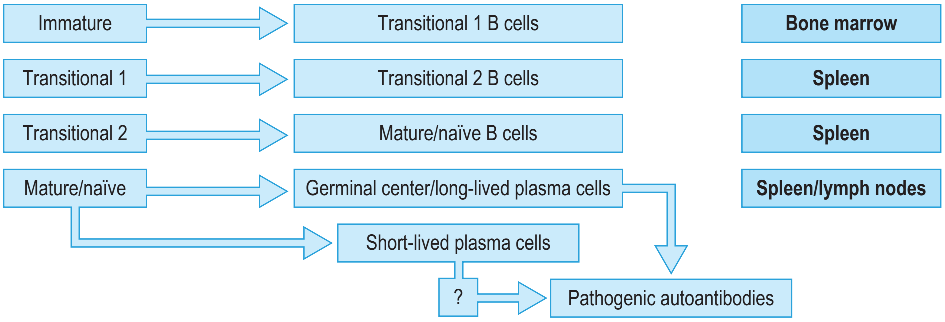


FIG 51.2 Autoreactive B-Cell Checkpoints. There are tolerance checkpoints at every stage of B-cell activation and maturation. How many checkpoints need to be breached to achieve a pathogenic state and clinical disease is not known.

has been shown to increase the percentage of ANA-positive B cells that are anergic in SLE patients.

Interactions between CD40 (B cell) and CD40 ligand (CD40L, T cell) are essential for B-cell proliferation, differentiation of memory cells into plasma cells, and germinal center formation. Immature autoreactive B cells can be rescued from antigen-induced apoptosis by engagement of CD40 or by IL-4.[31] SLE T and B cells have upregulated CD40L, providing a critical molecule to mediate B-cell rescue. The combination of IL-17 and BAFF facilitates B-lymphocyte proliferation and maturation. This combination can serve as an alternative stimulatory signal for B-cell activation and can replace CD40/CD40L interactions. Therefore in a permissible and proinflammatory cytokine milieu, B-cell activation and autoantibody production may occur in the absence of cognate T-cell help.[32] These studies are consistent with the increased numbers of plasmablasts present in blood of lupus patients with active diseases and with the increased numbers of class-switched CD27-negative B cells in blood of lupus patients.

B-Cell Pathogenicity Unrelated to Antibody Production

An essential role for B cells in SLE that is independent of antibody production is demonstrated by the T-cell activation, kidney inflammation, and increased mortality in MRL/lpr lupus-prone mice that lack the ability to secrete antibody and the absence of such cells in mice that lack B cells. B cells are efficient antigen-presenting cells; B cells with self-reactive specificity that have escaped tolerance are likely to present self-antigen to autoreactive T cells. This may explain, in part, the beneficial effects of treatment with rituximab, a B-cell–depleting drug, in SLE as autoantibody levels remain unaffected.

Neutrophils

An important link exists between neutrophils and autoimmunity. Neutrophil extracellular traps (NETS) are chromatin filaments released from neutrophils to trap microbes. In addition to microbial peptides, NETS also contain neutrophil peptides, including neutrophil-encoded antimicrobial peptide LL37 (cathelicidin). Circulating DNA-containing immune complexes from lupus patients have been shown to contain peptide LL37. These combined DNA/LL37 immune complexes are potent TLR9 stimulants, resulting in production of IFN-α by plasmacytoid DC (pDC).[33] Moreover, anti-RNP antibodies also induce "NETosis" in a process that is dependent on TLR7 and FcγRIIa signaling.[34] Thus it is possible that autoantibodies accelerate NET formation and increase formation of DNA/LL37 immune complexes and pDC production of IFN-α. These studies confirm earlier work that described a "granulocyte signature" with significant upregulation of granulocyte-specific transcripts within peripheral blood mononuclear cells in pediatric SLE patients.[35]

Dendritic Cells

DCs recognize pathogens through membrane pattern recognition receptors (PRRs) and are a critical component of the immune system connecting innate and adaptive immune responses.[36] They can be tolerogenic or highly inflammatory due to their high expression of costimulatory molecules. DCs function normally as surveillance cells, phagocytizing cellular debris and determining whether there is cause for alarm. PRRs on DCs bind pathogen-associated molecular patterns (PAMPs) and damage-associated molecular patterns (DAMPs) that are present in sterile inflammation. After internalization of the PAMPs or DAMPs, endosomal TLRs are activated by nucleic acids and other ligands. Enhanced IFN-α in SLE pathogenesis is now well described; the primary cells responsible for IFN-α production are pDC. TLR7 or TLR9 activation with RNA and DNA, respectively, induces immature DC to differentiate to immunocompetent, IFN-α-producing DC that have wide-reaching effects on T cells, B cells, neutrophils, and monocytes. Of note, type I IFN increases plasmablast differentiation favoring short-lived plasma cells over germinal center matured long-lived plasma cells. Plasmacytoid dendritic cells (pDC) have been demonstrated in skin and renal lesions, and upregulation of the "IFN signature" is observed in many, but not all, SLE patients and has been associated with disease activity. A number of susceptibility genes identified through GWASs are associated with IFN pathways in SLE including *IRF5, IRF7,* and *STAT4.*

T Cells

T cells are critical in the abrogation of self-tolerance by providing help to autoreactive B cells and facilitating the production of somatically mutated, high-affinity, pathogenic autoantibodies. Lupus patients can exhibit T-cell phenotypes consisting of increased numbers of $CD3^+CD4^-CD8^-$ T cells, increased Th17 cells, T-follicular helper (Tfh) cells, and decreased numbers or function of T regulatory cells (Tregs).[37] Additionally, lupus T cells display increased expression of activation markers and abnormal TCR signaling responses. A substitution of TCR ζ-chain by the γ-chain of the Fc receptor results in increased intracellular influx of Ca^{2+} after TCR stimulation and a concomitant decrease in IL-2 production, disfavoring the induction of Tregs.

Self-tolerance is maintained in part by the suppressive actions of Treg. "Natural" Tregs arise *de novo* in the thymus, whereas "induced" Tregs evolve from naïve T cells exposed to IL-2 and transforming growth factor (TGF)-β in the periphery. These cells are characterized by high surface expression of the IL-2 receptor α-chain, CD25, and high levels of FOXP3 intracellularly. Tregs act together with tolerogenic DC to maintain a steady state of immature DC. In contrast, effector T cells can secrete IFN-γ and IL-17 that promote immature DC differentiation to immunogenic DC capable of secreting IL-1, IL-6, IL-12, and TNF-α and can activate autoreactive T cells, thereby establishing a feedback loop with impaired Tregs and activation of autoreactive T cells.[38] Studies demonstrate alterations in Tregs in patients with SLE. Some have demonstrated reduced numbers and altered function of peripheral $CD4^+CD25^+$ $FOXP3^+$ cells in patients with active disease, whereas others fail to identify a defect.[39]

HORMONAL INFLUENCES

The most compelling evidence for the role of sex hormones in SLE is the observation that lupus preferentially affects women of childbearing age. The female-to-male ratio is 2:1 before menarche, 8–9:1 in the fourth decade, and 2:1 after menopause. Numerous case reports and studies of disease flares correlating with pregnancy, menstruation, and use of oral contraceptives containing high doses of estrogen suggest a role for estrogen in disease activity. A significant correlation between plasma levels of estradiol and clinical disease activity and increased α-hydroxylation of estrogen in SLE yielding the more active metabolite 16α-hydroxyestrone are also reported.[40] No significant differences in levels of sex hormones (including estrogen, testosterone, prolactin) are noted in male SLE patients, suggesting that the development of SLE in females may be more closely related to sex hormones than in men. Randomized controlled studies of estrogen in SLE suggest that the use of exogenous estrogen in patients with stable disease may be safe; however, a subset of patients appears to have an estrogen-sensitive disease. Mild to moderate flare rates were significantly increased in postmenopausal women treated with hormone replacement therapy.

Most of what we understand about hormonal modulation of B-cell development comes from mouse studies.[31] Estrogen treatment of NZB/W and MRL/*lpr* mice or castration of male lupus-prone mice exacerbates disease, whereas oophorectomy of female mice ameliorates disease. Treatment of lupus-prone mice with the selective estrogen receptor modulator tamoxifen also ameliorates disease.

Estrogen and prolactin promote the loss of B-cell tolerance in nonautoimmune mice. Estrogen results in diminished B-cell responsiveness to BCR cross-linking and in less stringent negative selection. Additionally, estrogen has also been reported to inhibit activation-induced T-cell death by downregulation of Fas ligand expression, thereby permitting increased numbers of autoreactive T cells.[41] Emerging data suggest that estrogen receptor signaling promotes differentiation of DC into immune-competent DC through increased expression of the transcription factor IRF4.[42]

Elevated prolactin levels are reported in 20% of patients with SLE, and increased prolactin exposure in lupus-prone mice exacerbates disease activity. Whereas treatment of patients with bromocriptine yielded equivocal results, treatment of NZB/W lupus mice with bromocriptine results in improved survival. Transmembrane prolactin receptors are present on a variety of cells, including T and B cells. Upregulation of both Bcl-2 and CD40 on B cells and CD40L on T cells occurs in response to prolactin, identifying pathways that may be involved in the prolactin-mediated rescue of autoreactive B cells.

Data on the microbiome demonstrate that sex hormones can modulate its composition. Studies in murine lupus suggest that androgens generate a microbiome that prevents the development of lupus, whereas estrogen maintains a disease-permissive microbiome. Studies of the microbiome in human disease are just beginning.[43]

CLINICAL MANIFESTATIONS

CLINICAL PEARLS

- Systemic lupus erythematosus (SLE) is a systemic disease with the potential to affect any organ system.
- Not all symptoms experienced by a patient with SLE are from SLE disease activity. Attribution is critical to determine before initiation of treatment. The differential diagnosis of a lupus flare mandates consideration of infection, drug toxicities, or other etiologies.
- Corticosteroid exposure should be minimized.
- In the absence of data from randomized trials, use of aggressive treatment must be balanced against associated toxicity.
- SLE patients accumulate damage from both repeated episodes of inflammatory disease and medication toxicities. Prompt recognition and appropriate treatment of disease flares should result in reduced exposure to corticosteroids and immunosuppressive agents.
- SLE patients are at increased risk of developing atherosclerotic disease, osteoporosis, malignancy, diabetes mellitus, and hypertension. It is essential to screen for and reduce modifiable risk factors.

The prevalence of the diverse clinical and laboratory features of lupus varies in published reports (Table 51.2). Both genetic and environmental factors are likely to account for much of the variability between cohorts. Characteristics of published cohorts, including inception versus long-standing lupus cohorts, community versus academic settings, socioeconomic factors, and ascertainment differences are also likely to contribute to observed differences in the prevalence of clinical manifestations. Over time, the course of lupus is characterized by flares and remissions of disease activity.

KEY CONCEPTS

- Continued heightened awareness of systemic lupus erythematosus to shorten the time between onset of symptoms and diagnosis will improve outcomes.
- Lupus is a disease characterized by recurrent flares.
- Attentive monitoring even during periods of disease remission leads to early recognition of impending flare, better control, and better prognosis.
- Lupus is a chronic disease; the importance of a therapeutic partnership between physicians and patients, emotional/social support, and patient education cannot be overemphasized.
- Advances in understanding mechanisms of disease pathogenesis correspond to advances in treatment strategies and development of therapies with improved efficacy and safety profiles.

The most common features of lupus are constitutional and include fatigue, malaise, low-grade fever, anorexia, and lymphadenopathy. These symptoms are believed to result from elevated serum levels of inflammatory cytokines and may accompany other organ system manifestations of active disease, or they may occur in isolation. Although frequent, these symptoms are nonspecific and do not aid in making the diagnosis of SLE. Symptoms of fatigue and malaise may also represent fibromyalgia, which can co-occur with SLE or confound the diagnosis.[44]

Musculoskeletal Involvement

The musculoskeletal system is the most common organ system affected in SLE; joint pain is the presenting symptom in approximately 60–70% of patients, and 85% have joint involvement after 5 years.[45]

Arthritis and Arthralgia

The pattern of joint involvement is usually symmetrical, affecting the small joints of the hands, wrists, and knees. Pain in the ankles, elbows, shoulders, or hips or monoarticular involvement is less typical. In contrast to rheumatoid arthritis, morning stiffness is typically limited to several minutes. Frequently, the subjective complaints of pain are greater than the objective findings of warmth, swelling, and erythema. Lupus arthritis is characteristically nonerosive on X-ray and nondeforming. Anticyclic citrullinated peptide antibodies occur frequently in the rare patients with erosive arthritis. Some lupus patients develop a nonerosive hand deformity with hypermobile joints secondary to tendon and ligamentous laxity (Jaccoud arthritis) (Fig. 51.3). Proliferative tenosynovitis, synovitis, small erosions not detectable on plain radiographs, and capsular swelling are features of joint and soft-tissue involvement that may be seen on MRI. The role of ultrasound and MRI in the evaluation of musculoskeletal symptoms is not established.

Joint effusions, when they occur, are usually small. The fluid is clear yellow with normal viscosity and forms a mucin clot. It is typically noninflammatory with a normal glucose level and a white blood cell (WBC) count of less than 2000 cells/mL that is predominantly lymphocytic. ANA performed on the synovial

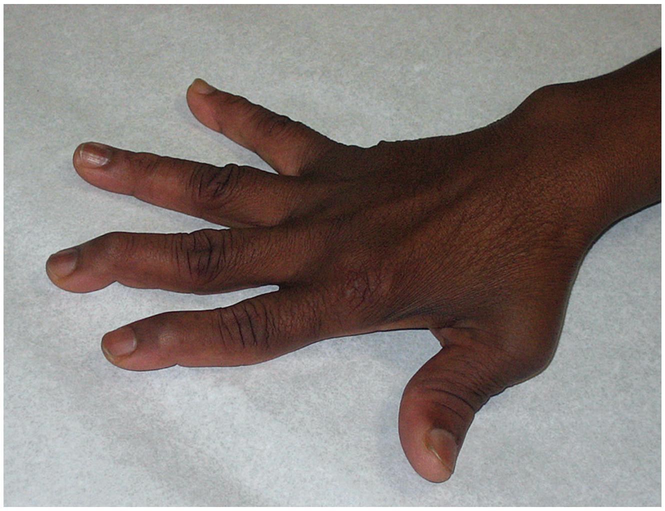


FIG 51.3 Jaccoud Arthritis in Systemic Lupus Erythematosus.

fluid may be positive, and lupus erythematosus (LE) cells may be present. Synovial fluid complement levels can be normal or depressed. Synovial histology in lupus is not specific and shows synovial hyperplasia with fibrin deposition and microvascular changes that include perivascular infiltrates in the majority of cases.

Tendinitis

Tendinitis is not usually attributed to SLE unless associated with tendon rupture. When present, it is usually located in the Achilles tendon or the tendons around the knee. Tendon ruptures are more common in males and have been associated with trauma, steroid use, long disease duration, and Jaccoud arthropathy.[46] Biopsy shows a mononuclear infiltrate with tendon degeneration and neovascularization. The diagnosis can be easily demonstrated on ultrasound or MRI.

Myositis/Myalgia

Generalized myalgia is extremely common in lupus. It frequently affects the deltoids and quadriceps and occurs during flares of active disease. Muscle disease secondary to treatment with corticosteroids, statins, and antimalarials or in association with hypothyroidism is also frequent and must be considered in the evaluation of a lupus patient with myalgia. Inflammatory muscle disease with weakness and an elevated creatine phosphokinase is less common, occurring in approximately 10% of lupus patients.[44] Electromyography can be normal or can be characteristic of the myositis observed in polymyositis or dermatomyositis. Muscle biopsy can also be normal or can show changes associated with dermatomyositis such as a perivascular or perifascicular infiltrate and immunoglobulin and complement deposition. Muscle atrophy, fiber necrosis, microtubular inclusions, and/or a mononuclear infiltrate have been documented. MRI findings are nonspecific.

Avascular Necrosis

Avascular necrosis (AVN) has been reported in up to 30% of lupus patients, is frequently asymptomatic, and is detected by MRI.[47] The most commonly affected site is the femoral head. Groin pain exacerbated with weight bearing is a common complaint. In addition to the hip, AVN can involve the knees, shoulders, and wrists. The majority of AVN is associated with previous administration of high doses of corticosteroids (>30 mg/day); associations with vitamin D deficiency, minority ethnicity, hypertension, and renal disease have also been reported. Bone biopsy in lupus patients affected by AVN does not reveal unique findings.

Mucocutaneous Manifestations

Skin

The skin is commonly affected in SLE, with a wide variety of lesions from malar erythema to severe bullous lupus and scarring discoid lesions (DLE). SLE-associated cutaneous lesions are generally categorized as acute cutaneous lupus erythematosus, subacute cutaneous lupus erythematosus (SCLE), and chronic cutaneous lupus erythematosus (CCLE). Known triggers for cutaneous SLE include ultraviolet (UV) light, infections, and drug reactions. Many lupus rashes arise in sun-exposed areas, and sun exposure can precipitate flares of systemic disease. Typically, a photosensitive rash erupts within hours of sun exposure and consists of tiny pruritic plaques and vesicles lasting several days. UV light induces DNA strand breaks in keratinocytes, resulting in apoptotic cell death and providing a rich source of autoantigen (*e.g.,* Ro52 antigen). SLE patients have increased numbers of apoptotic keratinocytes after exposure to UV light,[48] and these apoptotic cells have been identified in the basal layer of CCLE lesions. Particularly in a setting of impaired ability to clear apoptotic debris, the abundance of autoantigen provides stimuli for autoreactive T and B cells and recruitment of pDC with ensuing production of proinflammatory cytokines (IL-1, IFN-α, TNF-α, IL-6, IFN-γ). UV-oxidized DNA is resistant to cytosolic enzymatic degradation, thereby potentiating pDC expression of IFN-α. In addition to the proinflammatory cytokines, cutaneous lupus lesions demonstrate increased expression of IFN-α-inducible chemokines CXCL9, CXCL10, and CXCL11, which attract lymphocytes.[49] Transcript analysis of lesional skin in CCLE demonstrates a paucity of Tregs with increased numbers of T-helper 1 (Th1), IFN-γ producing cells. Increased NETS extruded from dying neutrophils are also seen in cutaneous lupus lesions. The DNA in NETS is resistant to degradation and may contribute to ongoing pDC activation through TLRs and IFN-α expression characteristic of cutaneous lesions.

Acute Cutaneous SLE

The malar rash typically occurs across the cheeks and nose but can include the forehead and chin, sparing the nasolabial folds (unlike seborrheic dermatitis) (Fig. 51.4). It usually begins as small discrete erythematous macules or papules that coalesce, is frequently associated with sun exposure, and heals without scarring. Some patients additionally have facial swelling. The differential diagnosis includes acne rosacea, seborrheic dermatitis, erysipelas, dermatomyositis, and contact dermatitis. Microscopical analysis reveals a sparse inflammatory lymphocytic dermatitis with occasional histiocytes engulfing nuclear debris resembling LE cells found close to the dermoepidermal junction. Immunofluorescent staining for complement components and immunoglobulin at the dermoepidermal junction is positive in 70–80% of patients.

Subacute Cutaneous SLE

SCLE is characterized by recurrent, nonscarring skin lesions. This distinctive rash consists of erythematous papules and plaques, with or without adherent pityriasiform scale, that erupt on the

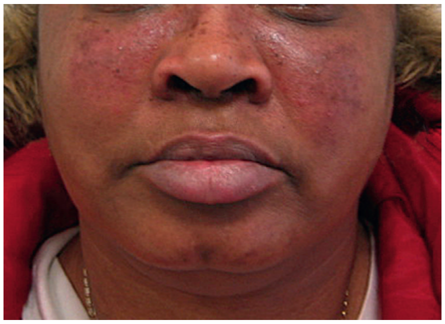


FIG 51.4 Malar Rash in a Systemic Lupus Erythematosus Patient.

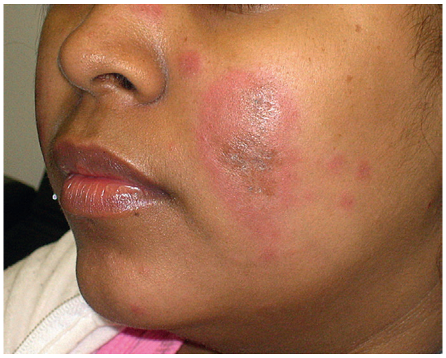


FIG 51.5 Discoid Lesion in Systemic Lupus Erythematosus.

extremities and trunk, usually sparing the head and neck. These lesions may assume an annular polycyclic form with central pallor and tiny vesicles at the active margins and can be mistaken for erythema multiforme. The differential diagnosis also includes psoriasis, polymorphic light eruption, and tinea corporis. SCLE is exacerbated by UV light and a growing list of medications, including thiazides and calcium channel blockers. Biopsy reveals a lymphocytic dermatitis confined to the superficial and mid dermis, frequently with associated dermal edema, mucinosis, and degenerating keratinocytes. Both TNF-α and IL-6 have been demonstrated in active SCLE lesions. SCLE is most commonly seen in Caucasian populations. Genetic analyses have revealed associations with HLA-A1, HLA-B8, and HLA-DR3 haplotypes as well as with deficiencies of C2, C4, and C1q. From 60–90% of SCLE patients have circulating anti-Ro antibodies; however, they are also deposited in nonlesional skin.

Chronic Cutaneous SLE

DLE are usually localized to the head and neck in photo-exposed areas, with a predilection for the ears and periorbital areas (Fig. 51.5). The lesions vary in size and result in scar tissue with significant disfigurement. Early lesions appear as erythematous plaques with or without follicular hyperkeratosis, plugging, and scale and progress to scarring annular lesions with an erythematous, indurated border, adherent scale, and a central area with atrophy and telangiectasias. There are no autoantibody associations with DLE, and only 5% of patients with DLE develop systemic lupus. High-titer ANA, Raynaud phenomenon, and the presence of arthralgias may identify patients at risk for systemic evolution. Histopathology characteristically reveals a lymphocytic interface dermatitis with CD4 lymphocytes and pDCs involving follicles and epidermis. There is vacuolar degeneration of the basal layer of epidermal keratinocytes and prominent keratotic follicular plugging. Dermal mucin deposition is also present, and there is usually dense granular deposition of immunoglobulin (predominantly IgG) and C3 at the dermal–epidermal junction. Patients with C2, C4, and C1q deficiencies may be predisposed to DLE, and promoter region polymorphisms leading to high IL-10 and low TNF-α are risk factors for DLE.

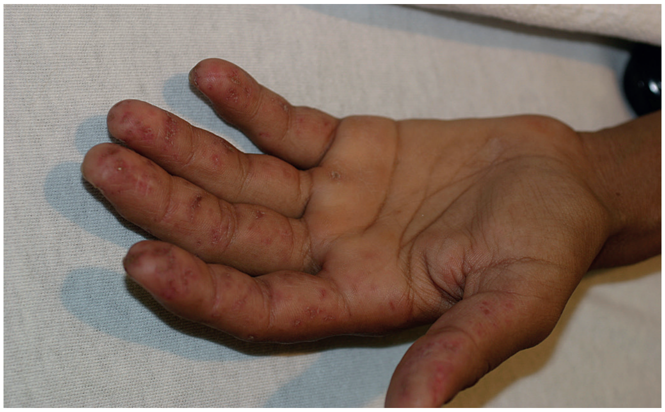


FIG 51.6 Cutaneous Vasculitis Affecting the Hands in a Patient With Active Systemic Lupus Erythematosus.

Lupus profundus typically presents as firm, tender, deep subcutaneous nodules that may atrophy over time. Overlaying epidermal changes include DLE, ulcerations, and dystrophic calcification. Biopsy reveals a lobular panniculitis with patchy lymphoplasmacytic infiltrate in subcutaneous fat lobules. Panniculitis occurs in 10–20% of patients and must be differentiated from a subcutaneous T-cell lymphoma, erythema nodosum, pancreatic panniculitis, and morphea.

Nonspecific skin lesions reported in SLE are typically seen during disease flares and are associated with greater disease severity. These lesions include, but are not limited to, entities such as cutaneous vasculitis (Fig. 51.6), urticaria, Raynaud phenomenon, livedo reticularis, alopecia, sclerodactyly, calcinosis cutis, atrophie blanche, bullous lesions, erythema multiforme, lupus tumidus, and leg ulcers. Nonspecific cutaneous lesions such as cutaneous vasculitis or ulcers are associated with a more aggressive disease course than most of the SLE-specific lesions.

The lupus band test (LBT) refers to the deposition of immunoglobulin (IgG, IgM, and/or IgA) and/or C3 along the dermoepidermal junction. Approximately 25% of normal individuals display weak IgM staining at the dermoepidermal junction, whereas 70–80% of SLE patients have a positive LBT in sun-exposed, nonlesional skin. Half of SLE patients have a positive LBT in non-sun-exposed, nonlesional skin. It is unclear whether the LBT is a consequence of circulating ANA targeting denatured DNA from UV light–damaged keratinocytes, or a result of immune complex deposition or antibasement membrane antibodies. The test may be useful diagnostically and prognostically, as it correlates with increased systemic disease severity.

Hair and Nail

Hair involvement in SLE includes scarring alopecia resulting in permanent hair loss, induced by DLE. This can be differentiated from other common forms of scarring alopecia by immunofluorescent studies. Patchy or diffuse nonscarring alopecia is frequently associated with disease activity. Histologically, there is a peribulbar lymphocytic infiltrate surrounding shrunken anagen hair bulbs similar to findings in alopecia areata. Nonscarring alopecia resolves with complete hair regrowth with control of disease activity.

A wide spectrum of nail abnormalities, including pitting, ridging, onycholysis, and dyschromia with blue or black hyperpigmentation, are reported in up to 30% of SLE patients, but none are lupus-specific. Nail fold erythema with ragged cuticles and splinter hemorrhages resembling the changes of dermatomyositis are common.

Oral Lesions

The spectrum of oral lesions reported in SLE includes cheilitis, ulcerations, erythematous patches, lichen planus–type plaques on the buccal mucosa and palate, and DLE.[50] Most oral lesions are asymptomatic and must be looked for on examination. Positive immunofluorescent staining on biopsy may be useful to differentiate DLE from lichen planus–like lesions and leukoplakia. Lupus mucosal ulcerations demonstrate an interface mucositis and not leukocytoclastic vasculitis.

Gastrointestinal Manifestations

GI symptoms occur commonly in SLE with reported incidences of 15–75%; attribution is critical, as at least half are attributable to side effects of medications and to infectious complications.[51]

Esophagus

The prevalence of esophageal involvement varies. Many reviews citing a high incidence of dysphagia and odynophagia predate the advent of proton pump inhibitors and H_2 blockers, and the relationship of medication use to symptoms is not clear. Dysphagia and heartburn are reported in 1–50% of patients, although esophageal dysmotility is observed in up to 72%. An inflammatory process involving esophageal muscle or vasculitic damage to the Auerbach plexus is thought to contribute to the esophageal dysmotility. Ulceration is rarely seen outside the context of infections such as invasive candidiasis, herpes simplex, or cytomegalovirus. SLE patients with a secondary Sjögren syndrome may have salivary gland dysfunction, resulting in decreased saliva contributing to dysphagia.

Abdominal Pain/Vasculitis

Acute abdominal pain is common in SLE with reported incidences as high as 40%. The differential diagnosis includes intestinal vasculitis (45.5%), pancreatitis (10.8%), hepatobiliary disease (18.8%), and intestinal pseudoobstruction (IPO) (3.3%).[52] The most catastrophic and potentially fatal GI disturbances are related to ischemia of the small and large intestines resulting from medium- and small-vessel vasculitis or thrombotic complications of antiphospholipid (APL) antibodies. Approximately half of SLE patients with acute abdominal pain will have intestinal ischemia; associated mortality is high, so early consideration and intervention is critical. Intestinal vasculitis is frequently associated with active disease elsewhere and increased SLE Disease Activity Index (SLEDAI) scores, whereas SLE patients with inactive disease and acute abdominal pain will have different intraabdominal pathology unrelated to SLE. The clinical presentation may be acute, severe abdominal pain or an insidious, stuttering course with nausea, vomiting, bloating, diarrhea, postprandial fullness, anorexia, and weight loss. Mesenteric vasculitis preferentially affects the superior mesenteric artery, involving the small intestine more commonly than the large bowel. Vasculitis can also occur in the esophagus, stomach, peritoneum, rectum, gallbladder, pancreas, and liver. Computed tomography (CT) and/or magnetic resonance with or without angiography are the preferred imaging tests for evaluation of abdominal pathology; the radiographic signs of intestinal ischemia do not differ based on pathogenesis.

In cases with an insidious clinical course, endoscopy and colonoscopy may provide evidence of ischemia demonstrating ulcerating or heaped-up lesions with overt vasculitis on biopsy. The lesions are segmental and focal. Histologically, there is a small-vessel arteritis and venulitis with neutrophilic, lymphocytic, and macrophage infiltrates and fibrinoid necrosis of the vessel walls, associated thrombosis, and mononuclear infiltrate in the lamina propria. There may be immunoglobulin, C3, and fibrin deposition in the adventitia and media.

Timely treatment with high-dose intravenous corticosteroids (1–2 mg/kg per day) usually results in a favorable outcome and avoidance of serious complications such as bowel necrosis and perforation. To date there are no randomized trials of immunosuppressive agents for the treatment of intestinal vasculitis.

Intestinal Pseudoobstruction

Though rare, SLE-associated intestinal pseudoobstruction (SLE-IPO) may be the initial manifestation of SLE.[51,53] Clinical symptoms (abdominal pain and distension with diminished or absent peristalsis) and radiographic findings of SLE-IPO mimic those of mechanical obstruction. Distinguishing features of SLE-IPO include concomitant active SLE in other organ systems and associations with hematological cytopenias, hypocomplementemia, and serositis. Findings of coexisting ureterohydronephrosis and hepatobiliary dilatation in the absence of obstructing lesions suggest underlying smooth muscle dysmotility. Poor prognosis is associated with older age at SLE diagnosis, GI symptoms as the initial SLE manifestation, longer disease duration, and delayed diagnosis of IPO.

Peritonitis

Inflammation of serosal membranes is well described in SLE; despite evidence of peritoneal inflammation in 63% of autopsy studies, symptomatic pericarditis and pleuritis occur far more commonly than peritonitis. Acute peritonitis may be attributed to peritoneal vasculitis or ischemia and presents with abdominal pain (see above). The finding of ascitic fluid by CT scan or ultrasound mandates an evaluation of the fluid to exclude infection and malignancy. Rarely, ascites may be attributable to hepatic or portal vein thrombosis. Chronic peritonitis characterized by large amounts of painless ascites attributable to SLE and not to heart failure, constrictive pericarditis, or severe hypoalbuminemia due to nephrotic syndrome, liver disease, or a protein-losing enteropathy (PLE) is rare. In lupus peritonitis, the ascitic fluid is generally exudative with a predominance of lymphocytes; LE cells, autoantibodies, and low complement levels are frequent. On biopsy, the peritoneum is usually edematous; it is sometimes hemorrhagic with lymphocytic perivascular infiltrates.

Pancreatitis

Pancreatitis attributable to SLE is rare, occurring in 8–11% of patients with abdominal pain and having an annual reported incidence ≤1/1000. Although both corticosteroids and azathioprine can trigger pancreatitis, 34% of reported cases are not on these medications at the onset of pancreatitis, and most patients respond to steroid therapy. The clinical presentation and diagnosis of pancreatitis is similar in patients with and without SLE. Specific findings in SLE include leukopenia, thrombocytopenia, and anemia. Inflammation and necrosis are common on biopsy; there is a single report of pancreatic vasculitis. Mortality rates are reported from 18–27%; poor outcome is associated with increased systemic SLE activity (particularly, low complement and thrombocytopenia).

Liver

The term "lupoid hepatitis" was coined in the 1950s to describe cases of young women with chronic active hepatitis characterized by a lymphoplasmacytic infiltrate on biopsy, hypergammaglobulinemia, and a positive LE cell test in the blood. With the development of ANA testing, it became apparent that "lupoid" and "chronic active hepatitis" were indistinguishable in terms of clinical presentation, pathology, and response to treatment, and the term "autoimmune hepatitis" (AIH; Chapter 76) was endorsed in 1993.[54] Importantly, only 10% of AIH meet criteria for SLE, and 2.4–4.7% of SLE patients have noninfectious hepatitis attributed to SLE.[51] Analysis of SLE liver pathology from 52 autopsies in Japan revealed a variety of findings, including congestion (40/52), fatty liver (38/52), arteritis (11/52), cholestasis (9/52), and a few cases each of chronic persistent hepatitis, nodular regenerative hyperplasia, hemangioma, and cholangiolitis. Cirrhosis is an extremely rare complication of liver involvement in SLE. The distinction between subclinical liver inflammation related to SLE (lupus hepatitis [LH]) and AIH is important because therapy and prognosis are different. Whereas LH is associated with mild enzyme abnormalities, AIH is a progressive disease frequently leading to hepatic failure. Both conditions share a predilection for young women, and both demonstrate features of autoimmunity, including hypergammaglobulinemia, arthralgias, and serum autoantibodies. Histologically, biopsies of LH reveal lobular and periportal lymphocytic infiltrates, in contrast to the periportal and piecemeal necrosis with dense lymphocytic infiltrates seen in AIH with progression to panlobular or multilobular necrosis and cirrhosis. Serologically, antiribosomal P antibodies have been associated with liver disease in LH, whereas AIH is associated with antibodies to liver and kidney microsomes. Antismooth-muscle antibodies are observed in 60–80% of patients with AIH, compared with 30% of patients with LH. Although both have a favorable response to steroids, AIH usually requires additional immunosuppressive agents.[54]

Distinguishing LH from hepatitis C (HCV) can be difficult. Up to 30% of patients chronically infected with HCV have low titers of ANA and other autoantibodies (anti-DNA, anticardiolipin antibodies, and rheumatoid factor). They can also have cryoglobulins and associated cryoglobulinemic vasculitis. It is necessary to confirm a positive enzyme-linked immunosorbent assay (ELISA) for HCV with polymerase chain reaction (PCR) in patients presenting with arthritis, cutaneous vasculitis, and a positive ANA, as SLE patients may have false-positive serological tests for HCV.

Protein-Losing Enteropathy (PLE)

Profound hypoalbuminemia in the absence of severe nephrotic syndrome, liver disease, or constrictive pericarditis should trigger concern for PLE. Clinically significant PLE is uncommon in SLE, with reported prevalence rates of 1–8%; it can present individually or in the context of severe disease activity with other organ involvement. The diagnosis is confirmed with an increased α_1-antitrypsin level in a 24-hour stool collection or with the presence of IV-administered labeled human albumin in the stool. A recent review of 112 cases reported an Asian predominance (65%) and clinical features of edema (805), ascites (38%), pleural effusion (38%), pericardial effusion (21%), and hypocomplementemia (79%). Approximately half the patients experienced diarrhea; intestinal histology revealed a range of findings, including lymphangiectasis, edematous villi, inflammatory infiltrate, vasculitis, and mucosal atrophy. Biopsy results suggest a role of TNF-α, IFN-γ, and IL-6 in the increased vascular and enterocyte permeability in PLE.

Pulmonary Involvement

Lupus affects the lungs in diverse ways involving the pleura, lung parenchyma, and blood vessels. The most frequent and important complicating feature is infection.

Pleuritis

Pleuritis is the most common pulmonary manifestation of SLE, reported in 40–56% of patients.[55] Pleural involvement in up to 93% of lupus patients at autopsy suggests that much pleuritis may be asymptomatic. Clinically, patients note typical pleuritic pain. On physical examination, the most frequent abnormality is tachypnea; a pleural friction rub is present in some cases, and pleural effusions occur in more severe cases. Pleural fluid is usually exudative with normal glucose and pH and elevated protein levels. The leukocyte count can range from several hundred to 20 000 cells/μL; both a lymphocytic and neutrophilic predominance are reported. When performed, immunological testing on pleural fluid may show reduced complement levels and the presence of ANA, anti-DNA antibodies, and LE cells. Although these tests are commonly obtained, these results are neither sensitive nor specific to diagnose lupus pleuritis.

Lupus Pneumonitis

Lupus pneumonitis occurs in up to 10% of patients. Patients present with dyspnea, cough, mild pleuritic chest pain, and fever. Pulmonary infiltrates are present on plain radiograph or CT. This presentation must be distinguished from an infectious etiology. Histological examination of affected lung tissue shows alveolar edema and hemorrhage with hyaline membrane formation; immunofluorescent staining reveals immune complex deposition.

Pulmonary Hemorrhage

Pulmonary hemorrhage is a rare but potentially fatal complication of SLE. Symptoms include shortness of breath and hemoptysis accompanied by a fall in hemoglobin, usually occurring in the context of multiorgan involvement from SLE. Imaging may show patchy infiltrates. Pulmonary function testing is marked by an increased diffusion capacity (DL_{CO}) secondary to the presence of alveolar blood, whereas arterial O_2 saturation is decreased. Histopathology shows bland intraalveolar hemorrhage and hemosiderin-laden macrophages, although microangiitis with an inflammatory infiltrate and necrosis of the alveolar septa can occur. As hemorrhage into the lung may be secondary to thrombotic thrombocytopenia, infections, or pulmonary hypertension, demonstration of an inflammatory process in the pulmonary vessels or tissue is helpful to establish a diagnosis of primary pulmonary hemorrhage.

Chronic Diffuse Interstitial Lung Disease

Chronic diffuse interstitial lung disease is a relatively uncommon manifestation of SLE and is occasionally associated with anti-Ro antibodies. It usually has a progressive course with a chronic nonproductive cough, dyspnea, and pleuritic chest pains. Physical examination is frequently remarkable for basilar rales with diminished diaphragmatic movement. Pulmonary function tests demonstrate a restrictive pattern with decreased diffusion capacity; oxygen saturation is decreased. Imaging often shows interstitial fibrosis that is more prominent at the lung bases. High-resolution CT (HRCT) may also help determine the extent of treatable disease, *i.e.,* fibrosis (honeycombing) versus inflammation (ground glass). However, the most reliable method to assess the extent of pulmonary inflammation in comparison to fibrotic damage is histological examination. Evaluation of bronchial alveolar lavage fluid helps to exclude infection.

Pulmonary Hypertension

Pulmonary hypertension unrelated to chronic pulmonary emboli or interstitial lung disease occurs in SLE. Severe cases are rare; the recent recognition of milder cases may be partially attributed to increased awareness and findings of elevated pulmonary artery pressures on echocardiograms obtained for evaluation of shortness of breath. Patients typically present with progressive dyspnea occurring in the absence of infiltrates on chest radiographs or significant hypoxemia. Chest pain and a chronic nonproductive cough are also frequently present. Pulmonary function testing reveals a reduced DL_{CO}. Elevated pulmonary artery pressure is confirmed with cardiac angiogram. Biopsy or autopsy specimens of the lung reveal "plexiform" lesions that resemble those seen in primary pulmonary hypertension.

Shrinking-Lung Syndrome

The shrinking-lung syndrome refers to the rare findings of shortness of breath occurring in the absence of pleuritis or interstitial lung disease plus a chest X-ray showing elevated hemidiaphragms. Pulmonary function testing shows a restrictive pattern with loss of lung volume. It had been generally accepted that this syndrome results from diaphragmatic weakness (from a myopathic process) or chest wall restriction; however, more recent studies suggest that pleuritis may play a causative role. There is no definitive therapy, although immunosuppressive therapy with cytotoxic agents usually results in an improvement of lung function and respiratory symptoms.

Cardiac Involvement

There are a number of ways in which lupus affects the cardiovascular system; targets include the myocardium, valves, pericardium, and vessels.

Myocardium

Myocardial dysfunction in SLE is likely to be secondary to factors other than lupus, such as hypertension, medications, or coronary artery disease (CAD). However, a cardiomyopathy resulting from immune-mediated myocardial inflammation does occur, either in isolation or concomitant with myositis or other manifestations of systemic disease. Inflammatory myocarditis is often associated with anti-RNP antibodies. Histopathology typically shows a mononuclear, inflammatory cell infiltrate. Perivascular or myocardial wall deposits of immune complexes and complement also occur. Myocardial biopsy is critical for diagnosis as well as for determining the extent of active inflammatory disease and fibrosis. Symptoms and signs of myocarditis include unexplained tachycardia, an abnormal electrocardiogram (with ST- and T-wave abnormalities), cardiomegaly, and heart failure. Echocardiography may show systolic and diastolic ventricular dysfunction. Myocardial involvement without overt clinical signs occurs commonly and may be documented using Doppler echocardiography. A noninflammatory cardiomyopathy may be seen in association with high-dose cyclophosphamide and, although rarely, with hydroxychloroquine.

Valvular Heart Disease

Valvular abnormalities, with thickening, regurgitation, or verrucous vegetations, occur commonly in SLE (50–60%) and are documented by transesophageal echocardiography.[56] The characteristic Libman–Sacks lesion, nonbacterial verrucous vegetations, is observed at autopsy in 15–60% of patients. Mitral, aortic, and tricuspid valves are most frequently involved. Clinically these lesions are usually asymptomatic, and hemodynamic compromise, rupture of the chordae tendineae, or infection are rare events. On histological examination, mononuclear cells, hematoxylin bodies, fibrin and platelet thrombi, and immune complexes are present. Antiphospholipid antibodies can be associated with the development of valvular heart disease, although their pathogenic role remains unproven.

Pericarditis

Pericardial inflammation in SLE is frequent. Although asymptomatic pericarditis occurs in more than 50% of patients, clinically apparent pericarditis occurs in only 25%, and cardiac tamponade and constrictive pericarditis are infrequent. Pericardial fluid and thickening are easily detected by echocardiography; cardiac silhouette enlargement on plain films is seen in the presence of large effusions. There are no unique signs and symptoms of pericarditis in lupus patients. The histological findings of acute pericarditis in lupus are inflammation with a mononuclear cell infiltrate accompanied by immunoglobulin and complement deposition. Results of pericardial fluid analysis are neither sensitive nor specific; the fluid is usually an exudate with elevated protein concentrations, normal or low glucose levels, and an elevated WBC count that is primarily neutrophilic. Complement levels in the fluid are low, and autoantibodies (ANA, dsDNA) and LE cells have been reported.

Coronary Artery Disease

Myocardial infarction, angina, and sudden death resulting from CAD are well described in SLE. Estimates of the prevalence of CAD vary widely depending on the methodology used for ascertainment; however, the risk of myocardial infarction has been estimated to be 50-fold greater in young women with lupus than in normal age-matched controls.[57] Cardiac events occur both late in the course of SLE as well as early, and may even predate the SLE diagnosis. Coronary artery vasculitis is a potential cause of CAD but is exceedingly rare, and bland atherosclerotic plaque is typically described in surgical and postmortem specimens. Traditional risk factors for CAD, such as hypertension, diabetes, and hyperlipidemia, are enriched in lupus patients, and an increased prevalence of the metabolic syndrome likely contributes additional risk. Lupus-related risk factors include duration of SLE, duration of corticosteroid use, renal disease, and absence of use of hydroxychloroquine. The potential contributing influence of antibodies to phospholipids or disease activity continues to be explored. There is a growing body of evidence supporting a risk conferred by inflammation *per se.* The atherosclerotic plaque is infiltrated by activated macrophages and T cells that produce proinflammatory cytokines. Inflammation may accelerate both the formation and the rupture of atherosclerotic plaque. In SLE, the pathology of acute events seems to focus on unstable plaque. If endothelial dysfunction and vascular injury are the events triggering atherosclerosis, there are multiple potentially responsible processes in lupus; these include, but are not limited to, autoantibodies directed to endothelial cells, immune complexes, and enhanced cleavage of membrane endothelial protein C receptor.[58]

Renal Involvement

Lupus nephritis is a common manifestation of disease with significant impact on morbidity and mortality (Chapter 68). The prevalence of nephritis ranges from 50–75% overall, with increased prevalence of proliferative nephritis and more aggressive disease in African Americans and Hispanics compared with Caucasians. Low socioeconomic status, independent of ethnicity, is predictive of poor prognosis, and pediatric lupus and male lupus are both associated with a greater incidence of, and more aggressive, nephritis. Onset of nephritis frequently occurs within 2 years after diagnosis, but it may occur at any time, so monitoring for potential renal activity is an ongoing obligation. Clinically, patients are asymptomatic unless they are nephrotic or have developed end-stage renal disease. Detection typically relies on examination of the urine, although a rising creatinine or hypertension may herald renal involvement. The presence of proteinuria on urinalysis, hematuria (>5 red blood cells (RBCs)/high-power field), or pyuria (>5 WBCs/high-power field) in the absence of other etiologies should prompt an evaluation for nephritis. A 24-hour urine collection remains the most accurate measurement of urinary protein loss; however, the protein/creatinine ratio in a spot urine is accepted and more commonly used when monitoring patients. Monitoring serum creatinine as a surrogate for the glomerular filtration rate is standard; however, creatinine is an insensitive marker of lupus renal disease and should be used in conjunction with other assays. Renal activity is usually preceded or accompanied by serological activity. Antibodies to dsDNA are almost always elevated or rising, whereas measurements of serum complement (C3, C4, or CH_{50}) are usually low or dropping.

Histological findings of the kidney in lupus nephritis can be defined using the classification proposed by the International Society of Nephrology/Renal Pathology Society (ISN/RPS).[59] In general, membranous nephritis (class V) presents with a bland urinary sediment (*i.e.,* no RBCs, WBCs, or casts), nephrotic-range proteinuria, a normal to mildly elevated creatinine, a normal blood pressure, and normal serologies. Patients with mesangial disease (class II) present with a bland or minimally active sediment, low-grade proteinuria (less than 500 mg/24 hours), and a normal serological profile. Class III (focal) and class IV (diffuse) proliferative nephritis are characterized by active urinary sediment, proteinuria (>500 mg/24 hours), active serologies, and, frequently, hypertension and elevated serum creatinine. Class III is defined as ≤50% glomerular involvement, and class IV is defined as >50%. The extent of proteinuria, urinary sediment activity, serological abnormalities, and creatinine elevation is often less in class III than in class IV renal disease. Most cases of class II nephritis do not require initiation of cytotoxic therapy, and progression to end-stage kidney disease is rare. The prognosis of class III disease is dependent on the degree of activity; patients with 40–50% of glomerular involvement have a prognosis similar to that of patients with class IV disease. In addition to the number of involved glomeruli, renal biopsy assessment includes measures of activity (proliferative response) and chronicity (sclerotic response). Therapeutic intervention in class III or IV disease usually requires cytotoxic therapy in addition to high-dose corticosteroids, provided the chronicity index is not too high, indicating irreversible damage. Even with potent immunosuppressant therapy such as cyclophosphamide or mycophenolate mofetil, a complete response is induced in only approximately 20%, with a partial response in about 80% of patients. Moreover, repeat histopathological studies of kidney biopsies in patients achieving a complete clinical response show ongoing renal inflammation in almost half. Relapses and flares of renal disease are not infrequent, particularly when tapering corticosteroids or discontinuing immunosuppressive treatment. Potential therapies that maintain podocyte integrity or prevent activation of renal endothelial cells as well as agents directed against inflammatory cytokines, B cells, or T cells may offer therapeutic advantage and improved renal outcome.[60]

Hematological

Hemocytopenias occur frequently in SLE, and prevalence varies among lupus cohorts (Fig. 51.7). Evaluation for medication effects is essential before attributing a cytopenia to an immune-mediated mechanism.

Anemia

Antibody-mediated peripheral destruction of RBCs, autoimmune hemolytic anemia (AHA), occurs in 5–14% of SLE patients.[61] In the multiethnic Latin American GLADEL cohort, AHA was an independent predictor of damage accrual and decreased survival and was associated with disease activity.[62] The antierythrocyte antibody is usually an opsonizing IgG. The specificities of the antierythrocyte antibodies in SLE have not been clearly defined; the only nonrhesus-specific antigen reported in SLE is the membrane band 3 anion transporter protein. The presence of AHA is readily diagnosed by a positive Coombs test, elevated lactate dehydrogenase and total bilirubin, low serum haptoglobin, and the presence of spherocytes on the peripheral smear. AHA in SLE has also been associated with APL antibodies, which may reflect cross-reactivity with erythrocyte membrane antigens.[63]

Anemia of chronic disease is the most common cause of anemia in SLE; however, if the hemoglobin is <10 g/dL, another cause of anemia should be considered. The inhibitory effects of proinflammatory cytokines on erythrocyte production can be compounded by increased hepcidin levels and an inadequate erythropoietin response. Low erythropoietin levels in SLE may be attributable to renal disease as well as antierythropoietin antibodies. Although hemophagocytosis of hematopoietic cells is frequently noted on bone marrow biopsies, the hemophagocytic syndrome characterized by spiking fevers, tender hepatosplenomegaly, anemia, leukopenia, and markedly elevated serum ferritin is rare.

Lupus is also associated with a thrombotic microangiopathic hemolytic anemia with schistocytes, helmet cells, and triangular fragments of RBCs. The clinical constellation of high fever, renal insufficiency, neurological symptoms, and thrombocytopenia is characteristic of thrombotic thrombocytopenic purpura (TTP). Coexistent TTP and SLE is a rare and frequently fatal phenomenon that has been associated with antibodies to ADAMTS 13 (von Willebrand factor cleaving protease) in 16% of SLE-associated TTP and APL antibodies.

Leukopenia

Leukopenia, either neutropenia or lymphopenia, occurs in 20–40% of patients.[64] Lymphopenia (lymphocytes <1500/mm^3) has a reported prevalence of 15–80%, whereas severe lymphopenia (<500/mm^3) occurs in 4–10% of SLE patients. The prevalence of neutropenia, variably defined in the literature as <1800 to <2500/mm^3, is 20–40%, whereas severe neutropenia (<1000/mm^3) is rare (1–4% prevalence). Both neutropenia and lymphopenia can reflect disease activity; however, associations with infection remain controversial. Carli et al. found that associations between leukopenia and major infection lost significance in half of published studies after controlling for confounding factors.[64]

Autoantibodies are the pathogenic mechanism generally implicated in the leukopenias. Antineutrophil antibodies directed against membrane components of mature and progenitor cells, resulting in decreased phagocytosis and accelerated apoptosis, are implicated, as are antibodies against granulocyte–colony-stimulating factor (G-CSF) and hyposensitivity of myeloid cells to G-CSF. Binding of TNF-related apoptosis-inducing ligand (TRAIL) to TRAIL receptors on neutrophils accelerates neutrophil apoptosis.[65] One of four known TRAIL receptors is unable to transduce the death signal and functions as a decoy receptor; this receptor is reduced in SLE patients with neutropenia and is upregulated by steroid therapy. Lymphocytotoxic antibodies, particularly to CD4 T cells, increased apoptosis related to Fas and Fas ligand upregulation, and increased serum IL-10 levels have all been implicated in the pathogenesis of lymphopenia. B cells expressing the 9G4 idiotype found on V_H 4.34 heavy chains may be responsible for production of antilymphocyte antibodies.

Treatment of leukopenia is guided by the clinical scenario, and exclusion of drug effects, malignancy, or myelofibrosis is required before attribution to SLE. In the absence of recurrent infection, leukopenia in SLE rarely warrants treatment, as increased steroids can also contribute to the risk of infection. Severe leukopenia has been treated with G-CSF; however, although it is rapidly effective, G-CSF has also been associated with disease flare in 30% of cases.

Thrombocytopenia

Low platelet counts are seen in approximately 25% of patients, although severe thrombocytopenia is reported in fewer than 10%. Immune-mediated consumption is the most frequent cause, but in rare instances thrombocytopenia occurs as a manifestation of microangiopathic hemolytic anemia, TTP, disseminated intravascular coagulation (DIC), or as part of the hemophagocytic syndrome, all of which are associated with high mortality and morbidity. A pathogenic role for antibodies against platelet membrane glycoproteins (IIb/IIIa antigen) is well established.[66] Other possible mechanisms include antibodies to the thrombopoietin receptor (TPOR), and APL and anti-CD40-ligand antibodies that bind to platelets, resulting in platelet sequestration.

Central and Peripheral Nervous System

Neurological and psychiatric manifestations of SLE are diverse and frequently occur irrespective of systemic disease activity. Neuropsychiatric symptoms can be focal or diffuse, peripheral or central, isolated or complex, and difficulties with assessment and attribution of individual symptoms have hindered progress in understanding mechanisms for neuropsychiatric disease attributable to SLE (NPSLE).

NPSLE Nomenclature

In 1999 a consensus committee established by the ACR developed reporting standards, case definitions, and recommendations for laboratory and imaging studies for 19 neurological, psychiatric, and cognitive syndromes. Diagnosis of NPSLE requires the exclusion of infections, metabolic disturbances, bleeding disorders, malignancy, and medication toxicities. Using this nomenclature, NPSLE is common overall, with a prevalence of 57–95%. Individually, headache, mood disorders, cerebrovascular events, cognitive dysfunction, and seizures are frequent; polyneuropathy, demyelinating disease, mononeuritis, myasthenia gravis, chorea, cranial neuropathy, myelopathy, and Guillain–Barré syndrome are uncommon. NPSLE can be further subdivided into syndromes occurring in the anatomically distinct central (CNS NPSLE) and peripheral (PNS NPSLE) nervous systems.

Manifestations of CNS NPSLE include focal neurological syndromes (stroke, seizures, aseptic meningitis, movement disorder, myelopathy, demyelinating syndrome) and diffuse psychiatric/neuropsychological syndromes (mood and anxiety disorders, psychosis, acute confusional state, cognitive dysfunction, headaches). PNS NPSLE syndromes include cranial neuropathy, polyneuropathy, mononeuropathy, myasthenia gravis, Guillain–Barre syndrome, plexopathy, and autonomic neuropathy (Fig. 51.7).

NPSLE Pathogenesis

Many of the focal CNS NPSLE syndromes occur in association with APL antibodies and thrombotic events or, more rarely, with small-vessel vasculitis leading to intraluminal thrombosis. In contrast, proposed mechanisms for the diffuse CNS NPSLE syndromes involve autoantibody-mediated neurotoxicity and complicated neuroimmune interactions, resulting in increased CNS expression of proinflammatory cytokines with deleterious effects on neuronal and microglial cell function and adult neurogenesis.[67,68] Proposed mechanisms for diffuse CNS NPSLE rely in large part on disruption of the blood–brain barrier (BBB)—a dynamic interface between the brain and circulating hormones and inflammatory molecules. Evidence for BBB disruption in SLE is supported by reports of an elevated albumin concentration gradient between cerebrospinal fluid (CSF) and plasma, an elevated IgG index, and elevated serum levels of proteins whose origins are exclusive to the brain parenchyma, such as S100B. Proposed mechanisms for BBB disruption in SLE include brain endothelial cell disruption mediated by inflammatory cytokines, chemokines, complement C5a, antiendothelial cell antibodies, anti-NMDAR antibodies, and TWEAK (TNF-like weak inducer of apoptosis).[69–71]

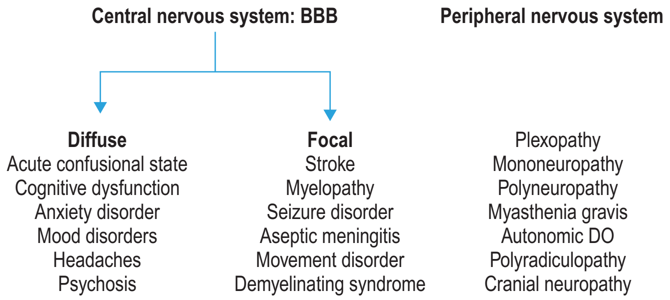


FIG 51.7 Neuropsychiatric Syndromes in Systemic Lupus Erythematosus.

This distinction between focal and diffuse CNS NPSLE and potential mechanisms is essential for the development of treatment strategies. Thrombotic events associated with APL antibodies are treated with anticoagulation, whereas high-dose corticosteroids and immunosuppression in combination with neuroleptic and antiseizure medications are generally required for treatment of diffuse CNS events. Because there is no intervening BBB, peripheral nerves are more accessible to circulating complement, autoantibodies, and inflammatory molecules, and vasculitis of epineural arteries is a common finding in PNS NPSLE. Treatment consists of corticosteroids and immunosuppression or intravenous immunoglobulin (IVIG).

NPSLE Assessment and Attribution

The diagnosis of diffuse CNS NPSLE is frequently difficult, and the clinical evaluation of a patient with presumed CNS NPSLE mandates an exhaustive search for other potential causes. CSF examination is useful for excluding infection or malignant cells. Although CSF in NPSLE can be characterized by a lymphocytosis and increased immunoglobulin with elevated total protein, IgG index, and oligoclonal bands, these abnormalities are not consistently present. Numerous autoantibodies (ANA, anti-dsDNA, anticardiolipin, antiribosomal P, anti–N-methyl D-aspartate receptor [NMDAR], antineuronal, anti-RNP, anti-Sm) and various cytokines have been identified in the CSF of patients with SLE; however, none is specific for active NPSLE or individual CNS NPSLE syndromes, and routine testing is not recommended. MRI is extremely sensitive for detection of structural lesions and new ischemic lesions associated with focal CNS NPSLE, but it is frequently not helpful in diagnosing active diffuse CNS NPSLE, as MRI studies may be normal in patients with psychiatric syndromes and global CNS dysfunction.[72] Imaging studies of neuronal function, positron emission tomography (PET), and single-photon emission computed tomography (SPECT) scans measuring cerebral glucose uptake and blood flow, respectively, magnetic resonance spectroscopy, and functional MRI (fMRI) correlate to a limited extent with diffuse CNS NPSLE. These scans must be interpreted carefully because atrophy with neuronal cell loss leads to changes in metabolism and blood flow. Cerebral angiograms can be helpful if a diagnosis of cerebral vasculitis is considered; however, CNS vasculitis is rare, with true vasculitis in only 5–8% of autopsies. More commonly, a bland vasculopathy with degenerative and proliferative changes in small vessels, perivascular infiltrates, microinfarcts, and microhemorrhages is present.

There are no serological tests specific for CNS NPSLE. Serological evidence of disease activity (elevated anti-DNA antibodies and low complement) may help to diagnose CNS NPSLE, particularly if combined with other clinical signs of active disease. However, CNS NPSLE can also flare in the absence of serological and clinical disease activity. Cognitive impairment in particular tends to develop insidiously, irrespective of peripheral disease activity. Recent data from murine models and *in vitro* studies strongly support the role of specific autoantibodies directed against NMDAR, phospholipid, α tubulin, neuronal surface P antigen, and ribosomal P in NPSLE.[73] Anti-NMDAR and anti-P antibodies have been identified in the serum, CSF, and brain of SLE patients and have been associated with cognitive and depressive syndromes in some SLE patients. CSF anti-NMDAR antibodies have also been reported to be increased in patients with acute confusional state. The current challenge is to identify easily accessible, sensitive, and specific serological or imaging biomarkers for the diagnosis of CNS NPSLE syndromes that can also be used to monitor therapy.

PNS NPSLE events are recognized on the basis of clinical presentation with diagnostic studies such as electromyography and peripheral nerve biopsy.

Antiphospholipid antibody syndrome is described in detail in Chapter 61.

Drug-Induced Lupus

Until recently, drug-induced lupus referred to a clinical syndrome characterized by constitutional, musculoskeletal symptoms and serositis that resembles mild lupus and occurs after exposure to a number of drugs (most notably procainamide, hydralazine, chlorpromazine, and methyldopa). Drug-induced lupus has been associated with ANA and antihistone antibodies, whereas generation of more specific autoantibodies such as anti-DNA antibodies has been rare. Symptoms generally begin weeks to months after initiation of the inciting therapeutic agent and resolve within weeks after the drug is discontinued; autoantibodies can persist for up to 12 months. Host factors affecting drug metabolization (*e.g.*, slow acetylation of procainamide and hydralazine) contribute to the risk of developing drug-induced lupus. Multiple potential mechanisms resulting in the loss of self-tolerance have been suggested for this classic model. One of the most extensively explored is inhibition of DNA methylation with TLR activation, resulting in overexpression of costimulatory molecules such as leukocyte function–associated antigen-1 (LFA-1) on T cells and enhanced T-cell help.[74]

Since the introduction of anti-TNF agents, a different variant of drug-induced lupus has been recognized. Up to 30% of patients receiving TNF blockade develop autoantibodies, including a positive ANA and antibodies to dsDNA. A minority of patients develop anti-DNA antibodies, nephritis, and vasculitis. The pathogenesis of this immunological deregulation is not known.

Treatment

THERAPEUTIC PRINCIPLES

Goals of Therapy in SLE

I. Suppression of inflammation
  - Induction of remission
  - Maintenance of remission
  - Preservation of organ function

II. Suppression of immune activation
  - Modulation of the immune response

III. Treatment/prevention of drug-related toxicities

The goals of lupus treatment are to stop and reverse ongoing organ inflammation, to prevent or limit irreversible organ damage, and to suppress the immune response driving the inflammation and prevent flares. Therapeutic agents, and combinations thereof, can be used for induction of remission, maintenance of remission, or prevention of flare. The efficacy of therapeutic agents must be balanced against their potential toxicity. Thus treatment must be tailored to the individual patient based on disease manifestations. In general, milder disease requires treatment with less potent or lower doses of antiinflammatory and immunosuppressive medications than more active, severe disease affecting major organs such as the kidney or brain (Table 51.4). Individual patient responses to a given medication will vary, and patients must be monitored closely for response as well as toxicity.

"Treat-to-target" recommendations in SLE from an international SLE task force include the following more specific guidelines: the treatment target should be remission of organ manifestations; factors such as pain that negatively influences health-related quality of life should be addressed; maintenance treatment should aim for the lowest glucocorticoid dose needed to control disease; and relevant therapies adjuvant to any immunomodulation should be considered to control comorbidity in SLE.[75]

Some genetic factors predicting risk of toxicity or therapeutic benefit for individual agents have been identified. Polymorphisms of a key enzyme in the metabolism of azathioprine, thiopurine methyltransferase (TPMT), are common; 0.3% and 11% of Caucasians are homozygous and heterozygous, respectively, for mutations associated with altered expression of TPMT.[76] TPMT-deficient patients are especially susceptible to leukopenia and pancytopenia associated with azathioprine. Cyclophosphamide is metabolized to its active form by cytochrome P450. Individuals heterozygous or homozygous for a specific cytochrome P450 polymorphism (*CYP2C19*2*) have a lower probability of developing premature ovarian failure, but they also show a poorer response to therapy.[77]

Although corticosteroids are the foundation of treatment, exposure must be minimized to the greatest extent given their multiple and frequent side effects, including hypertension, diabetes mellitus, increased susceptibility to infection, bone loss, and weight gain (Chapter 86). Additional disease-modifying agents that are also steroid-sparing include antimalarials (hydroxychloroquine), antimetabolites such as azathioprine (1–2.5 mg/kg per day), methotrexate (7.5–25 mg/week), leflunomide (10–20 mg/day), and mycophenolate mofetil ([MMF], 2–3 g/day), and alkylating agents such as cyclophosphamide (monthly pulse 0.5–1.0 g/m^2) (Chapter 87). When more conventional therapies have failed, anecdotal reports, case series, and open-label studies suggest the use of intravenous immunoglobulin (2 g/kg over 2–5 days), thalidomide (50–100 mg/day), or cyclosporine (3–5 mg/kg/day). Medications such as dapsone, danazol, and chlorambucil may be efficacious in cutaneous disease, hematological disease, and in severe refractory disease, respectively, but in general these medications are not commonly used due to their toxicity and the introduction and availability of better-tolerated, efficacious agents. Belimumab, a monoclonal antibody directed against BAFF (Chapter 89), was approved by the FDA for treatment of SLE in 2011 and represents a therapeutic option for patients with nonrenal, non-CNS active disease.

A number of newer agents are on the horizon for treatment of SLE. Some are currently available but approved for other indications, whereas others are still in the developmental stages. These therapies, in general, have more specific immunological targets than standard treatments. Although multiple agents have been tested in animal models of disease, we will focus only on therapies that have been given to humans. Several of these agents have been recently evaluated in phase III clinical trials and have been unable to demonstrate superiority over treatment with placebo or standard of care. Although imperfect trial design may account for some of these failures, the possibility remains that these agents are not effective treatments for SLE and do not confer a substantial improvement over current standard-of-care therapies. Rituximab (anti-CD20 antibody), a B-cell–directed therapy approved for use in rheumatoid arthritis, targeting all B cells except plasma cells, failed to demonstrate improved performance over standard of care in trials of patients with active SLE, or in patients with active lupus nephritis added to a background therapy of MMF and corticosteroids. Elevated BAFF levels are present after B-cell depletion. These can exacerbate the impaired tolerance of SLE. Thus studies are ongoing to evaluate the efficacy of rituximab followed by belimumab. Abatacept (CTLA4-Ig), which is also approved for rheumatoid arthritis and blocks T-cell activation, was no better than placebo in trials of active lupus and of lupus nephritis, again added to background therapy of corticosteroids and MMF. A redesigned trial of abatacept for SLE is now being conducted. Epratuzumab (anti-CD22 antibody) targeting all B cells failed a phase III clinical trial. TACI-Ig, which blocks both BAFF and APRIL, is undergoing clinical study, although previous clinical trials were terminated prematurely because of infectious complications. Tocilizumab, an antibody to the IL-6 receptor, a cytokine involved in B-cell survival and activation as well as in differentiation of Th17 and Tfh cells, which is approved for RA, demonstrated unacceptable toxicity in SLE. Other agents in earlier clinical development are inhibitors of intracellular B-cell signaling molecules designed to block B-cell activation; R406, a SYK inhibitor; and Janus tyrosine kinase (JAK) inhibitors. The Jak inhibitors are being studied using topical administration for discoid lupus and oral administration for systemic disease.

Potential therapies aimed at DCs include vitamin D, which blocks DC maturation and T-cell activation, and inhibitory oligodeoxynucleotides, which block TLR9 signaling and DC maturation. However, a study of vitamin D supplementation designed to assess vitamin D inhibition of DC activation failed to demonstrate efficacy. Hydroxychloroquine, a standard agent for lupus, interferes with TLR7 and TLR9 signaling by preventing acidification of the endosomal compartment. A study of a novel RNase is under way.

TABLE 51.4 **Treatments for Systemic Lupus Erythematosus Disease Manifestations**

	NSAIDs	Antimalarials[†]	Low-Dose Corticosteroids (<0.5 mg/kg per day)	High-Dose Corticosteroids (>0.5 mg/kg per day)[‡]	Azathioprine	MMF	CTX
Constitutional	+	+	+		+		
Musculoskeletal	+	+	+		+	+	
Mucocutaneous		+	+				
Serositis	+	+	+		+		
Hematological			+	+	+		
Renal				+	+	+	+
Central nervous system				+			+
Vasculitis				+	+	+	+

CTX, cyclophosphamide; *MMF,* mycophenolate mofetil; *NSAIDs,* nonsteroidal antiinflammatory drugs.

[†] Antimalarials should be given to all patients to sustain remission and reduce damage.

[‡] High-dose corticosteroids are required for remission induction or an inadequate response to low-dose therapy.

Two studies of monoclonal antibodies against IFN-α failed to demonstrate efficacy; a trial of an antibody to the IFN receptor is under way.

Novel interventions directed at T cells include anti-CD40 antibodies, abatacept (CTLA4-Ig) in combination with cyclophosphamide for proliferative nephritis, and studies targeting ICOS, an inducible T-cell costimulator.

Therapies in development for other autoimmune diseases may also prove useful in SLE. Eculizumab, currently approved for paroxysmal nocturnal hemoglobinuria, is an antibody directed against C5 that blocks cleavage of C5 and the subsequent triggering of the complement cascade. Alicaforsen, an antisense oligodeoxynucleotide that inhibits ICAM-1 expression, decreases inflammation in both rheumatoid arthritis and Crohn disease. Efalizumab, a monoclonal antibody against CD11a (an LFA-1 subunit that interacts with ICAM-1), benefits patients with psoriasis. Although no chemokine-targeted therapies are in clinical trial in lupus, a CCR1 antagonist slowed disease progression in a mouse model of lupus and has been tested in patients with rheumatoid arthritis. FTY720, an agonist for the sphingosine 1 phosphate receptor that prevents egress of lymphocytes from secondary lymphoid organs and inflamed tissues, has beneficial effects in the MRL/*lpr* mouse model of lupus; it has been given to transplant recipients and to patients with multiple sclerosis. The use of anti-TNF-α agents for SLE is inhibited by its lupus-like effects in some patients. Paradoxically, TNF-α is involved in renal inflammation in SLE, and short-term treatment with monoclonal antibody to TNF-α appears to improve lupus nephritis, showing that agents can abort inflammation and yet exacerbate autoimmunity.

TRANSLATIONAL RESEARCH

ON THE HORIZON

- Increased understanding of the functional consequences of single nucleotide polymorphisms identified as susceptibility loci in systemic lupus erythematosus may provide insight into:
  - Dysregulation of tolerance mechanisms and pathogenesis of disease in different ethnic populations
  - Predictors of response to therapy
  - New therapeutic targets
- Development of organ-specific biomarkers will allow for improved use of immunomodulating and immunosuppressive therapies and possibly for use of preventive therapies.

Laboratory research continues to inform us about dysregulated immunological pathways in SLE. The current challenge of translational medicine is to shed light on the clinical relevance of these altered pathways. Increased understanding of the qualitative and quantitative differences of these molecular perturbations not only should result in improved diagnostic capabilities and biomarkers for disease, but also will lead to the ability to subset patients for prognosis and response to therapy. Improved organ-specific biomarkers and methods for subsetting patients on a molecular level are likely to result in improved use of immunological therapies. In particular the brain, which has been relatively inaccessible to mechanistic and diagnostic probing, is likely to be an organ to be more fully explored. Ultimately, advances in translational studies should result in more effective and less toxic therapeutic interventions.

CONCLUSIONS

In the past several years much has been learned about genetic susceptibility to SLE, and it is encouraging that many of the genes identified are associated with pathways that also have been implicated in disease pathogenesis. The role of B cells in disease pathogenesis has been confirmed, but recent studies have also highlighted the role of T cells, DCs, and neutrophils. Enhanced understanding of each of these contributing factors and the cross-talk between them has allowed identification of numerous potential therapeutic targets. The disappointing results from clinical trials of immunobiological agents for lupus have highlighted the critical importance of study design and the complexity of disease in humans. The overall goal in therapy must be to eliminate autoreactivity while maintaining immunocompetence. Among the challenges now faced are the careful phenotyping of patients to identify etiopathologically distinct subpopulations and new clinical trial designs to allow the use of combinations of agents, each of which alone may have a minor effect on disease course.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Tan EM, Cohen AS, et al. The 1982 revised criteria for the classification of systemic lupus erythematosus. Arthritis Rheum 1982;25:1271–7.
2. Yu C, Gershwin ME, et al. Diagnostic criteria for systemic lupus erythematosus: a critical review. J Autoimmun 2014;48–49:10–13.
3. Borchers AT, Naguwa SM, et al. The geoepidemiology of systemic lupus erythematosus. Autoimmun Rev 2010;9:A277–87.
4. Gonzalez LA, Toloza SM, et al. Impact of race and ethnicity in the course and outcome of systemic lupus erythematosus. Rheum Dis Clin North Am 2014;40:433–54, vii–viii.
5. Fors Nieves CE, Izmirly PM. Mortality in systemic lupus erythematosus: an updated review. Curr Rheumatol Rep 2016;18:21.
6. Gatto M, Iaccarino L, et al. Clinical and pathologic considerations of the qualitative and quantitative aspects of lupus nephritogenic autoantibodies: A comprehensive review. J Autoimmun 2016;69:1–11.
7. Ueda H, Howson JM, et al. Association of the T-cell regulatory gene CTLA4 with susceptibility to autoimmune disease. Nature 2003;423:506–11.
8. Gregersen PK, Olsson LM. Recent advances in the genetics of autoimmune disease. Annu Rev Immunol 2009;27:363–91.
9. Deng Y, Tsao BP. Genetic susceptibility to systemic lupus erythematosus in the genomic era. Nat Rev Rheumatol 2010;6:683–92.
10. Orozco G, Sanchez E, et al. Association of a functional single-nucleotide polymorphism of PTPN22, encoding lymphoid protein phosphatase, with rheumatoid arthritis and systemic lupus erythematosus. Arthritis Rheum 2005;52:219–24.
11. Manderson AP, Botto M, et al. The role of complement in the development of systemic lupus erythematosus. Annu Rev Immunol 2004;22:431–56.
12. Baumann I, Kolowos W, et al. Impaired uptake of apoptotic cells into tingible body macrophages in germinal centers of patients with systemic lupus erythematosus. Arthritis Rheum 2002;46:191–201.
13. Anolik JH, Barnard J, et al. Rituximab improves peripheral B cell abnormalities in human systemic lupus erythematosus. Arthritis Rheum 2004;50:3580–90.
14. Mackay M, Stanevsky A, et al. Selective dysregulation of the FcgammaIIB receptor on memory B cells in SLE. J Exp Med 2006;203:2157–64.
15. Floto RA, Clatworthy MR, et al. Loss of function of a lupus-associated FcgammaRIIb polymorphism through exclusion from lipid rafts. Nat Med 2005;11:1056–8.

16. Wong M, Tsao BP. Current topics in human SLE genetics. Springer Semin Immunopathol 2006;28:97–107.
17. Tucci M, Barnes EV, et al. Strong association of a functional polymorphism in the monocyte chemoattractant protein 1 promoter gene with lupus nephritis. Arthritis Rheum 2004;50:1842–9.
18. Crow MK. Interferon pathway activation in systemic lupus erythematosus. Curr Rheumatol Rep 2005;7:463–8.
19. Nath SK, Harley JB, et al. Polymorphisms of complement receptor 1 and interleukin-10 genes and systemic lupus erythematosus: a meta-analysis. Hum Genet 2005;118:225–34.
20. Schotte H, Willeke P, et al. Extended haplotype analysis reveals an association of TNF polymorphisms with susceptibility to systemic lupus erythematosus beyond HLA-DR3. Scand J Rheumatol 2005;34:114–21.
21. Wu J, Metz C, et al. A novel polymorphic CAAT/enhancer-binding protein beta element in the FasL gene promoter alters Fas ligand expression: a candidate background gene in African American systemic lupus erythematosus patients. J Immunol 2003;170:132–8.
22. Mehrian R, Quismorio FP Jr, et al. Synergistic effect between IL-10 and bcl-2 genotypes in determining susceptibility to systemic lupus erythematosus. Arthritis Rheum 1998;41:596–602.
23. Ballestar E, Esteller M, et al. The epigenetic face of systemic lupus erythematosus. J Immunol 2006;176:7143–7.
24. Chung SA, Nititham J, et al. Genome-wide assessment of differential DNA methylation associated with autoantibody production in systemic lupus erythematosus. PLoS ONE 2015;10:e0129813.
25. Javierre BM, Fernandez AF, et al. Changes in the pattern of DNA methylation associate with twin discordance in systemic lupus erythematosus. Genome Res 2010;20:170–9.
26. Jeffries MA, Sawalha AH. Epigenetics in systemic lupus erythematosus: leading the way for specific therapeutic agents. Int J Clin Rheumatol 2011;6:423–39.
27. Watad A, Neumann SG, et al. Vitamin D and systemic lupus erythematosus: myth or reality? Isr Med Assoc J 2016;18:177–82.
28. Milner EC, Anolik J, et al. Human innate B cells: a link between host defense and autoimmunity? Springer Semin Immunopathol 2005;26:433–52.
29. Pugh-Bernard AE, Cambier JC. B cell receptor signaling in human systemic lupus erythematosus. Curr Opin Rheumatol 2006;18:451–5.
30. Liu Z, Davidson A. BAFF and selection of autoreactive B cells. Trends Immunol 2011;32:388–94.
31. Grimaldi CM, Hill L, et al. Hormonal modulation of B cell development and repertoire selection. Mol Immunol 2005;42:811–20.
32. Doreau A, Belot A, et al. Interleukin 17 acts in synergy with B cell-activating factor to influence B cell biology and the pathophysiology of systemic lupus erythematosus. Nat Immunol 2009;10:778–85.
33. Lande R, Ganguly D, et al. Neutrophils activate plasmacytoid dendritic cells by releasing self-DNA-peptide complexes in systemic lupus erythematosus. Sci Transl Med 2011;3:73ra19.
34. Garcia-Romo GS, Caielli S, et al. Netting neutrophils are major inducers of type I IFN production in pediatric systemic lupus erythematosus. Sci Transl Med 2011;3:73ra20.
35. Bennett L, Palucka AK, et al. Interferon and granulopoiesis signatures in systemic lupus erythematosus blood. J Exp Med 2003;197:711–23.
36. Kis-Toth K, Tsokos GC. Dendritic cell function in lupus: independent contributors or victims of aberrant immune regulation. Autoimmunity 2010;43:121–30.
37. Crispin JC, Kyttaris VC, et al. T cells as therapeutic targets in SLE. Nat Rev Rheumatol 2010;6:317–25.
38. Scheinecker C, Bonelli M, et al. Pathogenetic aspects of systemic lupus erythematosus with an emphasis on regulatory T cells. J Autoimmun 2010;35:269–75.
39. Ohl K, Tenbrock K. Regulatory T cells in systemic lupus erythematosus. Eur J Immunol 2015;45:344–55.
40. McMurray RW. Bromocriptine in rheumatic and autoimmune diseases. Semin Arthritis Rheum 2001;31:21–32.
41. Zhang W, Wu K, et al. Transforming growth factor beta 1 plays an important role in inducing CD4(+)CD25(+)forhead box P3(+) regulatory T cells by mast cells. Clin Exp Immunol 2010;161:490–6.
42. Cunningham M, Gilkeson G. Estrogen receptors in immunity and autoimmunity. Clin Rev Allergy Immunol 2011;40:66–73.
43. Mu Q, Zhang H, et al. SLE: Another autoimmune disorder influenced by microbes and diet? Front Immunol 2015;6:608.
44. Zoma A. Musculoskeletal involvement in systemic lupus erythematosus. Lupus 2004;13:851–3.
45. Urowitz MB, Gladman DD, et al. American College of Rheumatology criteria at inception, and accrual over 5 years in the SLICC inception cohort. J Rheumatol 2014;41:875–80.
46. Grossman JM. Lupus arthritis. Best Pract Res Clin Rheumatol 2009;23:495–506.
47. Aranow C, Zelicof S, et al. Clinically occult avascular necrosis of the hip in systemic lupus erythematosus. J Rheumatol 1997;24:2318–22.
48. Tebbe B. Clinical course and prognosis of cutaneous lupus erythematosus. Clin Dermatol 2004;22:121–4.
49. Kirchhof MG, Dutz JP. The immunopathology of cutaneous lupus erythematosus. Rheum Dis Clin North Am 2014;40:455–74, viii.
50. Orteu CH, Buchanan JA, et al. Systemic lupus erythematosus presenting with oral mucosal lesions: easily missed? Br J Dermatol 2001;144:1219–23.
51. Ebert EC, Hagspiel KD. Gastrointestinal and hepatic manifestations of systemic lupus erythematosus. J Clin Gastroenterol 2011;45:436–41.
52. Yuan S, Ye Y, et al. Lupus mesenteric vasculitis: clinical features and associated factors for the recurrence and prognosis of disease. Semin Arthritis Rheum 2014;43:759–66.
53. Zhang L, Xu D, et al. Clinical features, morbidity, and risk factors of intestinal pseudo-obstruction in systemic lupus erythematosus: a retrospective case-control study. J Rheumatol 2016;43:559–64.
54. Mackay IR. Historical reflections on autoimmune hepatitis. World J Gastroenterol 2008;14:3292–300.
55. Mittoo S, Fell CD. Pulmonary manifestations of systemic lupus erythematosus. Semin Respir Crit Care Med 2014;35:249–54.
56. Miner JJ, Kim AH. Cardiac manifestations of systemic lupus erythematosus. Rheum Dis Clin North Am 2014;40:51–60.
57. Manzi S, Meilahn EN, et al. Age-specific incidence rates of myocardial infarction and angina in women with systemic lupus erythematosus: comparison with the Framingham Study. Am J Epidemiol 1997;145:408–15.
58. Sesin CA, Yin X, et al. Shedding of endothelial protein C receptor contributes to vasculopathy and renal injury in lupus: in vivo and in vitro evidence. Kidney Int 2005;68:110–20.
59. Weening JJ, D'Agati VD, et al. The classification of glomerulonephritis in systemic lupus erythematosus revisited. J Am Soc Nephrol 2004;15:241–50.
60. Davidson A, Aranow C. Lupus nephritis: lessons from murine models. Nat Rev Rheumatol 2010;6:13–20.
61. Newman K, Owlia MB, et al. Management of immune cytopenias in patients with systemic lupus erythematosus—old and new. Autoimmun Rev 2013;12:784–91.
62. Gonzalez-Naranjo LA, Betancur OM, et al. Features associated with hematologic abnormalities and their impact in patients with systemic lupus erythematosus: Data from a multiethnic Latin American cohort. Semin Arthritis Rheum 2016;45:675–83.
63. Sultan SM, Begum S, et al. Prevalence, patterns of disease and outcome in patients with systemic lupus erythematosus who develop severe haematological problems. Rheumatology (Oxford) 2003;42:230–4.
64. Carli L, Tani C, et al. Leukopenia, lymphopenia, and neutropenia in systemic lupus erythematosus: Prevalence and clinical impact—A systematic literature review. Semin Arthritis Rheum 2015;45:190–4.
65. Matsuyama W, Yamamoto M, et al. TNF-related apoptosis-inducing ligand is involved in neutropenia of systemic lupus erythematosus. Blood 2004;104:184–91.
66. Michel M, Lee K, et al. Platelet autoantibodies and lupus-associated thrombocytopenia. Br J Haematol 2002;119:354–8.
67. Mackay M. Lupus brain fog: a biologic perspective on cognitive impairment, depression, and fatigue in systemic lupus erythematosus. Immunol Res 2015;63:26–37.

68. Jeltsch-David H, Muller S. Autoimmunity, neuroinflammation, pathogen load: A decisive crosstalk in neuropsychiatric SLE. J Autoimmun 2016;74:13–26.
69. Abbott NJ, Mendonca LL, et al. The blood-brain barrier in systemic lupus erythematosus. Lupus 2003;12:908–15.
70. Stock AD, Wen J, et al. Neuropsychiatric lupus, the blood brain barrier, and the TWEAK/Fn14 pathway. Front Immunol 2013;4:484.
71. O'Carroll SJ, Kho DT, et al. Pro-inflammatory TNFalpha and IL-1beta differentially regulate the inflammatory phenotype of brain microvascular endothelial cells. J Neuroinflammation 2015;12:131.
72. Zardi EM, Taccone A, et al. Neuropsychiatric systemic lupus erythematosus: tools for the diagnosis. Autoimmun Rev 2014;13:831–9.
73. Jeltsch-David H, Muller S. Neuropsychiatric systemic lupus erythematosus: pathogenesis and biomarkers. Nat Rev Neurol 2014;10:579–96.
74. Kaplan MJ, Deng C, et al. DNA methylation in the regulation of T cell LFA-1 expression. Immunol Invest 2000;29:411–25.
75. van Vollenhoven RF, Mosca M, et al. Treat-to-target in systemic lupus erythematosus: recommendations from an international task force. Ann Rheum Dis 2014;73:958–67.
76. Yates CR, Krynetski EY, et al. Molecular diagnosis of thiopurine S-methyltransferase deficiency: genetic basis for azathioprine and mercaptopurine intolerance. Ann Intern Med 1997;126:608–14.
77. Takada K, Arefayene M, et al. Cytochrome P450 pharmacogenetics as a predictor of toxicity and clinical response to pulse cyclophosphamide in lupus nephritis. Arthritis Rheum 2004;50:2202–10.

MULTIPLE-CHOICE QUESTIONS

1. Mortality in systemic lupus erythematosus (SLE)
   A. Has been stable for the past 4 decades
   B. Is greater in Caucasians than in other racial groups
   C. Is secondary to both the disease and the therapy

2. Autoantibodies in SLE
   A. Are an epiphenomenon
   B. Are generally tissue specific
   C. Form TLR-activating immune complexes

3. Treatment in SLE
   A. Relies on immunosuppression
   B. Is different for different races
   C. Has minimal toxicity

52

Rheumatoid Arthritis

Andrew P. Cope

Rheumatoid arthritis (RA) is one of the most common chronic inflammatory diseases and in modern times has become a prototype disease entity for defining the molecular and pathological basis of chronic inflammatory syndromes. The term *rheumatoid arthritis* was coined by Garrod in 1859. However, this was probably an inappropriate use of the term because it encompassed polyarticular osteoarthritis as well as inflammatory polyarthritis. In spite of references to inflammatory afflictions of joints by the likes of Galen, Sydenham, and Heberden, the first convincing case reports of the disease, described in terms that would be recognizable today, were published in 1800 by Landré-Beauvais, who labeled the disease "*la goutte asthénique primitive.*"[1] This description was distinct because all patients were female, an observation that was significant when the most important differential diagnosis at that time was polyarticular gout, a disease predominantly of males.

Today we recognize RA as a chronic inflammatory disorder of joints of unknown etiology in which the major target tissue is the synovial lining of joints, bursae, and tendon sheaths.[2] Although traditionally considered an autoimmune disease, RA differs from organ-specific autoimmune disease entities in several respects. From the outset of clinically apparent disease, the systemic immuno-inflammatory process, driven by cytokines and other inflammatory mediators, promotes the activation and proliferation of stromal joint tissues, in particular the fibroblastic synovial lining layer. This appears to contrast with organ-specific autoimmune diseases, such as type 1 diabetes or autoimmune thyroiditis, characterized by an antigen-driven immune-inflammatory response *in situ* leading to targeted cellular destruction of autoantigen-expressing pancreatic β-islet or thyroid tissue cells. In RA, once the inflammatory process is established, the inflammatory synovium, or pannus, may invade and erode underlying cartilage and bone. Unlike autoimmune diseases that target single organs or tissues, RA is a systemic inflammatory disease that likely encompasses a heterogeneous syndrome with marked variation in clinical expression that most clinicians today would acknowledge is more than one disease entity. Indeed, it is now apparent that the disease is heterogeneous not only clinically but also pathologically, serologically, and genetically.

EPIDEMIOLOGY

The incidence of RA (the rate of new cases arising in a given period) is 0.1–0.2 per 1000 of the population for males and 0.2–0.4 per 1000 for females.[3] Rates as high as 0.4 per 1000 have been reported when the 2010 American College of Rheumatology/European League Against Rheumatism (ACR/EULAR) classification criteria are applied. These rates plateau between the ages of 45 and 75 years in some series but can increase steadily with age until the seventh decade, declining thereafter. The largest difference in incidence between the sexes occurs in those under 50 years of age. Past evidence indicates that the incidence may be in decline, at least for seropositive disease. Large cross-sectional population samples indicate that disease prevalence, which ideally should include all past and inactive cases, ranges from 0.5–2% for Caucasian European and North American populations over the age of 15, with a female to male excess of 2–4 times.[3] Despite similar prevalence estimates for these geographically diverse populations, greater diversity has been documented for rural African populations, where the prevalence has been reported to be as low as 0.1%, and for Native Americans (including the Pima, Yakima, and Chippewa tribes), where the prevalence may be as high as 5%. Such variance across geographical borders likely reflects distinct environmental factors and sociodemographic determinants as well as a spectrum of genetic admixture. Lifetime risk has been estimated at 4% for women and 3% for men.

Complex polygenic autoimmune syndromes like RA are diseases of low penetrance, where thresholds of disease expression may be higher in males. This is based in part on the observation that the increased risk of disease in siblings of probands (denoted λs) is rather small; the λs for RA (risk to sibling divided by population risk) are estimated to be 3, whereas the λs for systemic lupus erythematosus (SLE) may be as high as 30; for celiac disease the λs are closer to 50. Recent insights into familial clustering indicate that family history remains a strong independent risk factor for RA but that this risk does not differ by gender.[4] This implies that nongenetic factors influence gender bias. Twin studies also provide compelling evidence for genetic effects, given the excess concordance rates for monozygotic (12–15%) compared with dizygotic twins (<5%, and probably nearer to 3.5%).[4] These concordance rates appear somewhat low, but compared with a background prevalence of 1% in outbred populations, genetic epidemiology studies report heritability estimates of 68% for those patients who carry antibodies to citrullinated protein antigens (ACPAs$^+$) and 66% for those who do not (ACPAs$^-$), indicating substantial genetic influence.[5] Human leukocyte antigen (HLA) is thought to contribute 35–40% of this value. This "missing heritability" leaves a substantial contribution to disease susceptibility from environmental factors, influenced by occupation, socioeconomic status, exposure to infectious pathogens, and lifestyle factors—a conglomeration of factors termed the "exposome."

KEY CONCEPTS

Important Risk Factors for Developing Rheumatoid Arthritis

- Female gender; impact of X chromosome, microchimerism, lifestyle
- Age; associated with accelerated immune aging
- Inheritance of genetic variants, *e.g., HLADRB1* and *PTPN22*
- Autoantibodies to citrullinated protein antigens (ACPAs), rheumatoid factor
- Family history; first-degree relatives have higher prevalence of genetic and serological risk factors
- Hormonal factors; nulliparity, the first 3 months postpartum, low androgen or high estrogen status (in males); longer-duration breastfeeding
- Smoking status; >25 cigarettes/day for >20 years confers a 15-fold risk in subjects who carry disease associated human leukocyte antigen (HLA)-DRB1 alleles
- Low alcohol intake
- Environmental antigens (the "exposome"); dietary factors; exposure to infectious (and noninfectious/microbiota) pathogens at mucosal surfaces such as the lung, periodontium, and gut; noninherited maternal antigens (NIMA)
- Urban dwelling

Two of the more intriguing factors contributing to disease occurrence are age and gender. Age-associated changes in susceptibility to infection, neoplastic disease, and autoimmunity suggest that a common mechanism could be responsible. Immune senescence is one possibility, where age-related decline in host immunity is characterized at the cellular and molecular level by dramatic expansions of lymphocyte clones, corresponding contractions of the naïve T- and B-cell repertoires, and telomere erosion of leukocytes, indicative of an extensive proliferative history.[6] When combined with oxidative stress and a range of biochemical derangements of pathways integral to antigen responsiveness and immune regulation, these factors may combine to (i) increase susceptibility to foreign pathogens, (ii) augment reactivity to self-tissue antigens (which may be modified post-translationally by the aging process), and (iii) generate a repertoire of lymphocytes defective in terms of tumor surveillance. Thus immune senescence could be considered a risk factor for RA.

The female sex preponderance implies that hormonal and reproductive factors strongly influence risk. On the one hand, nulliparity is a risk factor for RA. Women entering the first 3 months of the postpartum period are also at increased risk. By contrast, oral contraceptive use, pregnancy, and hormonal replacement therapy have all been associated with reduced risk or less severe disease,[7] whereas extended periods of breastfeeding appear to increase risk. Several possible immune mechanisms have been proposed. For example, maternofetal mismatch at the major histocompatibility complex (MHC) locus has been associated with higher disease remission rates during pregnancy. In contrast, noninherited HLA-DRB1 DERAA-positive (occupying the same amino acid positions as the QQRAA/QKRAA/KKRAA shared epitope motif) maternal antigens (NIMA) expressed on maternal cells constitute an environmental factor that is protective to a DERAA-negative child, perhaps through mechanisms relating to microchimerism. In mouse models pregnancy has been associated with quantitative changes in the numbers of regulatory T cells (Tregs). An influence of hormonal factors is further suggested in studies of men where disease is associated with lower androgenic testosterone and dehydroepiandrosterone (DHEA) levels, and increased estradiol, compared with healthy control male subjects.

ETIOLOGY AND PATHOGENESIS

Environmental and Nongenetic Factors

Most, but not all, studies have reported an association between RA and smoking. One of the largest studies comprising over 370 000 women from Women's Health Cohort Study reported a relative risk of 1.4 for women who smoked more than 25 cigarettes per day for more than 20 years, compared with nonsmokers.[8] The association appears to be more closely related to duration than to the amount of tobacco exposure, with smoking status being a risk factor for older age of RA onset;[9] smoking also may influence severity because smokers are more likely to have seropositive, erosive disease with extraarticular manifestations. The effect may be fully reversible in those individuals who stop smoking for 10 years or more. Further evidence of gene-environment interactions with respect to smoking have been documented in a population-based case–control study of Swedish RA patients.[10] In this study the relative risk of developing rheumatoid factor (RF) positive (RF^+) RA was calculated according to smoking status and HLA-DRB1 genotype. The relative risk of developing RA increased from 2.5 in nonsmokers with disease-associated HLA-DRB1 genes to 7.5 and 15.7 in smokers who carried one or two copies of the susceptibility alleles, respectively. Follow-up gene-environment interaction studies demonstrate robust associations between heavy smoking, HLA-DRB1 alleles encoding specific amino acids at positions 11 and 13, and the presence of ACPAs. Strongest associations were with antibodies to citrullinated α and β chains of fibrinogen (epitopes Fibα580–600 and Fibβ36–52) and α-enolase (CEP-1); IgA ACPAs were especially prevalent in smokers.

Being female, a smoker, and carrying specific disease-associated genetic variants may be necessary but not sufficient to initiate chronic inflammatory arthritis. Other environmental triggers may be involved. Not least among these is exposure to foreign pathogens. This association has gained credibility because of the presumed link between not only infection and autoimmunity but also immunodeficiency and autoimmune disease. Nonetheless, no single pathogen or group of pathogens has been defined. This could imply that aberrant host responses (either exaggerated innate inflammatory responses or failure to terminate such responses) may arise after a wide range of infectious insults. Indeed, bacterial products including superantigens, mycoplasma species, viruses (including herpes family, parvovirus, and retroviruses), and fungi have all been implicated, but data are insufficient to prove causation. Epstein-Barr virus (EBV) infection is common, antibodies to EBV nuclear antigens have been reported in patients with RA, EBV is a polyclonal activator of B lymphocytes, EBV-specific T cells reactive to EBV gp110 have been identified in RA synovial joints, and EBV RNA has been isolated from the synovium. There exists a tantalizing link between infection with *Porphyromonas gingivalis,* which expresses its own enzymatic machinery for generating bacterial or host-derived citrullinated proteins, severe periodontitis (which shares risk factors for RA), and RA.[11] According to this model, molecular mimicry is underpinned by autoimmune responses to bacterial-derived proteins such as enolase and cross-reactivity to the human orthologue α-enolase, expressed in abundance by myeloid cells in inflamed periodontal and synovial tissues.

A comparison of the microbial genomes from the small intestine and colon of mice housed in conventional versus germ-free facilities suggests that a single gut-residing species, in this case segmented filamentous bacteria, can profoundly influence the clinical expression of inflammatory arthritis in mouse models of autoimmune arthritis. In the K/B x N serum transfer model of arthritis, segmented filamentous bacteria enhanced generation of interleukin (IL)-17 expressing T cells in lamina propria.[12] Together with rapid advances in sequencing technologies, these data have prompted a systematic analysis of the symbiont microbial communities in patients with RA in comparison to those derived from healthy control populations. Microbiota are attractive environmental risk factors because they are acquired at around the time of birth and are modified by diet as well as the host genome. Studies from several groups have demonstrated irrefutable evidence of dysbiosis, with distinct patterns of microbiota depending on the stage of disease and the population studied. For example, the first US study reported enrichment of *Prevotella spp.* and *Bacteroides spp.*, gram-negative anaerobes in the gut of patients with early but not established RA. Metagenomic sequencing of oral and fecal microbiota from a cohort of Chinese RA patients revealed enrichment of *Lactobacillus*, particularly in severe disease, and depletion of *Haemophilus* species.[13] Interestingly, microbiota signatures were predictive of response to therapy, suggesting that their composition may be prognostic as well as indicative of the stage of disease. The prospect of fecal transplantation to prevent or treat established disease is an intriguing one.

Immunogenetics

RA is a clinically heterogeneous disease, and so comprehensive identification of disease susceptibility genes has been challenging, in spite of heritability estimates of at least 60%. With the exception of the MHC, where extensive gene polymorphism contributes about one-third of genetic susceptibility, and *PTPN22* (odds ratio 1.8), individual genetic variants confer low to moderate risk and have low penetrance (odds ratios in the range 1.1–1.5). Numerous genome-wide linkage scans of multiplex families with RA have established and confirmed an important contribution of the MHC (Chapter 5). This lends support to a wealth of epidemiological and genetic data describing associations between RA and specific HLA-DRB1 alleles, in particular HLA-DR4 subtypes. Although this association was first described by Stastny in the 1970s, it was shown more than a decade later that susceptibility to RA across different ethnic populations correlated closely with the expression of a specific consensus amino acid sequence (referred to as the "shared epitope," hereafter SE) within the HLA-DR β chain α-helix (Fig. 52.1).[14] This sequence was subsequently shown by several groups of investigators to be encoded by HLA-DRB1 alleles, including HLA-DR4 (*04:01, *04:04, *04:05, and *04:08), but also HLA-DR1 (*01:01), DR6 (*14:02), and DR10 (*10:01) alleles, among others. HLA-DR9

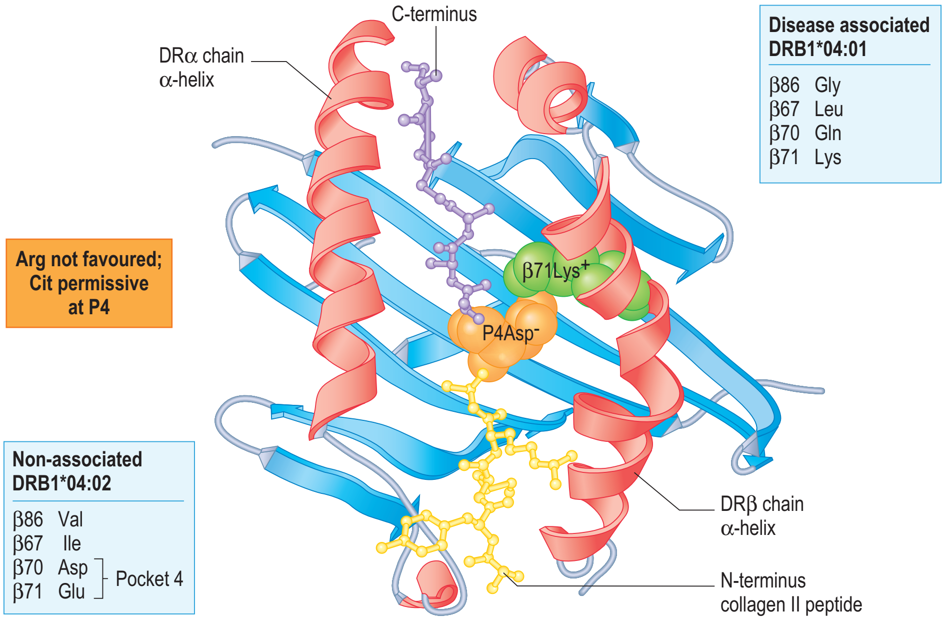


FIG 52.1 Crystal Structure of a Collagen II Peptide/Human Leukocyte Antigen (HLA)-DR4 Complex. Ribbon model of an immunodominant collagen II peptide (1168–1180) complexed to HLA-DR4 (DRA*01:01/DRB1*04:01); a view of the major histocompatibility complex/peptide complex as seen from the T-cell surface. DRα and DRβ chain helices are shown in red, whereas the β-pleated sheet comprising the floor of the peptide binding groove is shown in blue. Residues 67 to 74 of the DRβ chain, components of the third hypervariable region, derive the "shared epitope." The ball and stick model of the CII peptide is shown. Interacting residues of the peptide position 4 (Asp, orange) and DRβ chain residue (β71Lys, green), which make up part of pocket 4, are depicted as van der Waal's spheres. Differences in amino acid sequence between the closely related disease-associated DRB1*04:01 and nonassociated DRB1*04:02 gene products are illustrated. Note that although Arg would not be favored at position 4 in the peptide, modification of Arg → Cit by deimination would be permissive. Figure generated by R. Visse and A. Cope, based on crystal data derived by Wiley and colleagues (Courtesy Dessen et al, Immunity 1997;7:473.)

(*09) is also associated, but not in Caucasians. According to fine-mapping data, RA risk is linked not only to amino acids encoded by the RA SE sequence (71 and 74 positions of the alpha helix of the DRβ chain) but also to amino acid positions 11 and 13, located in the peptide binding groove of the HLA-DR heterodimer.[15] Single amino acid variants have also been defined in HLA-B (position 9) and HLA-DPβ (position 9). Differences in odds ratios associated with these five amino acid variants between ACPAs-positive and ACPAs-negative disease provide further evidence that seropositive and seronegative disease are genetically distinct. On the other hand, the finding of DERAA encoding HLA-DRB1 alleles (DRB1*01:03, *04:02, *11:02, *11:03, *13:01, *13:02, and *13:04), conferring lower disease risk, and more importantly the finding of reduced radiographic progression in RA even in the presence of one copy of the SE, raises the possibility that specific subsets of MHC class II genes may confer an independent protective role.

Specific genotypes cosegregate with distinct clinical features. For example, in population-based studies, different HLA-DRB1 alleles influence the severity of disease (including radiographic progression), with DRB1*04:01 being found in patients with severe, seropositive, erosive RA (often with extraarticular features such as vasculitis and Felty syndrome in *04:01 homozygous or *04:01/*04:04 compound homozygote individuals). Valine at position 11 of HLA-DRB1 confers the strongest association with radiological damage and highest all-cause mortality.[16] Statistical modeling points to HLA genetic polymorphism being associated with young-onset RA.[9] DRB1*01:01 and *10:01 are observed at a higher frequency in patients with less severe, seronegative, nonerosive disease. Inheriting two copies of alleles expressing the consensus sequence increases disease penetrance, time of onset, and severity. Thus these genetic associations manifest as clinically distinct disease entities.

On the basis of early observations, two principal models were proposed to account for the association between RA and the consensus DR β-chain sequence. Both were based on the assumption that the SE is the critical genetic element linked directly to disease. The first model proposed that the SE determines specific peptide binding and that "pathogenic" peptides bind only to disease-associated HLA class II molecules (Fig. 52.1). This model predicted that a gradient of affinities of disease inducing peptide for MHC class II molecules might account for the differences in susceptibility and/or severity conferred by different HLA-DR molecules. Along the same lines, disease-associated alleles may preclude the binding of peptides required for the generation of naturally occurring Tregs specific for self-peptide antigens. The second model proposed that the SE influences T-cell receptor (TCR) recognition by binding and selecting autoreactive T cells during thymic maturation and expanding these populations in the peripheral compartment; again perturbations of a repertoire of Tregs could arise through opposing influences of the SE sequence.

Two recent lines of experimental evidence provide further insights into HLA-disease associations. The first is the association between HLA-DR4 subtypes and telomere erosion in RA patients and healthy donors, detectable before the age of 20.[6] Telomere loss has been associated with aging and immune senescence. Whether this implies a direct contribution of HLA-DR4 expression to accelerated differentiation of hematopoietic cells in the context of a chronic inflammatory process is intriguing, but the relationship certainly warrants further investigation. The second line of evidence pointing to specific functions of SE^+ alleles has arisen through analysis of autoantibodies in RA patients typed at the HLA-DRB1 locus.[17] These studies, replicated in European and US RA cohorts, demonstrated associations between SE frequencies and antibodies to cyclic citrullinated peptides (anti-CCPs), as as opposed to rheumatoid factor (RF). Specifically, compared with healthy controls, the odds ratio for the association between one or two copies of the SE and anti-CCPs positivity was 4.4 and 11.8, respectively. These findings point to SE associating not with RA *per se,* but with a clinical phenotype, in this case autoantibodies to modified proteins.[17] An additional pathogenetic link between HLA and citrullination is supported by the observation that peptides carrying nonpolar residues such as citrulline (neutral), as opposed to arginine (positive charge), in pocket four of the HLA-DRαβ binding groove (see Fig. 52.1) convert a poorly binding peptide to a higher-affinity epitope that induces robust T- and B-cell responses *in vivo.* This notion is strongly supported by crystal structures of HLA-DRαβ (DRB1*04:01 and *04:04) complexed to peptide.[18] Here amino acids 11, 13, 71, and 74 form part of the fourth anchoring peptide binding pocket, the same positions that confer highest risk of disease, as noted above. Indeed, citrullinated human fibrinogen, but not the nonderivatized protein, induces chronic inflammatory arthritis in HLA-DR4 transgenic mice but not their nontransgenic C57BL/6 littermates. The importance of citrulline as a molecular feature of human T-cell–specific autoantigenic determinants is further suggested through identification of autoreactive $CD4^+$ T lymphocytes by flow cytometry using HLA-citrullinated-peptide tetramer complexes.[18] These tools also permit determination of the relative proportions of antigen-reactive effector and regulatory T-cell subsets.

Genome-wide association studies (GWASs) of large patient cohorts, stratified by autoantibody serology, have contributed significantly to our understanding of RA genetics. Meta-analyses and fine-mapping studies, including custom-designed SNP arrays covering nearly 130 000 genetic variants selected from 186 loci-poly defined in GWASs of autoimmune and immune-mediated inflammatory diseases (the Immunochip), have identified more than 100 susceptibility loci, many of which have been validated in RA populations of diverse ancestry.[19] Among the strongest associations outside HLA is *PTPN22,* initially identified in candidate gene association studies. The *PTPN22* gene variant confers risk with odds ratios ranging from 1.5 to 1.9 in Caucasian and European populations, and it is unusual among most genetic variants in that it is a coding frame mutation (R620W). *PTPN22* encodes a hematopoietic cytoplasmic protein tyrosine phosphatase whose substrates include Src and Syk family kinases, CD3ε, TCRζ, and signaling intermediates such as Vav.[20] It comes as no surprise that the *PTPN22* mutation, as a negative regulator of antigen receptor and integrin signaling, has also been described in type 1 diabetes, a subset of SLE patients, oligoarticular juvenile idiopathic arthritis, vitiligo, Addison disease, and autoimmune thyroid disease but not in multiple sclerosis or psoriasis. The link with autoimmune diseases associated with pathogenic autoantibodies is also intriguing. Whether the PTPN22-R620W variant is a loss- or gain-of-function mutant in man remains somewhat controversial. However, studies of Ptpn22-deficient mice, and mice expressing the R620W mouse orthologue Ptpn22-R19W, share features in T, B, and myeloid lineage indicative of loss of function for the R619W mutation, associated with increased antigen receptor signaling.[21] The relative contributions of the mutant to perturbationsin central and peripheral tolerance mechanisms remain to be determined.

Scrutiny of other, non-MHC susceptibility loci point to genes whose products are involved in proximal signaling pathways that regulate T-cell activation, differentiation, and persistence. Besides *HLA* and *PADI4*, which influence the molecular determinants of T-cell "input signals," these include *CD28*, *CTLA-4*, and *CD2-CD58* (regulation of T-cell costimulation); *CD247*, *PTPN22*, *PRKCQ*, *TAGAP*, and *REL* (transducer modules of TCR signaling); *STAT4* and *TNFRSF14* (inducers of lineage-specific cytokine gene expression and persistence of memory T cells); and *REL*, *IL2-IL-21*, *IL2RA*, and *IL2RB* (regulators of IL-2 gene expression and IL-2R signaling). Notable overlap with these and other allelic variants has been reported in other autoimmune diseases, indicating that susceptibility to RA is linked to fundamental perturbations of immune tolerance. Variants mapping to cell surface receptors (*IL6R*, *CCR6*, *CD40*, *CD5*, *FCGR2A*) and intracellular signaling intermediates (*TNFAIP3*, *TYK2*, *TRAF1*, *TRAF6*, *RASGRP1*, *BLK*) are well represented, as are transcription factors linked to cell differentiation and effector function (*GATA3*, *IRF5*, *IRF8*, *IKZF3*, *RBPJ*, *RUNX1*). Perturbations in expression or function of these genes will influence the function of a broad range of immune cell subsets.

Synovial Pathology

RA targets diarthrodial joints, structures characterized by hyaline cartilage lining opposing articulating surfaces and a cavity of viscous synovial fluid lined by synovial membrane lacking a basement membrane, but encased by a fibrous joint capsule. Normal synovial tissue comprises a lining layer, no more than a few cells in depth, of stromal fibroblast-like synoviocytes (FLSs—also known as type B synoviocytes) and sublining macrophages (type A synoviocytes). The synovium serves to line noncartilaginous surfaces, and although blood vessels are sparse, it functions to provide essential nutrients to avascular structures including cartilage, tendons, and bursae. Recent insights from synovial biopsies of "at-risk" subjects who have both joint pain (arthralgia) and serum RA-associated autoantibodies (*i.e.*, systemic autoimmunity) suggest that the transition from "normal" to inflamed synovium may occur over a period of only a few weeks or months.[22] Subtle T-cell infiltration was detected in those subjects who went on to develop clinically apparent synovitis.

Increased Vascularity and Cell Migration

The range of pathology observed in patients with RA perhaps most convincingly underscores the heterogeneity of the disease.[22] The earliest changes observed relate to increases in vascularity characterized by vascular congestion and thrombosis with obliteration of small vessels in association with perivascular inflammatory infiltrates. Hyperplasia of the synovial lining layer is another typical early finding. These changes are rather non-specific and certainly not diagnostic.

A key checkpoint defining the switch from acute to chronic persistent inflammation is the sustained activation of microvascular endothelium, phenotypic changes in the high endothelial venules (reminiscent of tissue injury), and the concomitant upregulation of adhesion molecules such as intercellular adhesion molecule (ICAM)-1 and vascular cell adhesion molecule (VCAM)-1 (Chapter 11). According to current thinking, the expression of chemoattractants derived from synovial stromal cells heralds the rolling, adhesion, and transmigration of mononuclear cells through endothelial barriers into the synovial membrane. It also contributes to the progressive synovial hypertrophy and hyperplasia, sometimes with villous-like projections more typical of chronic, established inflammation. Intravital imaging of synovial joints of mice injected with arthritogenic antibodies derived from the serum of *K/B* × *N* mice indicates that enhanced vasopermeability at sites destined to become arthritic is a crucial early event, at least in antibody-induced disease. This process is dependent on mast cells and neutrophils and on the release of the vasoactive amines histamine and serotonin, which contribute to heightened vascular permeability.

Neovascularization further promotes influx of inflammatory cells. To what extent this is driven by the hypoxic environment is not entirely clear, but angiogenic growth factors such as vascular endothelial growth factor (VEGF), angiopoietin, Tie-2, and hypoxia inducible factor (HIF), as well as lymphangiogenic factors VEGF-C and VEGF-R3, make important contributions. It has been proposed that influx of inflammatory lymphocytes and cells of monocytic lineage far outweighs the egress of cells from synovial tissue, likely due to chemokine gradients and engagement of sphingosine 1 phosphate receptor 1 (S1P1) by CD69 expressed on activated lymphocytes. This leads to downregulation of S1P1 and retention of cells in tissues. Once in the synovium, egress may be promoted by specific mediators such as sphingosine 1 phosphate, which actively regulates cellular egress, or blocked through integrin/adhesion molecule interactions, *e.g.*, between antigen-specific T cells and dendritic cells/APC, through TCR-dependent "stop" signals, or perhaps through highly selective and specific inactivation of chemokines by proteolysis, *e.g.*, stromal derived factor (SDF)-1 cleavage by cell surface dipeptidyl peptidase CD26. The abundance of lymphatic vessels in inflamed synovium, suggested by expression of podoplanin and CD31 *in situ*, suggests that active lymphangiogenesis exists that may promote efflux of cells and fluid from the synovium.[22] Interestingly, enhanced proliferation of lymphatic vessels and increased lymphatic drainage have been documented after tumor necrosis factor (TNF) blockade. Lymphatic growth factors such as VEGF-C (which signal through VEGFR-3/Flt-4) have been implicated in promoting this joint-protective effect.

Organization of Lymphoid Tertiary Microstructures

Tissue microstructure both dictates and facilitates immune responses in secondary lymphoid organs and mucosa-associated lymphoreticular tissues (MALTs). These structures have evolved to coordinate responses to pathogens and to direct lymphocyte recirculation, and although their role in immune homeostasis is established, quite how they contribute to pathological states is less well understood. Thus the inflamed synovium appears to be uniquely suited to supporting distinct patterns of cellular infiltrates—noncapsulated, as opposed to capsulated—inducible lymphoid structures, that could play a role in pathways of cell activation, differentiation, and survival.[22] These include diffuse, rather disorganized lymphocytic infiltrates which comprise the most common form of synovitis, occurring in ~30% in prospective cohort studies; up to 70% has been described in late-stage disease (at arthroscopy, joint replacement surgery). In 40–50% of patients more organized follicular structures may exist (Fig. 52.2). Based on immunohistochemical analysis, approximately 25% of these follicular structures include organized germinal centers in which there are zones of proliferating B cells with affinity maturation, in addition to a distinct T-cell zone. In aggregates lacking germinal centers, follicular dendritic cells (DCs) are absent.

A fourth histological pattern, characterized by granulomatous reactions, has been described in a much smaller subset of patients. The cellular and molecular determinants of these structures include the homeostatic lymphoid chemokines CXCL13 and the CCR7 ligands CCL21 and CCL19, VCAM-1$^+$ICAM-1$^+$LTβR$^+$ mesenchymal

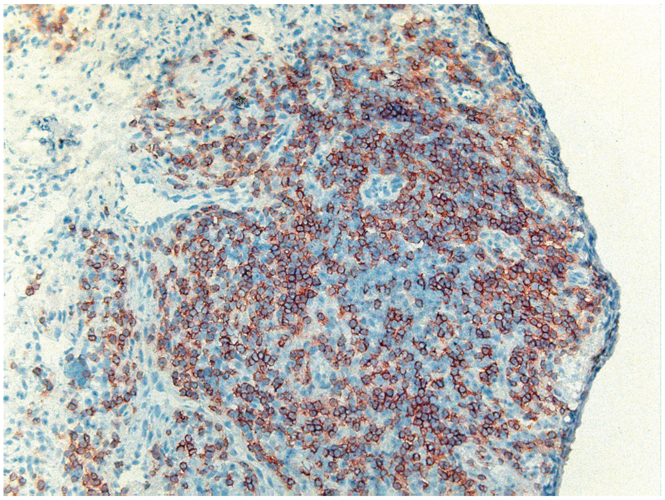


FIG 52.2 Lymphoid Follicular Structures in Inflamed Rheumatoid Arthritis Synovial Tissue. A characteristic hematoxylin and eosin–stained tissue section from a patient with active RA showing a large follicular-like structure (original magnification ×100). This section is also stained with monoclonal antibodies to CD3ε, followed by a three-step immunoperoxidase staining protocol ($CD3^+$ T cells stained dark red). (Courtesy Tak, et al. Arthritis Rheum 1996;39:1077, with permission from J. Wiley and Sons, Inc.)

organizer cells, and hematopoietic-derived $CD3^-CD4^+IL\text{-}7R^+RANK^+$ lymphoid inducer cells. Lymphoid tissue inducer cells produce LT-β, which is required for high endothelial venule differentiation, amplification of chemokine expression, and development of the stromal architecture. The mechanisms that sustain these lymphoid structures over time are not fully characterized but could involve persistence of antigen and augmented antigen-presenting function through the differentiation of follicular DCs, localized and autonomous expression of cytokines and growth factors, chemokine gradients favoring cellular interactions, and dysregulation of leukocyte homing and/or egress. Documentation of class-switch recombination and somatic mutation of immunoglobulin genes *in situ* is further evidence of active adaptive immunity within synovial lymphoid follicles, and suggests that supporting *in situ* production of pathogenic antibodies is of considerable pathobiological importance.

Cross-sectional studies have indicated that, from a clinical standpoint, these distinct patterns reflect not only the heterogeneity of the disease but also a spectrum of disease severity, with each pattern requiring a distinct targeted approach to therapy. More recent prospective studies, however, seem to indicate that the presence of lymphoid aggregates is one aspect of a rather dynamic process. Although the lymphoid aggregates are associated with the degree of synovial tissue cellularity and vascularity, they are neither diagnostic for RA nor predictive of radiographic outcome such as joint erosions. The dynamic nature of these lymphoid structures is supported by their dissolution after therapy with rituximab (anti-CD20) and TNF inhibitors, in the latter case enhancing lymphatic flow.

Gene Expression Signatures

Gene array technology has permitted unbiased and systematic analysis of synovial tissue at the whole-genome level and, as with genetic, serological, and histological analysis, has provided further evidence of disease heterogeneity. This strategy assumes that each disease phenotype should be represented in gene expression signatures of multiple genes. Remarkable variation in gene expression signatures has been observed, and initial reports seemed to suggest that these signatures could be divided broadly into two groups.[23] The first is characterized by upregulation of immune response and inflammatory genes; these tissues are enriched for MHC and immunoglobulin gene products. In some tissues this signature may resemble an antiviral response consistent with prior infectious insult and, more specifically, a signal transducer and activator of transcription 1 (STAT1)–dependent gene signature (promoting protective or proinflammatory effects).[23] This has also been observed in peripheral blood lymphocytes (PBL) from patients. The second group is more indicative of tissue remodeling, more reminiscent of that seen in osteoarthritis. Most interestingly, distinct fibroblast-like synoviocytes (FLSs), genotypes, and phenotypes match the type of RA tissue from which they derived. For example, FLS signatures from high-intensity inflammatory tissue resemble signatures of transforming growth factor (TGF)-β/activin-A–induced expression profiles in myofibroblasts, whereas growth factors such as insulin-like growth factor 2 and insulin-like growth factor binding protein 5 characterize FLSs from tissues with low-intensity inflammation. These data support the notion of RA heterogeneity being reflected in FLSs as a stable, if not transformed, trait. They would also be consistent with a model in which there is a transition from an antigen-driven phase of disease to a relatively antigen-independent process, characterized by autonomous, invasive, and aggressive FLSs and manifested as persistence of the inflammatory state. Subsequent disease flares might arise as a consequence of repeated exposure of lymphocytes to antigen after infectious or mechanical insult.

Gene expression signatures have clinical utility. For example, a type I IFN signature in peripheral blood cells negatively predicts the response to B-cell–depleting therapy rituximab, but its detection in neutrophils predicts a good response to anti-tumor necrosis factor (TNF) therapy. Enrichment of TNF-induced genes in early RA synovial biopsies is associated with higher disease activity and predicts poor response to first-line therapy, regardless of the drug therapy, while enrichment of synovial myeloid, but not lymphoid, gene expression signatures at baseline is associated with good clinical responses to TNF blockade at 16 weeks.

Aberrant expression of microRNA (miRNA) in cells or plasma of patients with RA has been widely reported,[24] with circulating miRNA, such as miR-223 and MiR-16, correlating with disease activity. This is of particular relevance because of the unique role for miRNA in regulating expression of multiple gene sets. A recent report indicates that baseline expression of miR-125b in peripheral blood mononuclear cells of drug-naïve early RA patients predicts response to therapy at 3 months.

Immunobiology of RA

Initiation of the Immune Response

Synovial fibroblasts are exquisitely sensitive to inflammatory cytokines such as IL-1, TNF-α, and IL-6. Accumulating evidence indicates that FLSs also express a range of toll-like receptors (TLRs) that can respond to exogenous, pathogen-associated molecular patterns (PAMPs) or indeed a growing range of self-tissue proteins, some of which could be considered damage associated molecular patterns (DAMPS). Endogenous ligands especially relevant to inflammatory arthritis include heat shock proteins, fibrinogen fragments, antibody-DNA complexes, high-mobility group box (HMGB)-1, and hyaluronan oligosaccharides. Recent data suggest that synovial FLSs, as well as synovial macrophages, express TLR2 *in situ.* Expression is upregulated after stimulation with IL-1 and the TLR2 ligand peptidoglycan. TLR2 engagement induces cytokines such as IL-6, matrix metalloproteinases, adhesion molecules, and an array of chemokines including granulocyte chemotactic protein (GCP)-2, RANTES, monocyte chemoattractant protein (MCP)-2, IL-8, growth-related oncogene-2 and, to a lesser extent, macrophage-inflammatory protein 1α, MCP-1, EXODUS, and CXCL-16. Data suggest that TLR3, TLR4, TLR7, and TLR9 are also expressed at messenger ribonucleic acid (mRNA) and possibly protein level[25] and may augment inflammatory cytokine expression by DCs from patients with RA.

DCs are thought to be the most important antigen-presenting cells in RA. Indeed, the proinflammatory environment favors DC maturation in regional lymph nodes as well as inflamed tissue.[26] Thus in peripheral blood, DC precursors express either an immature CD33$^{\text{dim}}$CD14$^{\text{dim}}$CD16^{-} phenotype or a more mature MHC class II$^{\text{bright}}$ CD11c^{+}CD33$^{\text{bright}}$CD14$^{\text{dim}}$ surface phenotype typical of conventional myeloid DC (mDC); neither population expresses costimulatory molecules. In contrast, synovial fluid and tissue DC subsets resemble mature peripheral blood cells; in addition, a subset expresses high levels of CD86 and can support allogeneic mixed leukocyte reactions. More recent data indicate that they may differentiate further *in situ* as suggested by nuclear translocation of RelB in DC localized within perivascular infiltrates, consistent with prior cytokine receptor or TLR engagement *in vivo.* Perivascular RA synovium also contains populations of MHC class II^{+}CD11c^{-}CD123^{+} plasmacytoid DC (pDC); in contrast to the conventional myeloid DC subset, these are RelB^{-} and comprise ~30% of all synovial DC. A subset of pDC express BDCA2, capable of producing IFN-α *in situ.* Unlike their peripheral blood counterparts, synovial pDC efficiently activate allogeneic T cells to proliferate as well as to produce IFN-γ, TNF-α, and IL-10.

Although the common myeloid precursor cell is the precursor for all myeloid cells, including DC and tissue macrophages, the precise role of monocytes—namely CD14^{+}CD16^{-}, CD14^{+}CD16^{+}, and the more recently described CD14$^{\text{dim}}$CD16^{+} subset—in synovial inflammation is uncertain. They are good candidates as persistence factors through their capacity to activate and polarize T-cell subsets, to respond to the environment through TLR expression, and to produce a wide range of inflammatory mediators, including IL-1, TNF, IL-6, IL-8, CCL2, NO, and type I IFN. In contrast a more recently described CD14$^{\text{dim}}$CD16^{+} subset may function to sense damage, to scavenge cell debris and higher-order molecular complexes, and to promote angiogenesis and tissue repair. More work is needed to establish the contribution of these distinct functional subsets to the chronic inflammatory process. Nonetheless, multiple cellular and molecular determinants exist in RA synovium that could serve both to initiate and perpetuate the immune inflammatory response.

This initial wave of inflammation has two major consequences. First, inflammatory cytokines will promote the activation of vascular endothelium, changes that occur very early in disease (see above and Fig. 52.3). Under the influence of locally generated cytokines and chemokines, synovial postcapillary venules undergo morphological changes to an extent that they resemble high endothelial venules similar to those observed in secondary lymphoid organs. The second major consequence is the migration of inflammatory leukocytes, including polymorphonuclear leukocytes and immature or undifferentiated monocytes, orchestrated by chemokines produced by resident stromal as well as infiltrating cells (Fig. 52.3). CXC, CC, C, and CX$_3$C chemokines all play a role, exerting chemotactic activity toward neutrophils, lymphocytes, and monocytes but also influencing the topology of inflammatory infiltrates. They are invariably early activation genes (*e.g.,* type I IFN response genes), in response to inflammatory stimuli. Besides the homeostatic chemokines described above, the key players include IL-8/CXCL8, RANTES/CCL5, MIP-1α/CCL3, SDF-1/CXCL12, IP-10/CXCL10, and MCP-1/CCL2.[27] Upregulation on endothelium of cell surface adhesion molecules, including ICAM-1, VCAM-1, and E-selectin, permits the rolling and adhesion of leukocytes as they migrate. In synovial joints, resident stromal cells and infiltrating macrophages are a dominant source of such factors. Crucially, the expression of cognate chemokine receptors such as CCR4, CCR5, CCR6, CXCR3, and CX3CR1 on inflammatory cell subsets contribute selectivity of cellular recruitment.[27] Within the T-cell compartment, there exist distinct profiles of chemokine receptor expression, patterns evolved perhaps to facilitate eradication of pathogens. For example, T-helper 1 cells preferentially express CXCR3 and CCR5; Th2 cells express CCR3; Th17 cells express CCR6 and CCR4; and Tfh cells express CXCR5. Data suggest that CCR5, CCR6, CCR7, CXCR3, CXCR4, and CXCR5 may all be important for B-cell migration into synovium. Together, these events characterize the acute phase of an innate immune response, a key checkpoint that precedes the progression to subsequent events that herald the onset of the chronic inflammatory phase.

Autoantigens in RA

Although current models of adaptive immune responses would suggest that DC carry antigens derived from damaged or dying synovial tissue, the molecular nature of disease-associated antigens has, until recently, remained an enigma. Many RA-associated autoantigens have been described (see Table 52.1 for examples), and for some there exist *in vivo* arthritis model correlates.[28] The best described are collagen II, proteoglycans, HCgp-39,

TABLE 52.1 Autoantigens in Rheumatoid Arthritis

Established	T or B Cell[a]	Molecular Specificity	Assay[b]
Immunoglobulin G	B	Human Fc IgG	Rheumatoid factor
Cyclic peptides	T and B	Citrullinated peptides	Anti-CCPs
Fibrin	T and B	α- and β-chain epitopes	Research[b]
Fibrinogen	T and B	Multiple epitopes	Research[b]
Enolase	T and B	CEP-1 dominates	Research[b]
Vimentin	T and B	Citrullinated vimentin	MCV assay
Collagen II	T and B	Multiple epitopes	Research[b]
HnRNPA2	B	Multiple epitopes	Research[b]
Aggrecan	T and B	Multiple epitopes	Research[b]
HCgp-39	T	Multiple epitopes	Research[b]
Glucose-6-phosphate isomerase	B	Multiple epitopes	Research[b]

[a] Denotes autoantigens recognized by T or B cells, or both.

[b] Assay is either not commercially available or not in routine clinical use. Details of assays may be found in primary research communications.

CCPs, cyclic citrullinated peptides; *MCV,* modified citrullinated vimentin.

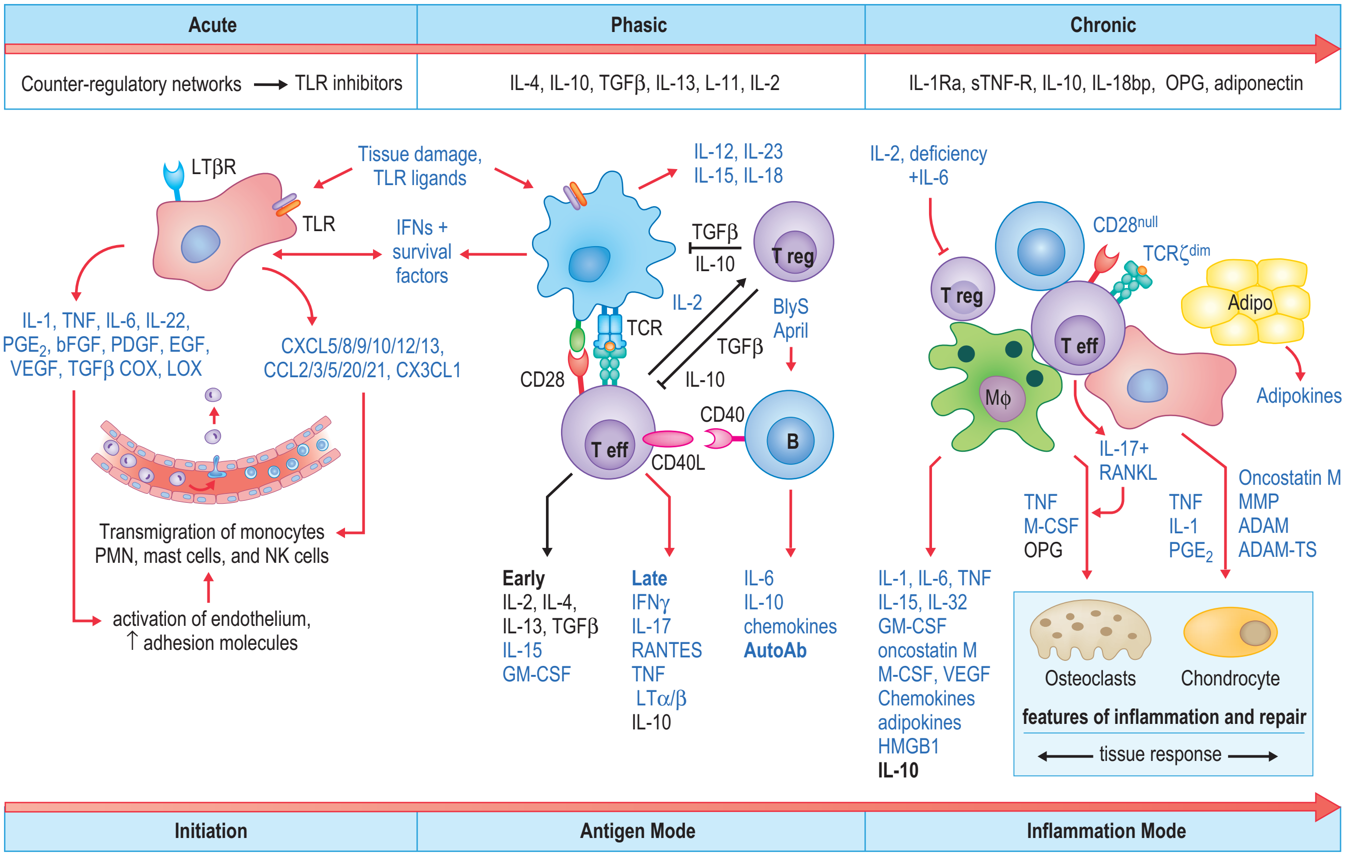

Acute
Phasic
Chronic
Counter-regulatory networks → TLR inhibitors
IL-4, IL-10, TGFβ, IL-13, L-11, IL-2
IL-1Ra, sTNF-R, IL-10, IL-18bp, OPG, adiponectin
LTβR
TLR
Tissue damage, TLR ligands
IFNs + survival factors
IL-1, TNF, IL-6, IL-22, PGE2, bFGF, PDGF, EGF, VEGF, TGFβ COX, LOX
CXCL5/8/9/10/12/13, CCL2/3/5/20/21, CX3CL1
Transmigration of monocytes PMN, mast cells, and NK cells
activation of endothelium, ↑ adhesion molecules
IL-12, IL-23
IL-15, IL-18
TGFβ
IL-10
T reg
IL-2
TCR
TGFβ
IL-10
BlyS
April
CD28
CD40
T eff
CD40L
B
Early
IL-2, IL-4, IL-13, TGFβ
IL-15
GM-CSF
Late
IFNγ
IL-17
RANTES
TNF
LTα/β
IL-10
IL-6
IL-10
chemokines
AutoAb
IL-2, deficiency +IL-6
T reg
CD28null
TCRζdim
Adipo
T eff
Mϕ
Adipokines
IL-17+ RANKL
TNF
M-CSF
OPG
TNF
IL-1
PGE2
Oncostatin M
MMP
ADAM
ADAM-TS
IL-1, IL-6, TNF
IL-15, IL-32
GM-CSF
oncostatin M
M-CSF, VEGF
Chemokines
adipokines
HMGB1
IL-10
Osteoclasts
Chondrocyte
features of inflammation and repair
← tissue response →
Initiation
Antigen Mode
Inflammation Mode

FIG 52.3 Cytokine Networks in Rheumatoid Arthritis. The pathogenesis of rheumatoid arthritis can be thought of as a series of complex and closely related pathways temporally and spatially regulated. These include (1) an acute insult that may trigger the disease, characterized by stimulation of FLSs by inflammatory stimuli and the generation of cytokines and chemokines that promotes the migration and infiltration by cells of the innate immune system; (2) repeated episodes of antigen-specific adaptive immune responses (in lymph node, bone marrow, and *in situ*). Failure to resolve adaptive immunity is a key checkpoint that may lead to (3) a cytokine-driven chronic inflammatory phase when multiple cellular and molecular components sustain the response. Through multiple pathways acting on many cell types, this process leads to tissue injury. Proinflammatory pathways are shown in blue (text) and red (arrows), whereas antiinflammatory, counter-regulatory pathways are shown in black (text and arrows). Adipo, adipocyte; AutoAb, autoantibodies; B, B cell; DC, dendritic cell; FLSs, fibroblast-like synoviocytes; MΦ, macrophage; TCR, T-cell antigen receptor; T_{eff}, effector T-helper cell; Treg, regulatory T cell.

glucose-6-phosphate isomerase, α-enolase, vimentin, and citrullinated fibrinogen. However, when used as recombinant native antigen, few have been found to elicit reproducible and/or robust PB or SF T- or B-cell responses in a significant proportion of patients, as opposed to healthy donors. There are several plausible explanations for this. The most obvious is that antigens that drive autoimmune arthritis are not the same in mouse and man; another factor is that effector T-cell responses are more important very early but are blunted in established disease, or they may simply present at very low frequency. Another possibility is that the autoantigens used to test lymphocyte reactivity *in vitro* do not carry the posttranslational modifications (*i.e.*, the neoepitopes) recognized by autoantibody or antigen receptor. Good examples are the carbohydrate moieties of collagen II epitopes that serve as key TCR contacts in collagen II immunity, or the citrullination of key arginine residues in triple-helical CII or fibrinogen peptides that appear to be the immunodominant autoantibody epitopes. Utilizing MHC class II tetramers loaded with derivatized peptides, it has been possible to demonstrate in peripheral blood and synovial fluid from RA patients low frequencies of T cells (<1 per 10^6), recognizing citrullinated cognate peptides derived from aggrecan, vimentin, α-enolase, cartilage intermediate layer protein (CILP), and fibrinogen α and β.[29] The impact of posttranslational modifications on the cleavage of peptide epitopes is just beginning to be appreciated, but it is an attractive and highly plausible mechanism for the generation of self neoepitopes.

The Discovery of Citrulline as a Key Target for Autoimmunity in RA

In 1998, van Venrooij and colleagues first reported that patients with RA carried antibodies that recognized deiminated peptides of fillagrin,[30] the substrate that was found to be the antigen recognized in rat keratinized epithelium. This substrate formed

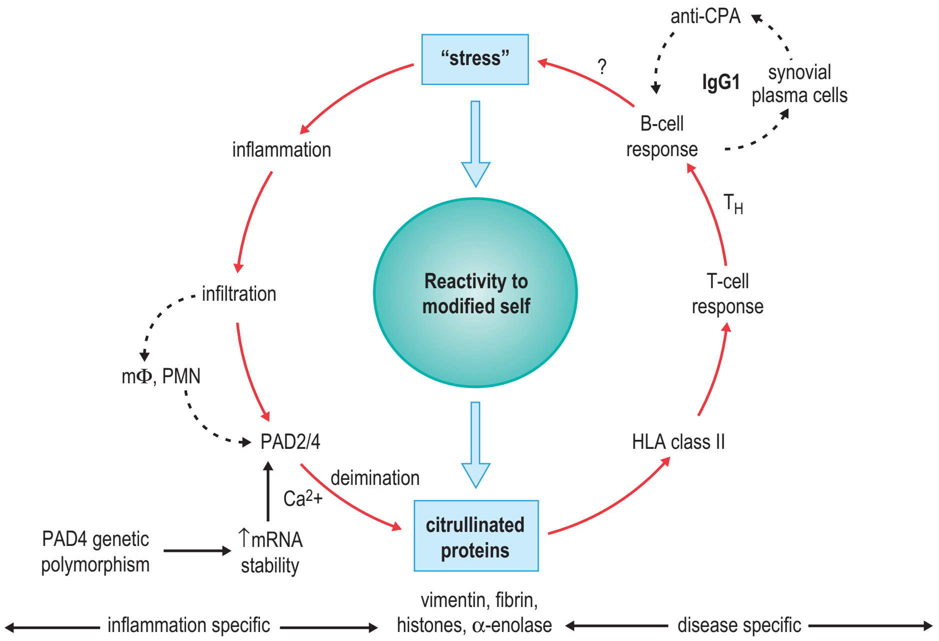


FIG 52.4 **The Generation of Autoantibodies to Citrullinated Protein Antigens.** The stressed and inflamed synovium is characterized by an influx of inflammatory cells, including macrophages and neutrophils that express peptidyl-arginine deiminases (PAD). In the presence of sufficient Ca^{2+}, PADs deiminate target proteins including, among others, vimentin, fibrinogen, aggrecan, type II collagen, histones, and α-enolase. This reaction is inflammation, but not disease-specific. The combination of environmental stimuli (including inflammation and exposure to tobacco smoke) and the inheritance of specific HLA-DRB1 alleles favor T- and B-cell immune responses to the host's derivatized neoepitope peptide antigens. Neoepitopes may also be generated by changes in peptide cleavage during antigen processing as a consequence of the Arg → Cit modification. Anti-CPA, autoantibodies to citrullinated protein antigens; T_H, T-helper effector cell.

the basis of the antiperinuclear factor (APF) assay, reported many years before. Using new-generation anti-cyclic citrullinated peptide (anti-CCP) based assays, the presence of these antibodies, now collectively termed ACPAs, is now confirmed to be both sensitive (up to 80%) and highly specific (>95%) for the diagnosis of RA.[30] Indeed, serum anti-CCPs levels are stable in established disease, they have been detected as early as 14 years before disease onset and have been shown to be predictors of radiographic progression. Changes in isotype usage and spreading of antigenic specificities suggest that the ACPAs response matures before disease onset. Citrullination is not specific for RA. Indeed, citrullination may be inflammation specific, because it has been documented in inflamed synovium derived from patients with reactive arthritis and psoriatic arthritis as well as RA, but not in those with OA. What appears specific for RA is the immune response to citrulline (Fig. 52.4). As outlined above, a link between ACPAs and HLA-DRB1 alleles, specifically SE$^+$ alleles, has now been established.[17] Linkage analysis across chromosome 6 has documented a peak with LOD scores in excess of 10 for ACPAs$^+$ patients but not for those that do not carry these antibodies. This relationship is independent of RF status, because the SE allele frequencies in ACPAs$^+$ patients are twice those of ACPAs$^-$ patients, even those patients who are RF$^+$. Indeed, the risk of carrying SE in RF$^+$ACPAs$^-$ patients is no different from that in the healthy control population. These data would be consistent with a model where T cells from patients with RA can recognize peptide autoantigens modified by citrullination if they carry SE$^+$ DRB1 alleles. Crystallographic studies suggest that the conversion of positively charged arginine at key residues in antigenic peptides from candidate autoantigens such as fibrinogen to neutral citrulline is permissive for peptide binding and recognition by autoreactive T cells *in vivo.*[18]

Finally, citrullination is widespread in multiple tissues in response to appropriate provocations. Although the molecular basis for these triggers is poorly understood, recent data point to a link between smoking—an environmental exposure known to be linked with RA—citrullination, and individuals carrying SE.[31] Thus cells derived from bronchoalveolar lavage from smokers, but not from nonsmokers, express citrulline. The association between the development of RA and smoking has now been linked to ACPAs$^+$ patients, whose relative risk increases 20-fold if they smoke and carry two copies of the SE$^+$ DRB1 alleles; in comparison, the relationship in ACPAs$^-$ patients appears to be much weaker or nonexistent. A second example relates to the fact that *Porphyromonas gingivalis,* a pathogen associated with severe periodontitis, expresses its own deiminating enzyme and has the capacity to modify host proteins such as fibrinogen and α-enolase in inflamed gingival tissue. Molecular mimicry is suggested by the finding that anti-α-enolase autoantibodies from patients with RA cross-react with *P. gingivalis*–derived α-enolase. The links between autoantibodies to the immunodominant α-enolase epitope CEP1, smoking, SE$^+$DRB1, and disease-associated *PTPN22* alleles are the strongest defined to date. These data are among the first to show a direct link between environmental exposure and disease-specific immune responses governed by immune response genes.

Lymphocyte Biology

Flow cytometric analysis of dissociated synovial mononuclear cell cultures indicates that infiltrating T lymphocytes make up approximately 10–35% of cells in inflamed tissue. The ratio of CD4 to CD8 T cells seen in peripheral blood is skewed in favor of the CD4 subset. Much of the information relating to phenotype and function is derived from detailed analyses of synovial fluid and, to a lesser extent, synovial tissue subsets. Similar enrichment of these subsets is commonly detected in PBL. Synovial T cells express phenotypic markers of antigen experienced, terminally differentiated T cells with enhanced migratory capacity. Thus synovial T cells typically carry cell surface markers such as HLA-DR$^+$, LFA-1$^+$, VLA-1$^+$, CXCR3, CD69$^+$, CD45RO$^+$, CD45RA$^-$, CD45RB^{dim}, CD29^{bright}, CD27$^-$, and CD25$^-$. Recent analysis of citrulline-reactive B cells from RA patients indicate that they display markers of class-switched memory B cells and plasmablasts (CD20$^+$CD27$^+$IgD$^-$). Analysis of the variable (V) regions of the immunoglobulin heavy and light chains confirm that antigen-specific activation and differentiation of B cells into plasma cells takes place in draining lymph nodes as well as in the chronically inflamed synovial tissue of patients with RA. Likewise, analysis of the T-cell repertoire indicates that the synovium provides a niche for supporting expansion of specific T cells, whose clonality is shared between different joints from the same patient, but not substantially with clones from paired blood samples.[32]

Analysis of synovial T cells reveals that expression of the TCR invariant chain subunits CD3ε and TCRζ is found at lower density than that found in corresponding peripheral blood T cells, evidence of prior antigen engagement. While synovial fluid T cells are also FasL$^+$, Bcl2$^-$, Bax^{bright}, favoring a proapoptotic state, it is thought that environmental cues transduced through common γ chain receptor signaling cytokines such as IL-2, IL-7, and IL-15, as well as type I interferons, prevent apoptosis of T cells *in situ.*[33] Synovial tissue–derived lymphocytes may be different. Consistent with their state of terminal differentiation, a subset of synovial T cells are CD28$^-$, while at the same time expressing a range of natural killer (NK) cell surface receptors that are thought to contribute to effector function independently of cognate antigen.

For over a decade it has been known that one of the dominant cytokines expressed in synovial T cells from patients with established disease is IFN-γ. Somewhat surprisingly, a significant proportion of IFN-γ$^+$ cells also express IL-10. Recent data support a model in which there exists, during differentiation, a transition from IFN-γ$^+$ Th1 T cells through an IFN-γ$^+$IL-10$^+$ double-positive stage to a single-positive IL-10$^+$ stage. This last phase could represent part of the normal lifecycle of an effector T cell, where IL-10 expression promotes the resolution of the adaptive immune response, attenuating the function of DCs. Interestingly data indicate that this Th1 lifecycle involving a switch from IFN-γ to IL-10 production may be defective in patients with active RA, providing an important mechanism for persistence of effector responses *in vivo.*[34] Treatment of RA patients with anti-TNF appears to restore IL-10 production through induction of specific transcription factors.[35]

Emerging evidence, supported by preclinical models, points to a role for IL-17 in the pathogenesis of inflammatory arthritis. Immunohistochemistry of RA synovium indicates that the dominant source is CD4 T cells, although γδ T cells, NK cells, and mast cells also contribute. Flow cytometric analysis reveals that the frequency of IL-17$^+$ T cells is less than that of IFN-γ$^+$ T cells and that a subset expresses both IL-17 and IFN-γ. The TNF$^+$IL-6$^+$IL-1$^+$–rich synovial environment is capable of supporting the differentiation of Th17 cells. Sustained IL-17 expression and stimulation of IL-17R–expressing stromal cells, macrophages, chondrocytes, and osteoclast precursors are implicated in promoting local tissue responses. These include induction of chemokines, matrix metalloproteinases, nitric oxide, and cyclooxygenase and promotion of osteoclastogenesis, which are key mechanisms underpinning pathways of cartilage destruction and bone erosion.

Molecular Basis of Persistence

Understanding the molecular basis of persistence is of major importance in diseases such as RA. It is now well established that enhanced expression of inflammatory cytokines is one of the hallmarks of chronic inflammatory diseases such as RA.[36] With few exceptions most cytokines are expressed in inflamed synovium at mRNA or protein levels. For example, unlike PBL from healthy donors, explants of synovial mononuclear cells constitutively express IL-1, IL-6, IL-8, granulocyte macrophage–colony-stimulating factor (GM-CSF), and a vast array of chemokines as well as growth factors including FGF, VEGF, and PDGF (Fig. 52.3). Although antiinflammatory or immunoregulatory cytokines, including IL-4, IL-10, IL-11, and TGF-β, as well as specific inhibitors of IL-1α/β (IL-1Ra) and TNF-α (soluble TNF-R) may also be detected, functional bioassays suggest that levels of these naturally occurring inhibitors may be deficient in relative terms. Indeed, levels of IL-1Ra and soluble TNF-R are significantly upregulated in synovial fluid compared with serum from patients with RA, and yet IL-1 and TNF bioactivity persists in spite of this, suggesting that attempts to suppress cytokine bioactivity are insufficient. The specific activity of cytokine subsets, their networks, and their contribution to RA are illustrated in Fig. 52.3. In many cases, the pathogenic roles of cytokines have been established in rodent models, either by selectively blocking their function (*e.g.,* with specific monoclonal antibodies or soluble receptor fusion proteins), by gene-targeting approaches, or by transgenesis.

Immune Regulation

Investigation of regulatory cell subsets and their antiinflammatory properties has perhaps more firmly established the concept that failure of the host's intrinsic mechanisms of immune regulation can underpin autoimmune diseases. Experiments in gene-deficient mice (*e.g.,* Foxp3, IL-2, IL-2R, IL-2R signaling, STAT5, IL-10, TGF-β) lend support to this concept. In RA the data remain less clear.[36] For example, there is *in vitro* evidence for a relative deficiency of constitutive IL-10 expression in synovial cell cultures, and yet clinical trials of IL-10 have been disappointing. These results may reflect the complex role of these cytokines in disease pathogenesis. The identification of defective numbers and/or function of CD4$^+$CD25^{bright} Tregs has been suggested by several investigators, but reports are conflicting.[37,38] Some studies have shown clear reductions in numbers of peripheral blood Tregs in patients with RA, whereas others have shown no difference; interestingly, Tregs appear deficient in cell number in subsets of children with juvenile idiopathic arthritis (JIA) whose disease tends to progress. In synovial joints, the data are more consistent, with many reports showing substantial increases in Treg numbers in synovial tissue and fluid compared with paired PB. However, some studies have reported normal function at a cellular level, whereas others have shown depressed regulatory function. An emerging concept, whose mechanism is yet to be proven, is that synovial effector T cells are refractory to regulatory pathways.

Impact of the Immune Response on Cartilage and Bone

For many years, it was considered that the terminal effector phase of chronic inflammation that led to cartilage destruction and bone resorption was driven almost exclusively by inflammatory cytokines and proteinases. IL-1, MMPs (MMP1, 3, 8, 13), and aggrecanases (ADAMTS 4 and 5) were, and remain, major culprits. Attempts to establish more directly a link between adaptive immunity and destruction of target tissue failed, not least because of the lack of a direct physical link between lymphocytes, chondrocytes, and bone. A breakthrough came in the late 1990s with the identification of the TNF/TNFR family member receptor for activation of nuclear factor (NF)-κB ligand (RANKL)/TRANCE/ODF and its counter-receptor RANK, and the dissection of the molecular and cellular components required for osteoclast differentiation from monocyte precursors.[39] According to contemporary paradigms, RANKL is necessary and sufficient for osteoclast differentiation. TNF, M-CSF, IL-1, and IL-17 contribute, and RANKL-independent pathways may also play a role. RANKL is expressed on synovial fibroblasts and osteoblasts but also on activated T cells, its counter-receptor being expressed on myeloid lineage cells including monocytes, osteoclast precursors, and DCs. Its expression is regulated by inflammatory mediators including TNF and PGE_2. RANKL is shed, probably through the action of several membrane-associated proteases including MT1-MMP (MMP14). Gene targeting of RANKL or RANK in mice leads to inhibition of osteoclastogenesis and a profound osteoporotic bone phenotype. Deletion of osteoprotegerin (OPG), the naturally occurring decoy soluble receptor for RANK, leads to unbridled osteoclast differentiation and bone resorption and substantially reduced bone mass. A good example of the link between adaptive immunity and bone resorption comes from cytotoxic T lymphocyte antigen-4 $(CTLA-4)^{-/-}$ mice, characterized by sustained chronic T-cell activation.[39] T cells from these mice overexpress RANKL. Importantly, the bone but not the inflammatory phenotype is rescued by inhibition of RANK/RANKL signaling. In RA, several studies have demonstrated perturbations of serum RANKL/OPG ratios, a parameter currently under investigation as a biomarker for bone homeostasis. Recent clinical experience with denosumab, a fully humanized monoclonal antibody that binds to RANKL, demonstrates increased bone mineral density and reduced bone turnover in patients with RA.

Distinct from enhanced bone resorption, bone formation is impaired in RA, although until recently the pathways involved were rather obscure. The inflammatory process itself, along with its effects on osteoblast maturation and function, has been directly implicated because bone formation at surfaces adjacent to bone marrow, as opposed to inflamed synovium, are relatively well preserved. Recent data also suggest that ACPAs can also directly promote osteoclastogenesis, providing a mechanistic link between autoantibodies and joint destruction that also depends on IgG glycosylation status.[40] Evidence now points to the canonical Wingless (Wnt) signaling pathway as being an essential control point in osteoblast function, based on the observation that in animal models of RA antibodies to dickkopf homologue 1 (DKK1), a secreted antagonist of Wnt blocking signals at the level of its cognate receptor Frizzled, promote bone formation and inhibit bone resorption indirectly by increasing production of OPG.[39] This indicates that Wnt signals negatively regulate osteoclastogenesis. Oxidative stress and hypoxia within the synovial compartment may contribute to attenuation of osteoblast-promoting *Wnt* signals.

The interplay between bone formation and resorption is fundamental to bone homeostasis, especially in the context of chronic inflammatory disease. Recent evidence suggests that the molecular basis for this so-called coupling lies with the osteoprotective factor semaphorin 3A. This factor regulates bone mass in both osteoblasts (through effects on the canonical Wnt/β-catenin signaling pathway) and osteoclasts (by inhibiting RANKL-dependent osteoclast differentiation).[41] Together these data support the view that bone and joint integrity are regulated by a delicate balance of catabolic and anabolic immune and inflammatory mediators influencing the maturation and function of osteoblasts (Wnt:DKK1) and osteoclasts (RANKL:OPG).

CLINICAL FEATURES

Disease Onset

RA is a heterogeneous disease that does not conform to a single clinical entity.[2] Whereas 10% of patients may have an acute severe onset and 20% a more subacute onset, the onset of signs and symptoms may be insidious in up to 70% of patients. A more episodic or palindromic onset has also been described. A common presentation, more likely during the winter months, will be that of a female in her fifth to sixth decade of life who complains of diffuse symmetrical joint pain, swelling, and stiffness of peripheral joints. Patients may complain that they can no longer make a fist, especially in the early morning. The targeting of afflicted synovial joints may be symmetrical in most but not all cases, typically affecting small joints of the hands and feet as well as wrists. Less frequent are those presenting with slow-onset monoarticular disease. Patients not fulfilling the diagnostic classification criteria for RA (see below) may, at this point, be ascribed the more appropriate diagnostic label of undifferentiated arthritis, because in a proportion of cases, signs and symptoms may resolve spontaneously. At 1 year, approximately 30% will progress to develop the syndrome recognized as RA. Systemic disease is much less common in current clinical practice, in part through application of earlier, more intensive treatment regimens (see below). Nonetheless, the systemic nature of the disease may be manifested through a wide array of systemic, extraarticular clinical features that may occur in patients with disease at the more severe end of the spectrum. A spectrum of disease severity may also be evident from hand radiographs; examples are shown in Fig. 52.5.

Diagnosis

Classification Criteria

The American College of Rheumatology criteria for the classification of RA are a set of clinical and laboratory parameters, established largely for epidemiological purposes, that serve as a guide for the diagnosis of RA. They are relatively straightforward and easy to apply, especially to patients with established disease. However, failure of a patient with early signs and symptoms of an inflammatory arthropathy to fulfill them does not mean that the individual does not have RA. The 1987 criteria were simplified further by removing the “probable,” “definite,” and “classic” subclassifications.[42] These criteria returned a sensitivity for RA of 91–94% and a specificity of 89% in the clinical setting. New criteria, established by a steering group comprising members of the ACR and EULAR and published in 2010,[43] are based on a

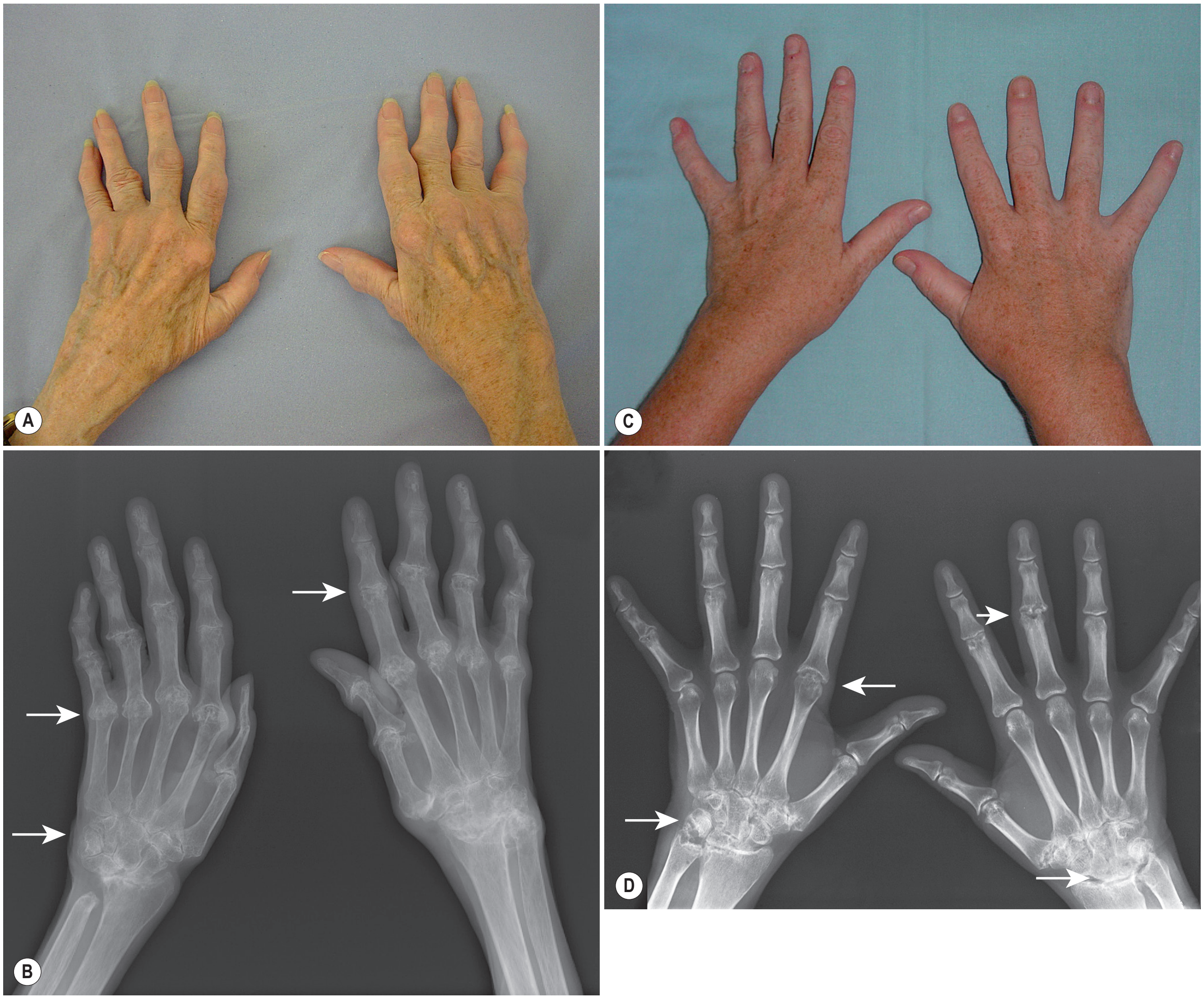


FIG 52.5 **Photomicrographs and Radiographs of the Rheumatoid Hand.** Chronic, severe, and erosive RA, refractory to treatment and showing joint swelling and classical deformities (A and B), is compared with erosive disease whose progression has been attenuated by combination and biological therapy (C and D). White arrows indicate major areas of bone and cartilage destruction. Reproduced with kind permission from the patients.

weighted score around four domains, including distribution and type of joint involvement (tender as well as swollen joints are included, scoring 0–5 points), serology (0–3) that includes RF or ACPAs and is weighted according to antibody levels, duration of synovitis (0–1), and acute-phase reactants erythrocyte sedimentation rate or C-reactive protein (0–1). A score of 6 or above is indicative of an early disease state requiring initiation of disease-modifying antirheumatic drugs (DMARDs) such as methotrexate (Table 52.2),[43] and so these criteria are thought to better reflect an intention to treat on the part of the supervising physician. Although the presence of synovitis in at least one joint is required (in the absence of an alternative diagnosis that better explains the synovitis), there remains debate as to whether subclinical synovitis of specific joints, defined by magnetic resonance imaging (MRI) or high-resolution ultrasonography (HRUS), should be included in the joint score.

Laboratory Findings

Until the late 1990s, IgM RFs, autoantibodies that recognize the Fc subunit of IgG, remained one of the few parameters of value in the clinical setting, forming the basis of the seropositive versus seronegative RA stratification and identifying those patients more likely to progress to erosive disease with or without extraarticular features. Nevertheless, RF can be detected in up to 5% of the healthy population and in 10–20% of the elderly population (>65 years of age); RF is found in a range of rheumatic conditions including Sjögren syndrome, SLE, and cryoglobulinemia as well as in acute infectious and neoplastic disease entities, influencing its diagnostic utility. In general RF is not of value for monitoring responses to therapy.

The discovery of ACPAs, which can be detected very early in the disease process, has had a major impact on diagnostic practice as the assays have become more widely available.[30] They also

TABLE 52.2 2010 ACR/EULAR Classification Criteria for Rheumatoid Arthritis

Domain	Weighted Score
Joint Involvement (0–5)	
1 medium to large joint	0
2–10 medium to large joints	1
1–3 small joints	2
4–10 small joints	3
>10 joints (with at least one small joint)	5
Serology (0–3)	
Neither RF- or ACPAs-positive (≤ULN)	0
At least one test, low positive titer (>1 ≤3 × ULN)	2
At least one test, high positive titer (>3 × ULN)	3
Duration of Synovitis (0–1)	
<6 weeks	0
≥6 weeks	1
Acute-Phase Reactants (0–1)	
Neither ESR or CRP abnormal	0
Abnormal ESR or CRP	1
TOTAL (≥6 indicates *definite* RA):	_____

ACPAs, antibodies to citrullinated protein antigens; *ACR,* American College of Rheumatology; *CRP,* C-reactive protein; *ESR,* erythrocyte sedimentation rate; *EULAR,* European League Against Rheumatism; *RF,* rheumatoid factor; *ULN,* upper limit of normal.

have prognostic value in terms of radiographic progression, and titers may alter with therapy. The new-generation anti-CCPs kits have demonstrated diagnostic sensitivity of 80% and specificity of 98%. As the range of RA-associated autoantigens has expanded and the repertoire of citrullinated target autoantigens has become better defined, multiplex assays of serum autoantibodies are likely to play an increasingly important role in the diagnosis and prognosis of subsets of autoantibody-positive inflammatory arthritides.

CLINICAL PEARLS

Predictors of Poor Outcome in Rheumatoid Arthritis

- Chronic, unremitting disease onset, especially at advanced age
- High disease activity scores at baseline
- Female gender
- Poor functional status as determined by validated functional disability indices such as the Stanford Health Assessment Questionnaire (HAQ) and the Arthritis Impact Measurement Scale (AIMS)
- Low socioeconomic status
- Systemic and extraarticular features
- Depression and anxiety
- Comorbidity, *e.g.,* infection, cardiovascular disease, renal impairment
- Early erosive disease (in first 6–12 months; may be associated with ACPAs autoantibodies)
- Persistent acute-phase response (*e.g.,* time-integrated CRP levels)
- Autoantibodies (RF and ACPAs) and HLA-DRB1 status (SE$^+$)
- Significant delay in early use of DMARDs and corticosteroids

ACPAs, antibodies to citrullinated protein antigens; *CRP,* C-reactive protein; *DMARDs,* disease-modifying antirheumatic drugs; *RF,* rheumatoid factor; SE, shared epitope.

Insights Into the Preclinical Phase of Disease

ON THE HORIZON

The Preclinical Phase of Rheumatoid Arthritis

- High-risk individuals include those with inflammatory joint pain (arthralgia) and serum RA–associated autoantibodies. From 40–50% of those with high-titer ACPAs (with or without rheumatoid factor) may develop clinical synovitis within 24 months.
- Targeting these high-risk individuals with preventative strategies provides the best chance of achieving cure. Trials of rituximab, abatacept, or hydroxychloroquine are ongoing.
- In-depth molecular and cellular studies for characterizing the preclinical phase of disease will be critical to the success of this endeavor.
- Models for progression to RA have identified the following phases:
  - (I) Genetic risk
  - (II) Genetic risk with autoimmunity (*e.g.,* ACPAs, rheumatoid factor)
  - (III) Genetic risk with autoimmunity and arthralgia (but absence of clinically apparent synovitis)
  - (IV) Undifferentiated arthritis (clinically apparent synovitis, not fulfilling RA classification criteria)
  - (V) Early RA (fulfilling disease criteria; need for DMARDs)
- Stratification of risk, including genetic, serological, and demographic factors, will permit the identification of subjects most suitable for intervention studies.
- Studying the impact of lifestyle modifications, *e.g.,* diet, stopping smoking, would be of great interest.
- Reestablishing immune homeostasis and/or induction of immune tolerance may provide the best chance of achieving cure in subjects during the preclinical phase of disease. Early-phase studies of peptide immunotherapy indicate that this approach is well tolerated.
- Establishing robust assays for signatures of immune tolerance (*e.g.,* immune phenotyping by flow cytometry, extended autoantibody serotyping, multiplex assays for detection of inflammatory mediators in serum, whole-blood transcriptomic profiles) that reflect a healthy immune system will greatly facilitate these endeavors.

ACPAs, antibodies to citrullinated protein antigens; *DMARDs,* disease-modifying antirheumatic drugs; *RA,* rheumatoid arthritis.

The identification of a preclinical phase of RA, through the detection in serum of autoantibodies up to 14 years before disease onset, has sparked considerable interest in characterizing the RA prodrome in more detail. Work has centered around evaluation of acute-phase proteins and serum cytokines and chemokines, which appear to rise 1–2 years before onset of clinical disease; ACPAs isotype usage and epitope spreading; whole-blood gene expression analysis, which shows similarities to patients with established disease (including type I interferon-inducible gene signatures in a subset); imaging by MRI and ultrasonographic modalities; and, most recently, analysis of lymph node immune phenotypes, showing evidence of lymphocyte activation. The strongest predictors of progression in those at-risk subjects are joint pains without clinically detectable synovitis (arthralgia) combined with the presence of high-titer ACPAs and RF autoantibodies. Risk scores, which take into account other demographic factors, have been developed and are currently being evaluated. Progression rates for ACPAs$^+$ arthralgia patients with subclinical synovitis detected by ultrasound examination are of the order of 20% over a median of 28 months,[44] while arthritis-free survival curves suggest that 40–50% of those at highest risk will develop clinically apparent synovitis in 2–3 joints within 24 months. This rate of progression is considered to be sufficient to justify secondary prevention intervention studies. Somewhat surprisingly, genetic markers, such as HLA-DRB1 SE, have not contributed dramatically to risk estimates.

TREATMENT

Disease-Modifying Antirheumatic Drugs

THERAPEUTIC PRINCIPLES

Treatment Paradigms in RA

- Education and counseling through early involvement of a multidisciplinary team, including specialist nurse and other health care professionals; appropriate balance of rest and exercise during disease flares
- Adequate nutrition (especially important with severe, active disease)
- Comprehensive assessment of disease activity, especially during early phase of disease to achieve rapid disease control
- Complete suppression of inflammation early in the disease, with tight control through regular and frequent reassessments focused around disease activity scores
- Yearly imaging assessments, *e.g.,* X-rays or high-resolution ultrasonography of hands and feet to monitor radiographic progression
- Early use of DMARD/SAARDs
- Early use of corticosteroids, including use of intraarticular joint injections to suppress inflammation, and use of step-up combination therapy in severe disease
- Appropriate use of biologicals, *e.g.,* early use of anticytokine, T- or B-cell–targeted therapies in severe disease
- Early relief of pain with judicious use of NSAID or COX2 inhibitors according to safety/risk profile
- Monitoring for drug toxicity
- Effective contraception, where appropriate
- Bone protection
- Monitoring for risk factors of key comorbidities, including cardiovascular disease
- Prevention of infection through vaccination (preferably before instituting immunosuppressive agents), *e.g.,* against influenza, pneumococcus, and herpes zoster

COX, cyclooxygenase; *DMARD,* disease-modifying antirheumatic drugs; *NSAID,* nonsteroidal antiinflammatory drugs; *SAARD,* slow-acting antirheumatic drugs.

Over the past two decades there has been a dramatic paradigm shift in the therapy of RA from control of symptoms to the control of the disease process and aggressive suppression of inflammation. This shift has come about through a growing appreciation of the relationship between joint inflammation and joint destruction and also through the development of imaging technology that has documented evidence of erosive changes within the first 6–12 months of disease. The impact of this paradigm shift in therapeutic terms is striking. Traditional "go-low, go-slow" regimens of the 1970s and 1980s included the initiation of nonsteroidal antiinflammatory drugs (NSAIDs), followed by implementation of DMARDs only after destructive disease became evident. Depending on the clinical response, sequential monotherapy was the norm. Although this strategy may still be appropriate for patients with mild disease, current practice now dictates aggressive combination therapy (two or more conventional DMARDs) from the outset for patients with poor prognostic factors, with preference for the faster-acting DMARDs such as methotrexate, leflunomide, and sulfasalazine (onset 3–6 weeks) over slower-acting agents such as hydroxychloroquine, gold, and D-penicillamine (onset 3–6 months), but with the addition of oral or parenteral prednisolone. More recent data suggest that the specific choice of therapy may be less important than the strategy. For example, the TICORA and BeST studies both indicate that intensive treatment when combined with intensive control most convincingly influences outcome measures, including clinical response, retention, functional status, and radiographic progression. The benefits of tight control have been substantiated in many subsequent clinical trials as well as in "real life" clinical practice.[45,46]

Anticytokine Therapy

The introduction of targeted therapy to the clinic using biological agents (*e.g.,* chimeric or fully humanized antibodies to ligands or receptors, soluble receptor fusion proteins, or recombinant receptor antagonists) has transformed the treatment of RA. The prototype biological is TNF-α blockade.[47] The rationale for inhibiting TNF-α bioactivity is based upon its pleiotropic effects on cell activation, cellular adhesion and migration, induction of cytokine and inflammatory gene mRNA and protein, neoangiogenesis, and the regulation of cartilage catabolic factors such as IL-1 and matrix metalloproteinases (Fig. 52.3). TNF-α and other inflammatory cytokines such as IL-1, IL-6, IL-15, IL-17, and GM-CSF are expressed constitutively in inflamed synovial tissue at mRNA and protein level. In many cases the expression of their high-affinity cognate receptors is upregulated and the functional activity of the corresponding naturally occurring inhibitors (*e.g.,* soluble TNF-R or IL-1Ra) is reduced (although levels of protein may be increased, reflecting an attempt at restoring homeostasis). As proof of principle, inhibition of cytokine activity, including TNF-α, IL-1, IL-6, IL-15, and IL-17, have all been shown to have benefit in animal models of arthritis.

Chimeric anti-TNF monoclonal antibodies (infliximab) were first used to treat RA in open-label clinical trials in 1992.[47] Humanized antibodies (adalimumab) and the soluble p75 TNF-R IgG fusion protein (etanercept) were tested subsequently with comparable therapeutic effects; golimumab and a construct comprising a PEGylated anti-TNF antibody Fab fragment (certolizumab) are also licensed for use in patients with RA, as are a growing number of "biosimilar" TNF inhibitors. TNF-α blockade leads to dramatic and rapid reductions in symptoms (pain, stiffness, and fatigue) and signs (joint pain and swelling) of arthritis in a dose-dependent fashion, and in a significant proportion of patients (~60–70%) who have failed conventional DMARDs.[47] As a class, these biological agents can be used repeatedly, and when used in combination with methotrexate, they are superior to either drug alone. The mechanism is not clear but may be related to reductions in immunogenicity of the therapeutic antibody (the antidrug antibody [ADA] response), effects on the half-life of the biological, or perhaps the combined immunological effects of anticytokine and anti-T-cell agents. Radiographic progression may be attenuated, even in subsets of patients with poor clinical responses; in some patients radiographic scores improve, suggesting that healing of joint tissues may take place. TNF-α blockers when used early demonstrate superior remission rates. Some studies have shown that a significant proportion of patients remain in remission even after withdrawal of anti-TNF, providing evidence suggestive of immune modulation. Regardless of the TNF inhibitor used, the contribution of TNF to host defense is further suggested by an approximately twofold increased risk, especially during the first 6 months, of serious infections of the upper respiratory tract, skin, and joint as well as by a significantly increased risk of reactivation of latent tuberculosis. The global risk of malignancy, particularly lymphoma, which already is increased in the RA population as a whole, does not appear to be increased with anti-TNF therapy, although an increased risk of melanoma has been reported through biologicals registries. The development of antinuclear antibodies in 8–15% of patients, as well as rare cases of demyelination, points to effects that may be related to restoration of autoimmune responses in predisposed subjects after TNF-α inhibition.

The clinical benefit of TNF-α blockade has prompted extensive mechanism of action studies.[47] Anti-TNF reduces the acute-phase response, including IL-6 serum levels. Leukocyte trafficking is inhibited, as demonstrated through an early (within hours) and dramatic rise in lymphocyte counts through demargination, a more prolonged and sustained exclusion of leukocytes based on reductions in cellularity of synovial tissue biopsies after treatment and suppression of markers of angiogenesis, including VEGF; TNF blockade promotes lymphangiogenesis and may facilitate cell egress from inflamed synovium. TNF-α blockade has also been shown to downregulate markers of cartilage and bone destruction, including the collagenases MMP1 and 3, and to reduce the ratio of RANKL and OPG in serum, effects that might explain in part the joint-preserving effects of anti-TNF *in vivo.* Finally, there is additional evidence of anti-TNF-induced immune homeostasis, based on restoration of cell-mediated immune responses to recall antigens as well as mitogens, but also on an increase in both the numbers and function of $CD4^{+}CD25^{high}$ Tregs and on increases in expression of the immunoregulatory cytokine IL-10.[35,37] This effect could be explained in part by the effect of TNF blockade on restoring the expression of the Treg-associated transcription factor FOXP3.

The IL-1 receptor antagonist (IL-1Ra) is the only IL-1 inhibitor currently licensed for use in RA. It has disease-suppressing effects in animal models of arthritis, with potent joint protection, but has proven less effective than anti-TNF in patients with RA. Nevertheless, it has been used effectively to treat patients who have failed TNF-α blockade, slowing radiographic progression, and may be efficacious in a subset of individuals with systemic autoinflammatory syndrome–like features. Anti-IL-6R blockade (tocilizumab) is licensed for use in patients with established RA. Evidence suggests that in early disease, tocilizumab as monotherapy is as effective in suppressing signs and symptoms of RA as when it is combined with methotrexate. These effects are mediated through blocking IL-6 actions on the immune response, the acute-phase response, osteoclastogenesis, B-cell activation and immunoglobulin production, angiogenesis, and cell adhesion (reviewed in Kishimoto).[48] Clinical responses are comparable to those observed with TNF blockade. Unlike in psoriasis, the effect of inhibiting IL-17 in RA (with humanized monoclonal antibodies ixekizumab or secukinumab that block IL-17A) has been more modest, although there may exist a subset of RA patients in whom IL-17-driven inflammatory responses dominate. A role for blocking GM-CSF is awaited. Blockade of cell adhesion or chemokine function has proved disappointing in RA for reasons that may lie in the redundancy of these complex systems.

Finally, the impact of using small molecule inhibitors of cytokine signals, transduced through receptors that utilize members of the Janus kinase (JAK) family of tyrosine kinases,[49] has been evaluated in some detail. These are interesting, immunologically important kinases, because gain-of-function mutants are associated with leukemia and lymphoma, whereas loss-of-function is associated with primary immunodeficiency. Jakinibs, including tofacitinib and baricitinib, have been studied in phase III clinical trials of RA patients. Tofacitinib has high affinity for JAK3 but is also able to partially inhibit JAK1 and, to a lesser extent, JAK2; baricitinib blocks JAK1 and JAK2 to a similar extent. Tofacitinib has been studied in head-to-head studies with methotrexate and with TNF inhibitors. In methotrexate-naïve patients, tofacitinib was superior to methotrexate in reducing signs and symptoms of RA and in inhibiting the progression of structural joint damage. In patients taking background methotrexate therapy, tofacitinib was similar in efficacy to adalimumab. JAK1- and JAK3-selective jakinibs are under evaluation. Inhibition of another kinase, spleen tyrosine kinase (Syk) with R788 (fostamatinib), has also shown clinical efficacy in RA.

Anti-T-Cell Therapy

The contribution of costimulatory signals (“signal 2”) transduced through CD28 to priming and activation of naïve T cells and amplification of cytokine gene expression and proliferative responses has provided a rationale for testing costimulatory blockade in patients. This has been achieved using a nondepleting humanized IgG1-CTLA-4 fusion protein that prevents CD80 and CD86 (expressed on antigen-presenting cells) from engaging CD28 (but also CTLA-4) expressed on T cells. Initial studies confirmed that the agent was safe and well tolerated.[50] When administered as an intravenous infusion on days 1, 14, and 28, and then monthly thereafter at a 10-mg/kg dose in combination with methotrexate, 60% of patients achieved a 20% improvement in tender or swollen joint counts as well as 20% improvement in three of the other five ACR criteria (ACR20) compared with 35.3% of controls; ACR50 responses were 36.5% versus 11.8%, respectively. CTLA-4-Ig (licensed as abatacept) has since been shown to inhibit radiographic progression and structural damage by ~50%, and it is also effective in treating those patients who have had inadequate responses to TNF blockade as well as to methotrexate. Abatacept can also be administered subcutaneously on a weekly basis with efficacy comparable to that of the intravenous preparation. From an immunological standpoint the data are interesting in several respects. First, head-to-head studies indicate that the kinetics of response are remarkably similar to those of anti-TNF. This likely reflects the significant number of costimulation-dependent T cells actively participating in the ongoing inflammatory response. Clinical improvement may continue beyond 12 months. Second, recent data suggest that inhibition of CD28 signals does reduce the number of Tregs, as might have been predicted from mouse studies.

Anti-B-Cell Therapy

Rituximab is a humanized monoclonal antibody that recognizes human CD20, a 33- to 37-kDa membrane-associated phosphoprotein expressed on pre-B, immature, and mature B cells but not on plasma cells. It was initially developed, and then licensed, for the treatment of non-Hodgkin lymphoma. While CD20 ligation promotes B-cell activation, differentiation, and cell cycle progression, the function of CD20 is still poorly understood. The therapeutic effects of anti-CD20 are related to B-cell depletion, which can vary between patients due to antibody-dependent cell cytotoxicity, complement-mediated cell lysis, and/or triggering of intracellular pathways for apoptotic cell death.[51]

The initial open-label studies in RA patients combined rituximab with cyclophosphamide and steroids, and so the precise contribution of rituximab to the clinical response was unclear. A pivotal placebo-controlled trial of rituximab therapy randomized 161 patients with RF-positive RA to compare the efficacy and safety of methotrexate alone (standard therapy) versus methotrexate plus rituximab (1000 mg on days 1 and 15), rituximab alone, or rituximab plus cyclophosphamide.[51] All groups received a 17-day course of steroids, which may facilitate B-cell depletion. Up to 43% of patients receiving the combination of rituximab and methotrexate achieved 50% improvement in clinical and laboratory parameters after 24 weeks (ACR50), and this was at least as good as the rituximab/cyclophosphamide combination (41% achieving ACR50) and superior to methotrexate (13%) or rituximab alone (33%). The overall incidence of infection was similar across treatment groups, but serious infections were reported in rituximab-treated patients. Postmarketing surveillance has documented progressive multifocal leukoencephalopathy in a small number of patients.

Detailed analysis of this and subsequent studies in RA patients indicate that the B-cell depletion, which may depend on *FCGR3* genetic variants, is profound (<2%), with no major differences in infection rates between groups with the possible exception of lower respiratory tract infections. Repopulation occurs at a mean of 8 months after treatment, comprises immature IgD$^+$ CD38$^+$CD27$^-$CD5$^+$ B cells, and is associated with increased serum B cell–activating factor (BAFF) levels. Early relapse is associated with reconstitution of CD27$^+$ memory B cells. The effects on serum immunoglobulin levels are modest with levels remaining in the normal range, unless patients undergo repeated cycles of B-cell depletion. Selective and rapid decreases in pathogenic IgM, IgG, and IgA RF (~60%) and IgG anti-CCPs autoantibodies have been documented; these changes are more striking in rituximab responders. Antibody titers to tetanus toxoid or to pneumococcal capsular polysaccharides were only modestly affected (~25%). These changes are intriguing given that terminally differentiated plasma cells do not express CD20. One possibility is that germinal center and marginal zone B cells may be resistant to anti-CD20. B-cell depletion is surprisingly well tolerated; however, in contrast to other biological agents, it may induce long-lasting immunocompromise.

Future Prospects for Therapy

There remain unmet needs in the treatment of RA. Principal among these are the fact that until relatively recently, treatment has been considered lifelong for the majority of patients, imposing greater risk of toxicity and, as the immune system senesces, increased risk of infection or lymphoproliferative disease. There is little doubt that early treatment with tight control offers the best outcomes. Indeed, several synthetic and biological DMARDs are now being tested in at-risk subjects to establish whether they can delay or even prevent development of clinically apparent disease. A better appreciation of the profile of RA-specific autoantigens has prompted early- phase studies to establish the safety and tolerability of peptide immunotherapy delivered by autologous DCs. From an immunological perspective, there remains a pressing need to develop immunological tools or immune biomarkers that can redefine disease subsets and measure effector and regulatory cell subsets. This must be achieved with technology that is readily accessible to routine clinical laboratories and that uses methods that will inform disease activity states at a biological level. While such tools may provide better insights into disease pathogenesis, they can also be adapted to monitor the impact of therapeutic intervention, whether this turns out to be cell-based therapy or the application of novel immune modulators. It is possible that similar approaches to developing immune signatures could be used in the future to identify those ACPAs$^+$ arthralgia patients who have a genetic predisposition to RA and are at highest risk of developing the disease.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Snorrason E. Landre-Beauvais and his goutte asthénique primitive. Acta Med Scand 1952;142:115.
2. Harris ED. Rheumatoid arthritis; pathophysiology and implications for therapy. N Engl J Med 1990;322:1277.
3. Silman AJ, Hochberg MC. Epidemiology of rheumatic diseases. Oxford: Oxford University Press; 1993.
4. Frisell T, Saevarsdottir S, Askling J. Family history of rheumatoid arthritis: an old concept with new developments. Nat Rev Rheumatol 2016;12:335.
5. van der Woude D, Houwing-Duistermaat JJ, Toes RE, et al. Quantitative heritability of anti-citrullinated protein antibody-positive and anti-citrullinated protein antibody-negative rheumatoid arthritis. Arthritis Rheum 2009;60:916.
6. Schönland SO, Lopez C, Widmann T, et al. Premature telomeric loss in rheumatoid arthritis is genetically determined and involves both myeloid and lymphoid cell lineages. Proc Natl Acad Sci U S A 2003;100(23):13471–6.
7. Spector TD, Roman E, Silman AJ. The pill, parity, and rheumatoid arthritis. Arthritis Rheum 1990;33:782.
8. Karlson EW, Lee IM, Cook NR, et al. A retrospective cohort study of cigarette smoking and risk of rheumatoid arthritis in female health professionals. Arthritis Rheum 1999;42:910.
9. Scott IC, Seegobin SD, Steer S, et al. Predicting the risk of rheumatoid arthritis and its age of onset through modelling genetic risk variants with smoking. PLoS Genet 2013;9:e1003808.
10. Padyukov L, Silva C, Stolt P, et al. A gene-environment interaction between smoking and shared epitope genes in HLA-DR provides a high risk of seropositive rheumatoid arthritis. Arthritis Rheum 2004;50:3085.
11. Lundberg K, Wegner N, Yucel-Lindberg T, et al. Periodontitis in RA—the citrullinated enolase connection. Nat Rev Rheumatol 2010;6:727.
12. Wu HJ, Ivanov II, Darce J, et al. Gut-residing segmented filamentous bacteria drive autoimmune arthritis via T helper 17 cells. Immunity 2010;32:815.
13. Zhang X, Zhang D, Jia H, et al. The oral and gut microbiomes are perturbed in rheumatoid arthritis and partly normalized after treatment. Nat Med 2015;21:895.
14. Gregersen PK, Silver J, Winchester RJ. The shared epitope hypothesis. An approach to understanding the molecular genetics of susceptibility to rheumatoid arthritis. Arthritis Rheum 1987;30:1205.
15. Raychaudhuri S, Sandor C, Stahl EA, et al. Five amino acids in three HLA proteins explain most of the association between MHC and seropositive rheumatoid arthritis. Nat Genet 2012;44:291.
16. Viatte S, Plant D, Han B, et al. Association of HLA-DRB1 haplotypes with rheumatoid arthritis severity, mortality, and treatment response. JAMA 2015;313:1645.
17. Huizinga TW, Amos CI, van der Helm-van Mil AH, et al. Refining the complex rheumatoid arthritis phenotype based on specificity of the HLA-DRB1 shared epitope for antibodies to citrullinated proteins. Arthritis Rheum 2005;52:3433.
18. Scally SW, Petersen J, Law SC, et al. A molecular basis for the association of the *HLA-DRB1* locus, citrullination and rheumatoid arthritis. J Exp Med 2013;210:2569.
19. Eyre S, Bowes J, Diogo D, et al. High-density genetic mapping identifies new susceptibility loci for rheumatoid arthritis. Nat Genet 2012;44:1336.
20. Burn GL, Svensson L, Sanchez-Blanco C, et al. Why is PTPN22 a good candidate susceptibility gene for autoimmune disease? FEBS Lett 2011;585:3689–98.
21. Rawlings DJ, Dai X, Buckner JH. J Immunol 2015;194(7):2977–84. doi:10.4049/jimmunol.1403034. Review. PMID: 25795788.
22. Manzo A, Bombardieri M, Humby F, et al. Secondary and ectopic lymphoid tissue responses in rheumatoid arthritis: from inflammation to autoimmunity and tissue damage/remodeling. Immunol Rev 2010;233:267.

22. de Hair MJ, van de Sande MG, Ramwadhdoebe TH, et al. Features of the synovium of individuals at risk of developing rheumatoid arthritis: implications for understanding preclinical rheumatoid arthritis. Arthritis Rheumatol 2014;66:513.
23. van der Pouw Kraan TC, van Gaalen FA, Kasperkovitz PV, et al. Rheumatoid arthritis is a heterogeneous disease: evidence for differences in the activation of the STAT-1 pathway between rheumatoid tissues. Arthritis Rheum 2003;48:2132.
24. Vicente R, Noël D, Pers YM, et al. Deregulation and therapeutic potential of microRNAs in arthritic diseases. Nat Rev Rheumatol 2016;12:211.
25. Sacre SM, Drexler SK, Andreakos E, et al. Could toll-like receptors provide a missing link in chronic inflammation in rheumatoid arthritis? Lessons from a study on human rheumatoid tissue. Ann Rheum Dis 2007;66:81.
26. Pettit AR, Thomas R. Dendritic cells: the driving force behind autoimmunity in rheumatoid arthritis? Immunol Cell Biol 1999;77:420.
27. Szekanecz Z, Kim J, Koch AE. Chemokines and chemokine receptors in rheumatoid arthritis. Semin Immunol 2003;15:15.
28. Kollias G, Papadaki P, Apparailly F, et al. Animal models for arthritis: innovative tools for prevention and treatment. Ann Rheum Dis 2011;70:1357.
29. James EA, Rieck M, Pieper J, et al. Citrulline-specific Th1 cells are increased in rheumatoid arthritis and their frequency is influenced by disease duration and therapy. Arthritis Rheumatol 2014;66:1712.
30. van Venrooij WJ, Vossenaar ER, Zendman AJ. Anti-CCP antibodies: the new rheumatoid factor in the serology of rheumatoid arthritis. Autoimmun Rev 2004;3(Suppl. 1):S17.
31. Klareskog L, Rönnelid J, Lundberg K, et al. Immunity to citrullinated proteins in rheumatoid arthritis. Annu Rev Immunol 2008;26:651.
32. Klarenbeek PL, de Hair MJ, Doorenspleet ME, et al. Inflamed target tissue provides a specific niche for highly expanded T-cell clones in early human autoimmune disease. Ann Rheum Dis 2012;71:1088.
33. Salmon M, Scheel-Toellner D, Huissoon AP, et al. Inhibition of T cell apoptosis in the rheumatoid synovium. J Clin Invest 1997;99:439.
34. Cope A, Le Friec G, Cardone J, et al. The Th1 life cycle: molecular control of IFN-γ to IL-10 switching. Trends Immunol 2011;32:278.
35. Evans HG, Roostalu U, Walter GJ, et al. TNF-α blockade induces IL-10 expression in human CD4+ T cells. Nat Commun 2014;5:3199.
36. McInnes IB, Schett G. Cytokines in the pathogenesis of rheumatoid arthritis. Nat Rev Immunol 2007;7:429.
37. Miyara M, Gorochov G, Ehrenstein M, et al. Human FoxP3(+) regulatory T cells in systemic autoimmune diseases. Autoimmun Rev 2011;10(12):744–55.
38. Walter GJ, Fleskens V, Frederiksen KS, et al. Phenotypic, functional, and gene expression profiling of peripheral CD45RA+ and CD45RO+ CD4+CD25+CD127(low) Treg cells in patients with chronic rheumatoid arthritis. Arthritis Rheumatol 2016;68:103.
39. Schett G, Hayer S, Zwerina J, et al. Mechanisms of disease: the link between RANKL and arthritic bone disease. Nat Clin Pract Rheumatol 2005;1:47.
40. Harre U, Georgess D, Bang H, et al. Induction of osteoclastogenesis and bone loss by human autoantibodies against citrullinated vimentin. J Clin Invest 2012;122:1791–802.
41. Hayashi M, Nakashima T, Taniguchi M, et al. Osteoprotection by semaphorin 3A. Nature 2012;485:69.
42. Arnett FC, Edworthy SM, Bloch DA, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum 1988;31:315.
43. Aletaha D, Neogi T, Silman AJ, et al. 2010 Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Arthritis Rheum 2010;62:2569.
44. Bos WH, Wolbink GJ, Boers M, et al. Arthritis development in patients with arthralgia is strongly associated with anti-citrullinated protein antibody status: a prospective cohort study. Ann Rheum Dis 2010;69: 490.
45. Grigor C, Capell H, Stirling A, et al. Effect of a treatment strategy of tight control for rheumatoid arthritis (the TICORA study): a single-blind randomised controlled trial. Lancet 2004;364:263.
46. Goekoop-Ruiterman YP, de Vries-Bouwstra JK, Allaart CF, et al. Clinical and radiographic outcomes of four different treatment strategies in patients with early rheumatoid arthritis (the BeSt study): a randomized, controlled trial. Arthritis Rheum 2005;52:3381.
47. Feldmann M, Maini RN. Anti-TNF alpha therapy of rheumatoid arthritis: what have we learned? Annu Rev Immunol 2001;19:163.
48. Kishimoto T. Interleukin-6: from basic science to medicine—40 years in immunology. Annu Rev Immunol 2005;23:1.
49. Ghoreschi K, Laurence A, O'Shea JJ. Janus kinases in immune cell signaling. Immunol Rev 2009;228:273.
50. Kremer JM, Westhovens R, Leon M, et al. Treatment of rheumatoid arthritis by selective inhibition of T-cell activation with fusion protein CTLA4Ig. N Engl J Med 2003;349:1907.
51. Edwards JC, Cambridge G. B-cell targeting in rheumatoid arthritis and other autoimmune diseases. Nat Rev Immunol 2006;6:394.

MULTIPLE-CHOICE QUESTIONS

1. Which of the following statements is correct regarding autoantibodies in rheumatoid arthritis (RA)?
   A. Antibodies to citrullinated peptides are highly specific for rheumatoid arthritis.
   B. Citrullination is found only in joint tissues.
   C. Citrullination requires the presence of the amino acid asparagine.
   D. Citrullination generates antigenic epitopes by posttranslational modification of pathogen-derived peptides.
   E. Antibodies to citrullinated peptides are generated after onset of disease.

2. Which of the following statements is correct regarding the shared epitope?
   A. The shared epitope refers to a common bacterial antigen.
   B. The shared epitope refers to a common viral antigen.
   C. Antibodies to the shared epitope are characteristic for RA.
   D. The shared epitope refers to a common sequence in a subset of major histocompatibility complex class II alleles.
   E. The shared epitope causes rheumatoid arthritis.

3. Which of the following characteristics are NOT typical features of individuals at high risk of developing RA?
   A. A family history of RA
   B. Low body mass index (BMI)
   C. Inflammatory joint pain
   D. A history of tobacco exposure
   E. Low alcohol intake

53

Juvenile Idiopathic Arthritis

Randy Q. Cron, Peter Weiser, Timothy Beukelman

Juvenile idiopathic arthritis (JIA) is a collection of conditions that manifest with chronic arthritis in childhood. By definition, JIA occurs prior to the 16th birthday (juvenile), has no known causes (idiopathic), and has evident chronic (>6 weeks' duration) swelling/inflammation of a joint or joints (arthritis).[1] The nomenclature and subcategorization of JIA have been published, with a second revision in 2004, partly in an attempt to unify the American (juvenile rheumatoid arthritis [JRA]) and European (juvenile chronic arthritis [JCA]) classification schemes.[1] Removal of the word "rheumatoid" was also important to distinguish these unique disorders from adult rheumatoid arthritis (RA) (Chapter 52). The new classification of JIA also includes the human leukocyte antigen (HLA)-B27–associated spondyloarthropathies (Chapter 57), which occur frequently during childhood (Table 53.1).

Because JIA is a diagnosis of exclusion, other causes of childhood chronic arthritis need to be ruled out. For example, sarcoidosis (Chapter 73), Sjögren disease (Chapter 54), systemic lupus erythematosus (SLE) (Chapter 51), Lyme disease (Chapter 28), dermatomyositis (Chapter 56), and a variety of vasculitides (Chapters 58 and 59) may all present as chronic arthritis in childhood. JIA, by definition, is idiopathic; however, chronic arthritis as part of inflammatory bowel disease (IBD) (Chapter 75) or associated with psoriasis (Chapter 64) is categorized under the JIA umbrella. When considering the seven categories of JIA together, the estimated prevalence of JIA is roughly 1 in 1000 children, with rates varying in different regions of the world, likely because of genetic risk factors and/or environmental triggers. Moreover, the relative rates of JIA categories vary in different regions of the world (*e.g.,* oligoarticular JIA is very common in Scandinavia, and enthesitis-related JIA is more typical in Latin America).

ETIOLOGY AND PATHOGENESIS

Genetic Contribution

Because JIA is an eclectic group of disorders with joint inflammation, is far less common than RA, and is classified into unique categories based largely on subjective clinical phenotypes, understanding its genetic causes is extremely difficult. Indeed, the relatively arbitrary distinction of having ≤4clinically arthritic joints to be considered as oligoarticular JIA is influenced by whether temporomandibular joint (TMJ) arthritis (or in essence any other suspected joint involvement) is screened for with magnetic resonance imaging (MRI) because clinical examination is neither sensitive nor specific for diagnosing TMJ arthritis.[2] Distinctions based on the types and location of joints manifesting arthritis may prove useful in identifying more homogeneous subtypes. Thus the criteria will likely continue to evolve and benefit from more precise clinical, laboratory, and genetic distinctions. One such suggested reclassification of JIA categories is based on shared similarities among children with JIA who are positive for antinuclear antibody (ANA)[3] regardless of total joint count, and a recent proposal used probabilistic principal component analysis of a large set of biological and clinical data to subdivide patients with JIA.[4] Moreover, there have been several attempts to characterize the genetics of the JIA categories by using a variety of approaches.[5]

Polygenic Disorder

Like most autoimmune conditions, JIA is a polygenic disorder with ≥20 potential genes contributing to the phenotypes of chronic arthritis presenting during childhood. An exception is the systemic JIA (sJIA) category, which resembles an autoinflammatory disorder rather than a solely autoimmune process.[6] Unlike autoimmunity, which is believed to be the result of an imperfect adaptive immune system, autoinflammatory conditions are thought to result from genetic perturbations in the innate immune response. Many autoinflammatory recurrent fever syndromes have been identified as resulting from single gene defects, such as pyrin in familial Mediterranean fever (FMF), and tumor necrosis factor (TNF) receptor in *TNFR1*-associated periodic syndrome (TRAPS) (Chapter 60).[6] Whether sJIA is the result of a single gene defect or is one of many potential single gene defects in the same immune pathway is currently unknown. However, a common complication of sJIA is the life-threatening condition of macrophage activation syndrome (MAS), which resembles familial hemophagocytic lymphohistiocytosis (fHLH) (Fig. 53.1). MAS may be present in ≈50% of children with sJIA in either an overt or occult/subclinical fashion.[7] Recently, genetic polymorphisms and heterozygous (single copy) mutations in genes associated with fHLH when present as homozygous defects have been identified in children with sJIA and overt MAS.[7] Protein products of these genes, including perforin 1 and MUNC 13-4, are critical in the pathway mediating cytolysis by CD8 T cells and natural killer (NK) cells (Chapter 17). This cytolysis is crucial to the ability to shut down an immune response following control of infection. Defects in this pathway can result in a hyperinflammatory state (cytokine storm) leading to pancytopenia, coagulopathy, and multisystem organ failure.[7] The most common condition resulting in MAS during childhood is sJIA. Therefore sJIA may be genetically related to fHLH but associated with

TABLE 53.1 Clinical and Laboratory Features of the Juvenile Idiopathic Arthritis (JIA) Categories

	JIA					
Classifications	**Former Juvenile Rheumatoid Arthritis (JRA)**				**Spondyloarthropathy**	
JIA categories/ features	Systemic	RF^- polyarticular	RF^+ polyarticular	Oligoarticular (persistent & extended)	Psoriatic (early and late onset)	Enthesitis-related arthritis (ERA) (including ankylosing spondylitis (AS) and inflammatory bowel disease [IBD])
Human leukocyte antigen (HLA)	DRB1*11	DRB1*08	DRB1*04	A2	DRB1*01	B27
Gender	Equal	$F \gg M$	$F \gg M$	$F \gg M$	$F > M$	$M \gg F$
Age at onset	Peak 2 years	Dual peaks	Teenage	1–3 years	Dual peaks	Teenage
Antinuclear antibody (ANA)	Rare	Yes	Yes	Yes	Yes	Rare
Uveitis	Rare	Yes	Yes	Yes	Yes	Non-silent
Temporomandibular joint (TMJ)	Yes	Yes	Yes	Yes	Yes	Yes
Enthesitis	No	No	No	No	Older age	Yes
Arthritis	Erosive	Symmetric	Erosive	Mixed	Dactylitis	Axial

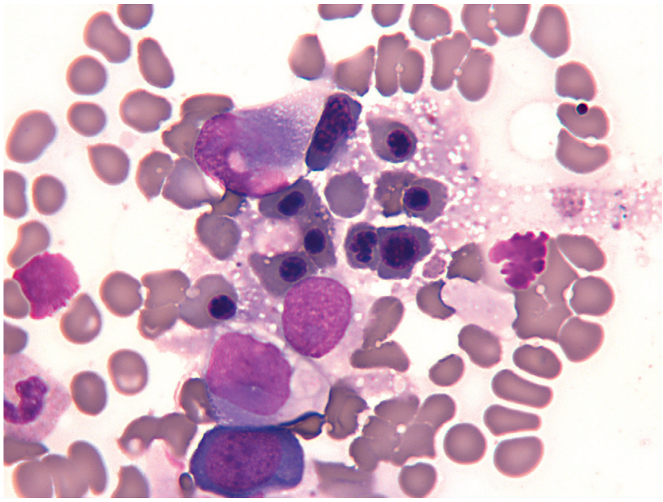


FIG 53.1 **Hemophagocytosis as Part of Macrophage Activation Syndrome (MAS).** A centrally located, vacuolated histiocyte is pictured engulfing numerous nucleated immune cells and non-nucleated mature red blood cells. Wright stain at ×198 magnification. (Courtesy of Dr. David Kelly.)

heterozygous, rather than homozygous, mutations in genes critical for cytolysis.[7]

HLA Associations

For all other JIA categories, polygenic influences, including the major histocompatibility complex (MHC), contribute to disease pathology. The MHC is complex and densely packed (>200 genes) with immune-associated genes (Chapter 5). The MHC is also the most polymorphic region of the human genome and gives rise to MHC class I and class II genes, complement proteins, TNF, and others. The critical role of MHC proteins in preventing autoimmunity by shaping the T-cell repertoire (Chapter 8) likely explains why the MHC is the most consistently and strongly associated genetic locus for most JIA categories. However, weak MHC associations and lack of gender predilection in the development of sJIA are additional arguments to consider sJIA as autoinflammatory rather than autoimmune.

In contrast, enthesitis-related arthritis (ERA) and psoriatic JIA share one of the strongest known MHC associations, HLA-B27, for any described autoimmune disorder. HLA-B27 is present in 70–90% of Caucasian children with ERA. Inheritance of ERA follows an autosomal dominant pattern and is the only JIA category likely to have similarly afflicted first-degree relatives. The pathophysiological explanation for the association of HLA-B27 remains unknown, but hypotheses vary from molecular mimicry of pathogens presented by HLA-B27 (Chapter 50) to the unfolded protein response.[8] The remaining JIA categories have all been linked to various other MHC proteins to lesser degrees (see Table 53.1).

Non-HLA Associations

Genes outside of the MHC have also been linked to developing JIA. Genetic approaches have highlighted the importance of genetic differences in the development of JIA, but at most, 11% of the contribution can be linked to the MHC.[5] Recent developments in high-throughput genetic sequencing combined with the identification of densely present restriction fragment length polymorphisms (RFLPs) throughout the human genome have allowed for powerful new genetic approaches. These include genome-wide association studies (GWAS) designed to identify genes associated with a variety of disorders, including autoimmune diseases, such as JIA.[9] To date, several genes have been reported to be associated with JIA.[5]

Strong evidence that genetic factors contribute to JIA susceptibility include twin and family studies. Data from a JIA national registry revealed that monozygotic twin pairs have a higher-than-expected proportion of twins with JIA. Moreover, siblings of children with JIA have as high as a 30-fold increased risk of JIA compared with the general population. Interestingly, siblings with JIA typically share the same JIA category, age of onset, and disease course.[5] Polymorphisms in MHC class II genes have been estimated to account for <20% of the recurrence risk of JIA in siblings and thus support the concept that JIA is a complex genetic trait. JIA likely shares general or common autoimmunity gene risk factors with many autoimmune disorders but may also be influenced by specific JIA risk-associated genes.[9]

Using linkage and association studies, researchers have identified a variety of potentially JIA-associated genes. However, outside of the HLA genes, only a small percentage of these genes have been independently confirmed by other investigators. A few of the independently confirmed JIA-associated genes/gene products include *PTPN22* (a phosphatase involved in inhibition of T-cell activation), *WISP3* (a signaling protein), interleukin-1α (a proinflammatory cytokine), TNF (another proinflammatory cytokine), macrophage migratory inhibitory factor (MIF), and *SLC11A1* (a resistance factor to intracellular pathogens in macrophages).[5] Thus gene products linked to both innate and adaptive immune responses likely contribute to JIA disease susceptibility. Ongoing GWAS may help validate the importance of these genes and may identify other candidate gene risk factors for JIA in the near future. Recently, epigenetic (modifications to DNA nucleotides rather than changes in the DNA sequence) risk factors that confer susceptibility to JIA are beginning to be explored.[10]

ENVIRONMENTAL FACTORS

In addition to genetic factors, there are likely a variety of environmental triggers for developing JIA in genetically susceptible hosts. A number of infectious agents have been explored as risk factors for development of JIA. Perhaps two of the best studied have been parvovirus and rubella virus infections. However, replication of these associations has been difficult. In addition, some scientists have suggested that heat shock proteins, resulting from cells undergoing environmental stress, may contribute to the development of JIA.[11] There is also a clear link between gut pathogens and potentially commensal organisms (gut microbiome) and the development of HLA-B27–associated spondyloarthropathies.[12] Last, prior to the identification of *Borrelia* as a cause of childhood arthritis, Lyme disease associated arthritis was difficult to distinguish from oligoarticular JIA.[13] Perhaps because there a variety of potential environmental triggers for JIA, combined with the number of different JIA categories, no single environmental trigger has been conclusively identified as contributing to the development of JIA.

IMMUNE ABNORMALITIES

Autoantibodies

As JIA is considered an autoimmune disease, except for sJIA as discussed earlier, there have been a variety of explorations into the role of various components of the immune system involved in JIA pathogenesis. The importance of the immune system in JIA pathology is highlighted by the increased incidence of childhood chronic arthritis among children with various immunodeficiencies. For example, children with immunoglobulin A (IgA) deficiency are at increased risk of developing chronic arthritis. In contrast, the presence of specific autoantibodies is associated with various forms of JIA. ANAs are present in up to 40% of patients with JIA, particularly those with oligoarticular JIA and are associated with silent uveitis (Chapter 74).[14] Similarly, antibodies to the nuclear oncoprotein DEK have been associated with uveitis as well as joint inflammation in children with JIA. IgM rheumatoid factor (RF) is present in a smaller subset of children with polyarticular JIA and is associated with a more aggressive/erosive form of arthritis as in adults with rheumatoid factor–positive (RF$^+$) RA (Chapter 52).[1] More recently, identification of anti–cyclic citrullinated peptide (CCP) antibodies have been found in a partially overlapping subset of children with RF$^+$ poly-JIA.[15] Thus a variety of autoantibodies are associated with JIA.

T-Helper Cells

T lymphocytes are also thought to play a major role in the development of JIA as evidenced by their relative predominance among mononuclear cells in synovial fluid of chronically inflamed joints in children with JIA. CD4 T-helper (Th) cells have been categorized into a variety of cytokine-producing subsets (Chapter 16). CD4 Th1 cells characterized by the production of interferon-γ (IFN-γ) have been identified in chronically inflamed joints in children with JIA, whereas IL-4–producing Th2 CD4 T cells are more commonly involved in the joints in oligoarticular JIA (generally less aggressive arthritis) than in polyarticular JIA. More recently, two relatively newly described CD4 Th subsets, regulatory T cells (Tregs) and Th17 cells, are being examined as potential players in JIA pathology.

CD4$^+$, CD25^{high} Tregs are characterized by the transcription factor, FoxP3, and by the ability to suppress immune activation (Chapter 18).[16] The ability of Tregs to suppress other T cells likely occurs by both cell-contact dependent (through the surface protein cytotoxic T lymphocyte antigen-4 [CTLA-4]) and independent (via suppressive cytokines) mechanisms.[16] Tregs secrete the antiinflammatory cytokines, IL-10 and transforming growth factor-β (TGF-β). Th17 cells, characterized by the transcription factor, retinoid orphan receptor (ROR)γT, produce the proinflammatory cytokine, IL-17.[17] IL-17 is thought to contribute to a variety of autoimmune disorders, including JIA. The balance between these two juxtaposed Th subsets may determine whether autoimmunity develops (Th17-dominant) or a state of immune tolerance to self (Chapter 12) persists (Treg-dominant). Indeed, recent studies have identified a predominance of Th17 cells and associated proinflammatory cytokines in the inflamed joints of children with JIA.[17] Thus a balance of proinflammatory cytokines and suppressive cytokines may dictate the expression of autoimmunity in the form of JIA.

Cytokines

Cytokines (Chapter 9), and particularly their inhibition, have taken a prominent status in the pathology and treatment of chronic arthritis, including JIA. The bench-to-bedside translation of anti-TNF therapy to the treatment of chronic arthritis has revolutionized the care of adults with chronic arthritis, as well as children with JIA. Inhibition of this proinflammatory cytokine in the circulation (via specific monoclonal antibodies [mAb] or receptor fusion proteins) rapidly and effectively treats most forms of JIA.[1] The one exception is sJIA, which may or may not respond to anti-TNF treatment. However, other proinflammatory cytokines, including IL-1, -6, and -18, are thought to be central to sJIA pathogenesis.[1] Indeed, serum from sJIA patients was shown to induce transcription of a variety of innate immunity genes, including IL-1, in normal peripheral blood mononuclear cells. Fortunately, novel therapies that target either IL-1 or IL-6 have proven highly successful in treating even the most severe forms of sJIA, including associated macrophage activation syndrome (MAS).[1,7]

Macrophage Activation Syndrome

The sometimes fatal complication MAS is most commonly seen in sJIA among rheumatic diseases. Clinically, MAS resembles many features of a sJIA disease flare-up, and it has been suggested

KEY CONCEPTS

Macrophage Activation Syndrome (MAS)

- MAS is present in up to 50% of children with systemic juvenile idiopathic arthritis (sJIA) in a subclinical or overt (10%) form.
- MAS manifests as fever, liver dysfunction, pancytopenia, central nervous system disturbance, hyperferritinemia, hemophagocytosis, and coagulopathy.
- MAS resembles hemophagocytic lymphohistiocytosis (HLH) and is thought to result from defects in perforin-mediated cytolysis by CD8 T cells and natural killer (NK) cells.
- Patients with sJIA and MAS have been noted to have NK-cell defects and mutations in perforin-1 and MUNC13-4 cytolytic pathway genes.
- MAS can be fatal if not recognized and treated early. Mainstays of therapy include high-dose corticosteroids and cyclosporine.
- Recently, interleukin-1 (IL-1) blockade with biological therapies has been found to be quickly and dramatically beneficial in treating MAS associated with sJIA.

that MAS may be inherent to sJIA disease pathology in up to half of all patients with sJIA.[7] MAS is likely part of the spectrum of HLH disorders. Primary HLH, or fHLH, typically presents in infancy following infection and results from homozygous mutations in genes involved in the cytolytic pathway employed by NK cells and CD8 T cells. Recent evidence suggests that patients with sJIA who have MAS have heterozygous defects in these same cytolytic pathway genes.[18] MAS can be triggered by a variety of infectious organisms, particularly members of the herpes virus family, but the precise role of infectious triggers of MAS in children with sJIA remains unknown. Nevertheless, the inability to effectively shut down an immune response via cytolytic mechanisms results in a "storm" of proinflammatory cytokines, such as IL-1, IL-6, IL-18, TNF, and IFN-γ. Mouse models of MAS/HLH have suggested that IFN-γ is the pivotal cytokine in MAS. In practical terms, inhibition of IL-1, and potentially of IL-6, has proven rather effective at treating MAS in children with sJIA.[1,7,18] It is quite remarkable how dampening of one critical cytokine can help restore the immune imbalance of multiple proinflammatory cytokines and rapidly reverse the life-threatening clinical scenario of MAS.

JIA CLINICAL SUBTYPES

As already mentioned, JIA is a group of chronic inflammatory joint diseases, which last for at least 6 weeks and commence prior to the age of 16 years with no identifiable cause. There is heterogeneity of clinical presentation and progression of the various subtypes, which is largely addressed by the International League Against Rheumatism (ILAR) classification schema (see Table 53.1). Nevertheless, as more information about genetic factors, response to medication, and subsequent outcomes becomes available from multicenter studies, revisiting the classification in the near future seems inevitable. A recent elegant report has suggested five distinct groups of patients categorized on the basis of clinical disease trajectories, all with subsets different from those defined by the ILAR classification.[4]

The current classification of JIA relies heavily on the number of the joints involved (≤4in oligoarticular JIA; ≥5 in polyarticular JIA, with or without the presence of RF).[1] Other categories of JIA are also classified on the basis of associated symptoms and signs, including fever, rash, and laboratory indicators of inflammation in sJIA, or associated diseases, such as psoriasis. There is a separate JIA category affecting boys more than girls—ERA, with inflammation of the entheses, attachments of the tendons, and ligaments to bone (see Table 53.1).[8] An important finding of recent years is that imaging (MRI and, in certain cases, ultrasonography) can identify ongoing inflammation in clinically asymptomatic joints. This has led to screening for TMJ inflammation primarily with MRI, as arthritis of this joint is frequently clinically silent but capable of resulting in facial dysmorphism from micrognathia in all JIA categories, including oligoarticular JIA.[2]

Oligoarticular JIA

Oligoarticular JIA is likely the most common category of JIA, affecting children who have 1–4 joints inflamed, most commonly knees, ankles, the TMJ, and fingers.[1] The archetype of this group is a preschool-aged, blonde-haired, blue-eyed girl who limps and has swollen knees. The diagnosis of arthritis may be delayed, as it is often painless. Because of this, by the time she is seen in the medical office, she may already have developed bony hypertrophy and limb length discrepancy, as chronic articular inflammation stimulates the osteoblasts of the nearby growth plates. In addition, there is often notable muscle wasting around the arthritic joint that can last into adulthood. The oligoarticular JIA category has the highest percentage of positive ANA blood tests and associated potentially damaging silent uveitis.[14] Monoarticular involvement calls for a careful differential diagnosis. JIA rarely presents with isolated hip involvement; toxic synovitis, septic hip, and malignancy all need to be considered. Oligoarticular JIA is also quite uncommon in middle and high school aged children, where the diagnosis of reactive arthritis, IBD-related arthritis, and Lyme disease should be entertained. The ILAR classification has a separate sub-category for children who develop additional joint involvement after the first 6 months based on clinical presentation, extended oligoarticular JIA,[1] which might be a variant of the RF-negative polyarticular category. Wrist involvement is considered to be a bad prognostic factor, as is the extended oligoarticular phenotype, and elevated laboratory indicators of inflammation.

Polyarticular JIA

Arthritis of ≥5 joints includes two main JIA categories (see Table 53.1) based on the presence of serum RF, an IgM antibody against the IgG Fc receptor.

RF-positive serum on two occasions at least 3 months apart is required for a child to be diagnosed with RF$^+$ polyarticular JIA. Joint involvement is typically bilateral and symmetrical, involving the small joints of hands and feet.[1] However, large joints and cervical involvement are often present. RF$^+$ polyarticular JIA usually presents in adolescent girls and is considered a form of early-onset adult RA.[1] It is a relatively infrequent category of JIA with <5 % of all patients with JIA classified in this category. Antibodies to CCP are much less common in children with polyarticular JIA than in adults,[15] but affected children usually have active arthritis for many years. The presence of RF predicts more destructive/erosive disease calling for early aggressive therapy. As in adult RA, arthritis of the wrists and fingers can lead to ulnar deviation and boutonniere and swan neck deformities. Destructive TMJ involvement is also common.[2] Similar to adult RA, RF$^+$ polyarticular JIA in children commonly has extraarticular manifestations, such as low-grade fever and occasionally rheumatoid nodules over bony surfaces.

RF-negative polyarticular JIA usually presents as asymmetrical involvement of the large joints, mostly knees, wrists, and ankles.

As in psoriatic JIA, small joint involvement tends to occur later in life. Another similarity between psoriatic JIA and RF^{-} polyarticular JIA is the bimodal distribution in preschool children and early adolescent patients.[1] There is more frequent silent uveitis in the former group, and this form of JIA is often difficult to distinguish from ERA in the latter. TMJ inflammation occurs frequently, leading to condylar damage and facial dysmorphism.[2]

Psoriatic Arthritis

Arthritis with concurrent psoriasis, or arthritis with two of three factors—dactylitis, psoriatic nail changes, or family history of a first-degree relative with psoriasis—comprises a separate category of JIA. There seems to be a bimodal distribution of age at onset for psoriatic JIA.[19] Preschool-aged children have mostly large joint involvement like those in the oligoarticular category but may also have dactylitis, whereas middle school–aged patients have JIA that resembles ERA with enthesitis, sacroiliac joint involvement (albeit milder), and even spondylitis.[8] In general, there is asymmetrical involvement of the joints, and if untreated, it will progress to polyarticular joint disease. Since up to 50% of the patients develop psoriatic skin findings several years after arthritis presentation, it is often difficult to diagnose this condition at onset. It is important to carefully examine children with JIA for dactylitis (tenosynovitis causing swelling of the digit beyond the joint capsule) and nail pits and onycholysis. Psoriatic JIA seems to be more resistant to therapy, and approximately 40% of children have active disease into adulthood while on medications. Insidious onset of anterior uveitis is more typical for the younger age group, whereas those with enthesitis have JIA that resembles the adult type of psoriatic arthritis with an associated HLA-B27 genotype and chronic symptomatic, often painful, eye disease (see Table 53.1).

Enthesitis-Related Arthritis

ERA affects boys more than girls[1] and may sometimes be a manifestation of IBD. Entheses are the attachments of the tendons, ligaments, or joint capsules to bone. These can be tender even in healthy children, but usually ≥3 tender entheses are associated with disease (see Table 53.1). ERA often occurs in boys 8 years of age and older.[1,8] They typically complain of joint pain related to playing sports, but there is also morning stiffness and pain that gets better during the day and worsens toward the end of the day or after engaging in a lot of activities. Many consider ERA a potential prelude to ankylosing spondylitis (Chapter 57), a HLA-B27–associated inflammatory condition resulting in irreversible fusion of the vertebrae. Recent therapeutic efforts are focusing on preventing the calcifying hypercorrection of inflamed vertebral edges using TNF blockade. Outcomes are still under investigation but appear promising with early treatment.

Systemic JIA

Approximately 10% of children with chronic arthritis belong to the sJIA category, also known as *Still disease.* The peak incidence of sJIA is ages 1–5 years, but it can present in adulthood. The ILAR criteria to classify sJIA require fever for 2 weeks, with at least three episodes of daily spiking (quotidian) fever, with at least one of the following: fleeting pink macular rash, arthritis, lymphadenopathy, and hepatosplenomegaly.[1] When febrile, children appear rather ill, and the rash is more prominent and can be evoked by contact (Koebner phenomenon). Typically, the fever subsides, and children are visibly better in the morning hours. High levels of indicators of systemic inflammation are typical at onset and may subside later in the disease. Arthritis is often very aggressive and frequently involves wrists, ankles, and knees but also causes ankylosis of the hip and neck leading to long-term damage and gait abnormalities. Occasionally, joint involvement begins months after fever onset, making sJIA diagnosis more difficult. The initial presentation mimics those of infections and malignancies, and this has to be taken into consideration, as sJIA is a diagnosis of exclusion. Fifty percent of children may develop (only 10% clinically overt) MAS, described previously.[7]

Laboratory Evaluation

There is no one laboratory indicator that will establish or rule out chronic arthritis *per se.* The complete blood count (CBC) is largely normal in oligoarticular involvement, as is the erythrocyte sedimentation rate (ESR). White blood cells (WBCs) are highly elevated in sJIA but are mostly within normal limits in other groups. Anemia of chronic disease presents as normocytic and normochromic and is often found in polyarticular involvement. In those cases, the ESR can also be elevated. Intermittent joint effusion of a single large joint with an elevated ESR necessitates further evaluation, and IBD should be considered, especially if there is a low serum albumin level and/or growth delay. Of note, elevated ESR on presentation predicts a worse outcome for those with the oligoarticular subtype. The platelet count, as a marker of inflammation, can be elevated in polyarticular disease and substantially so in sJIA.

Liver function tests are used for monitoring certain disease modifying antirheumatic drugs (DMARDs), such as methotrexate and leflunomide. They can be elevated as a result of prolonged, and frequently concomitant, use of nonsteroidal antiinflammatory drugs (NSAIDs).

As mentioned above, in 10% of patients, sJIA can progress to overt MAS. Ferritin, an acute-phase reactant, is a very sensitive indicator of this condition. A sudden drop of at least two cell lines in the CBC, a rising cross-reactive protein (CRP) level with decreasing ESR, elevated liver transaminases, prolonged prothrombin or partial thromboplastin times, high D-dimer levels, elevated triglycerides, and low fibrinogen should all alert caregivers about the likelihood of MAS in a child with sJIA.[7]

Although 75–85% of adult patients with RA have either a RF or CCP antibodies, <5% of patients with JIA have the RF, and those are mostly patients with early-onset RA, often teenage girls with symmetrical small joint involvement. Another rather frequently ordered test is the serum ANA titer. Similar to the RF, ANA is not suitable for screening, as it is of no diagnostic utility in either making or excluding a diagnosis of JIA.[20] ANA serves as a prognostic factor by identifying patients already diagnosed with JIA who have the highest risk for developing uveitis.[3,14] In addition, ANA levels may alert the clinician to the possibility of juvenile Sjögren disease or SLE being the etiology for the chronic arthritis.

The prevalence of the HLA-B27 antigen is 8% in the general Caucasian population but nearly 90% in the ankylosing spondylitis group.[8] HLA-B27 is useful to predict axial involvement in ERA, psoriatic JIA, and IBD-related arthritis but should be evaluated in patients with clinically established arthritis and/or enthesitis, rather than as a routine screening test during a workup for back pain.

Additional laboratory indicators may be helpful. Elevated serum lactate dehydrogenase and uric acid levels may indicate malignancy. Elevated angiotensin converting enzyme (ACE) and lysozyme levels are useful when considering sarcoidosis as the etiology of childhood arthritis and uveitis in early-onset sarcoidosis cases, but noncaseating granulomas seen on biopsy are much more revealing. As for imaging techniques, radiography can be helpful in ruling out structural changes resulting from disease processes other than arthritic disease. Nevertheless, periarticular osteopenia is commonly seen around an inflamed joint. The clinical presentation usually precedes the development of bony erosions, but sclerosis of the sacroiliac joint may indicate axial involvement in patients with ERA. The role of ultrasonography for diagnosis and monitoring of disease progression in pediatric patients is promising, but further establishment of normative data is still required. MRI, with and without intravenous contrast, may help identify synovitis, but even this imaging modality can sometimes yield false-positive results in certain joints in otherwise healthy children.[8,21,22]

DIFFERENTIAL DIAGNOSIS

A wide variety of conditions can mimic the symptoms and signs of JIA. Although both acute and chronic arthritis may present with joint pain, tenderness, swelling, and warmth, a single acutely involved joint at onset should be considered septic until proven otherwise. Septic arthritis frequently has an erythematous discoloration of skin, whereas JIA does not. Prompt evaluation with arthrocentesis is necessary, especially in indeterminate cases. Infection of the joint may lead not only to rapid destruction of the joint but also to systemic dissemination of infection. Specifically, *Kingella* species should be considered in painful monoarticular cases of young children when symptoms have sudden onset without prolonged morning stiffness.

Parainfectious arthritis, often resulting from viral disease, is usually short-lived and typically requires only NSAID treatment. In contrast, Lyme disease is characterized by a chronic extensively swollen joint(s), commonly the knee.[13] It appears several weeks to months after the usually unnoticed tick bite and should be considered in areas of high incidence (*e.g.,* northeastern United States). There are a few diseases associated with migratory arthritis, including malignancy, gonococcal arthritis, and rheumatic fever arthritis, the last having a typically very tender, red, hot joint persisting in each location for a couple of hours before moving to another.

As sJIA is a diagnosis by exclusion, multiple other etiologies of systemic inflammation should be considered. Fever and elevated WBCs, platelets, and ESR may accompany polyarteritis nodosa, Kawasaki disease, Henoch-Schönlein purpura, and other vasculitides. Ehrlichiosis and recurrent fever syndromes (*e.g.,* familial Mediterranean fever, TNF receptor–associated periodic fever syndrome) may also manifest as systemic inflammation, arthralgias, and rashes. A typically nonerosive symmetric arthritis may occur in SLE, often with cytopenias. Arthritis is also seen in related diseases (*e.g.,* Sjögren syndrome, mixed connective tissue disease [MCTD]), and evidence of antibodies to extractable nuclear antigens are common in SLE (anti-Smith, anti–double-stranded DNA) and MCTD (anti–ribonuclear protein [RNP]).

Patients presenting with joint pain only, without associated swelling or morning stiffness, most commonly have overuse syndromes or benign hypermobility syndrome.[23] Little league shoulder and elbow are rather common in middle school–aged baseball players; wrist pain occurs in gymnasts frequently. Pain in the hip, without morning stiffness, may indicate Legg-Calve-Perthes disease or slipped capital femoral epiphysis. Relentless knee pain without improvement in teenagers could be caused by osteochondritis dissecans. Other common causes of teenage knee pain include Osgood-Schlatter disease and patellofemoral syndrome.

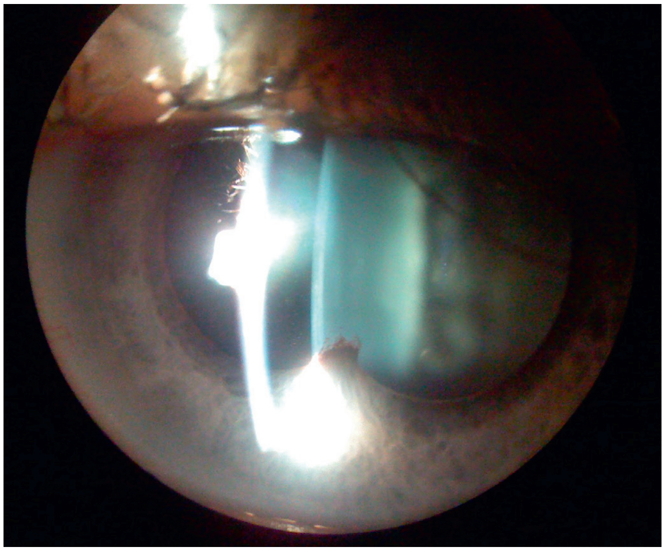


FIG 53.2 Posterior synechia, a complication of chronic anterior uveitis, is associated with several categories of juvenile idiopathic arthritis (JIA). The irregularities of the inner margins of the iris reflect fibrous adhesions between the iris and lens capsule. (Courtesy of Dr. Scott Olitsky.)

Clinically Silent Complications

There are two common findings that present insidiously during the course of JIA and yet can cause major damage/morbidity—uveitis and TMJ arthritis. Uveitis is a relatively common complication of JIA with potentially long-term morbidity. Risk factors for uveitis among patients with JIA include young age of onset, positivity for ANA, oligoarticular disease, and female sex.[14] It usually presents as iridocyclitis, but the choroid may also be affected. Although extremely rare in sJIA, approximately 20% of patients with oligoarticular JIA and 5% of patients with polyarticular JIA develop eye inflammation.[14] Some children with psoriatic JIA are also at risk of developing silent uveitis, especially in the ANA-positive subgroup.[19] Uveitis may lead to a great deal of morbidity, including cataracts, increased intraocular pressure, band keratopathy, and posterior synechiae (Fig. 53.2), with decreased vision developing in up to 40%.[14] The danger of most JIA-associated uveitis is its asymptomatic presentation, with the exception of symptomatic uveitis in children with ERA. Considering that the highest prevalence is in patients with oligoarticular JIA, which is most common in preschool and younger children, it is not surprising that many of the cases go unnoticed. A routine ophthalmological examination may fail to detect uveitis, and children need to have a formal slit-lamp examination to identify inflammatory cells.[14] The most common presenting signs are synechiae (an irregular iris border resulting from adhesions to the lens), hypopyon, and band keratopathy (see Fig. 53.2). Uveitis may develop many months or years after joint symptoms, and therefore close follow-up is warranted. ANA positivity and young age are associated with increased incidence; thus ANA-positive children with oligoarticular JIA need to be screened the most often, but those in other categories should be followed up on a set schedule as well.[14] Failure to do so and missed eye involvement may lead to the occurrence of cataracts, glaucoma, impaired vision, and even blindness.

KEY CONCEPTS

Temporomandibular Joint (TMJ) Arthritis

- TMJ arthritis is extremely common, present in up to 80% of children with juvenile idiopathic arthritis (JIA). It is typically asymptomatic and thus requires early screening with magnetic resonance imaging (MRI) with contrast.
- TMJ arthritis is often active despite therapy with disease-modifying antirheumatic drugs (DMARDs) and biologicals (*e.g.,* methotrexate plus tumor necrosis factor [TNF] inhibitors) and thus treatment with intraarticular long-acting corticosteroids or TNF inhibitors may help prevent mandibular growth damage and associated micrognathia and facial dysmorphology.

Another frequently asymptomatic complication of JIA is TMJ arthritis. TMJ arthritis in children with JIA has been recognized increasingly in recent years as a joint inflammation leading to silent destruction and facial deformity despite systemic therapy.[2] TMJ arthritis is quite common, with 40–80% of all patients with JIA patients affected.[2] The overall true prevalence is likely closer to the higher range, since not all children with JIA receive TMJ MRI screening, which frequently reveals synovial thickening (Fig. 53.3) at disease onset, and premicrognathic arthritis may be missed.[2,22] The highest rates of TMJ arthritis have been found in the extended oligoarticular JIA and RF-negative polyarticular JIA groups, as well as in children with upper extremity and neck involvement, as well as an elevated ESR.[2] Currently, intraarticular corticosteroid injection seems to be beneficial in certain cases, whereas TMJ arthritis often develops despite systemic use of methotrexate and TNF inhibitors.[2] TMJ arthritis needs to be recognized early in children with JIA so that it can be treated prior to growth disturbance.

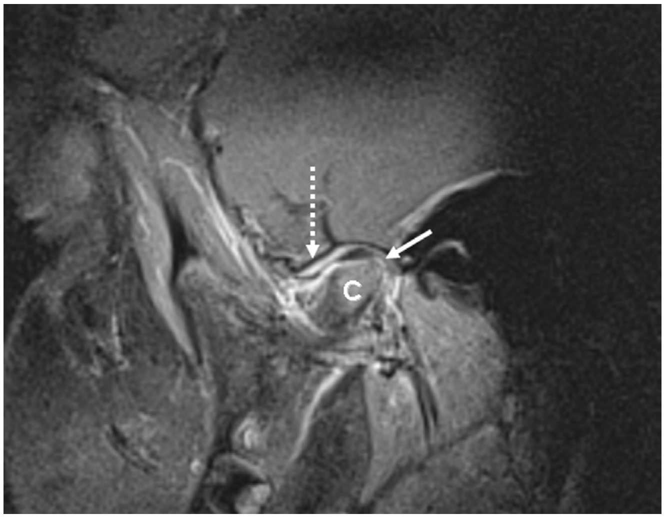


FIG 53.3 Acute and Chronic Temporomandibular Jaw (TMJ) Arthritis in a Child With Juvenile Idiopathic Arthritis (JIA). Synovial thickening and enhancement (long dashed arrow) and mandibular condyle ("C") flattening with contour irregularity/erosion (short arrow) are noted in this parasagittal postcontrast T1-weighted magnetic resonance imaging (MRI) image. (Courtesy of Dr. Dan Young.)

TREATMENT

Overview

Despite significant advances in the understanding of the pathogenesis of JIA, there are currently no curative treatments. JIA frequently persists into adulthood and may result in significant morbidity, including physical disability. The objective of treatment is to prevent disability and preserve normal growth and development while providing relief of symptoms and improved quality of life by controlling the inflammatory process.[1]

Over the past 15 years, remarkable advances have been made in the treatment of JIA.[1] Chief among these advances was the advent of targeted biological therapeutic agents (Table 53.2) (Chapter 89). These agents have been shown to be quite beneficial against active disease and are generally well tolerated.[1] The early initiation of biological therapeutic agents may, in fact, alter and improve the subsequent disease course.[1] These new breakthroughs have prompted pediatric rheumatologists to "invert the treatment pyramid," that is, to rapidly incorporate more effective therapeutic agents instead of slowly progressing to them in a stepwise fashion.[1]

THERAPEUTIC PRINCIPLES

Early Aggressive Therapy

- Accumulating evidence suggests that early aggressive therapy that includes targeted biological agents near the time of clinical diagnosis (during the "window of opportunity") may improve the future disease course.

In response to the growing number of treatment options for JIA and the advent of the biological therapeutics, the American College of Rheumatology (ACR) issued Recommendations for the Treatment of JIA in 2011.[24] These recommendations were developed by using a rigorous methodology to produce evidence-based and consensus-based guidance that reflected the current state of the field. The guidelines were updated in 2013 to include advances in the treatment of sJIA.[25]

With new and effective therapies continuing to be introduced to the therapeutic armamentarium, treatment goals have become elevated and more stringent. The current goal is to achieve a status of clinically inactive disease,[26] that is, the absence of any significant signs or symptoms of active arthritis. Although currently this is not possible for all children, until that becomes a reality, curative therapy for JIA (achievement of inactive disease status) for as many children as possible remains the goal.

Recent advances in the treatment of adults with inflammatory arthritis have elucidated some differences in the effectiveness of specific biological agents for specific forms of arthritis. For example, many non-TNF inhibitor biologicals, such as abatacept, rituximab, and tocilizumab, are highly effective in the treatment of RA but are far less effective in the treatment of ankylosing spondylitis. In contrast, some of the more recently introduced biological agents, such as the IL-17 inhibitor secukinumab, appear effective against ankylosing spondylitis but are likely less effective for RA.[27]

Despite these recent advances, the treatment of JIA is not currently strongly influenced by the distinct categories of JIA, with the exception of sJIA. For example, there are no specific therapies currently approved for the treatment of children with psoriatic arthritis as opposed to RF-negative polyarthritis. Accordingly, the discussion of treatment in this chapter will not detail all categories of JIA but will focus rather on the "treatment groups" as defined

TABLE 53.2 Biological Therapeutic Agents Used in the Treatment of Juvenile Idiopathic Arthritis (JIA)

Biological Target	Name	Structure	Frequency of Administration	Route of Administration
Tumor necrosis factor (TNF)	Etanercept	TNF receptor-immunoglobulin G (IgG) fusion protein	Twice weekly to weekly	Subcutaneous (SQ) injection
	Infliximab	Monoclonal antibody (chimeric)	Every 4–8 weeks	Intravenous (IV) infusion
	Adalimumab	Monoclonal antibody (humanized)	Every 1–2 weeks	SQ injection
	Golimumab	Monoclonal antibody (humanized)	Every 4 weeks Every 8 weeks	SQ injection IV infusion
	Certolizumab pegol	Monoclonal antibody (humanized and PEGylated)	Every 2 weeks	SQ injection
CD80/86	Abatacept	Cytotoxic T lymphocyte antigen-4 (CTLA-4–IgG fusion protein	Every 4 weeks Weekly	IV infusion SQ injection
Interleukin-1 (IL-1)	Anakinra	Receptor antagonist	Daily	SQ injection
	Canakinumab	Monoclonal antibody (humanized)	Every 8 weeks	SQ injection
	Rilonacept	Receptor-fusion protein	Weekly	SQ injection
IL-6 receptor	Tocilizumab	Monoclonal antibody (humanized)	Every 2–4 weeks Every 1–2 weeks	IV infusion SQ injection
CD20^{+} B cells	Rituximab	Monoclonal antibody	2 infusions 2 weeks apart, repeat every 6 months	IV infusion
IL-12 and IL-23	Ustekinumab	Monoclonal antibody	Every 12 weeks	SQ injection
IL-17A	Secukinumab	Monoclonal antibody	Every 4 weeks	SQ injection

by the 2011 ACR Recommendations for JIA treatment.[24] As our knowledge about pathophysiology and response to treatment continues to expand, differential therapeutic approaches for the disparate JIA phenotypes are eagerly anticipated.

Nonsteroidal Antiinflammatory Drugs

NSAIDs formed the foundation of the treatment of JIA for decades, and numerous NSAIDs have been shown to have beneficial effects. In the absence of significant numbers of head-to-head trials, it is believed that in general all NSAIDs are similarly effective,[1] although indomethacin is considered by some to be the most effective NSAID (especially in sJIA). One NSAID may be found to be more effective than another for a particular individual child.

In general, NSAIDs are not considered to be disease-modifying agents, that is, they are not felt to slow the progression of disease or prevent the appearance of radiographic damage. For this reason, monotherapy with NSAIDs initially began to decline with the advent of agents, such as methotrexate, that have been shown to modify the disease process. NSAIDs are frequently used for symptomatic relief, but most children who require daily chronic use of NSAIDs would likely benefit from the addition of systemic immunosuppression, such as from DMARDs. Gastrointestinal (GI) discomfort is a frequent adverse effect of NSAID therapy, although frank GI bleeding appears to occur at a lower incidence than in adults. Scarring pseudoporphyria of sun-exposed skin is another risk associated with NSAID use. The long-term cardiovascular effects of NSAIDs in children have not been studied.

Glucocorticoids

Similar to many rheumatological diseases, JIA has been shown to respond to treatment with glucocorticoids (Chapter 86). Intraarticular glucocorticoid injections typically result in a near-immediate decrease in inflammation that is maintained for many months.[28] Accordingly, intraarticular injections may form the foundation of therapy for children with mild or limited oligoarthritis. In addition, intraarticular injections may be effective in children with more extensive arthritis and in those who are receiving concurrent systemic therapy. Randomized trials have shown unequivocally that injected triamcinolone hexacetonide has the longest duration of effect.[26]

JIA also may be effectively treated with systemic glucocorticoids. They are frequently used in the treatment of systemic features of JIA and may form the foundation of the treatment for the sJIA category, although biological therapies may significantly lessen the need for this practice.[26] The use of systemic glucocorticoids for the treatment of synovitis in children with JIA is not an uncommon practice. However, the risks, benefits, and appropriate use of this approach are less clear. The recent 2011 ACR Recommendations for JIA treatment do not include this use of systemic glucocorticoids because of the near-complete absence of any published evidence.[24] In general, most pediatric rheumatologists would agree that the use of DMARDs is preferable to the use of moderate-dose or high-dose systemic glucocorticoids for the treatment of synovitis in JIA. The anticipated adverse effects of long-term use of moderate doses of glucocorticoids includes growth failure, osteoporosis, cataract formation, glaucoma, hyperglycemia, hypertension, avascular necrosis of bone, striae, and others.[26]

Nonbiological DMARDs

The use of DMARDs was introduced in the 1980s (Chapter 87). The most widely used and studied is methotrexate, which has been shown in randomized clinical trials to be efficacious in treatment of JIA.[1] Following these studies, methotrexate became the cornerstone of therapy for many children with JIA. Methotrexate is typically administered weekly through either the oral or the subcutaneous route, although studies have shown subcutaneous methotrexate to be better absorbed and more effective.[1]

Methotrexate may be associated with several typically minor adverse effects, such as nausea and fatigue. Occasionally, methotrexate may cause liver toxicity, necessitating periodic measurement of serum aminotransferase levels for routine monitoring.

Leflunomide has been shown to be slightly less efficacious than methotrexate in the treatment of JIA.[1] It may serve as an alternative therapy for children who are intolerant of methotrexate. Sulfasalazine is used, to varying degrees, by pediatric rheumatologists and may be of particular benefit to children with JIA of the ERA category.[1] Hydroxychloroquine monotherapy has been demonstrated inefficacious in treating JIA.[1]

Biological DMARDs

The use of biological DMARDs for the treatment of JIA began in the late 1990s (Chapter 89). Etanercept, a TNF inhibitor, was the first studied and was shown to be efficacious in a randomized clinical trial.[1] Additional TNF inhibitors have been introduced and used in the treatment of JIA; adalimumab was also shown to be efficacious in a randomized clinical trial.[1] Notable differences have been discovered regarding the effectiveness of the TNF receptor fusion protein (etanercept) and the mAbs (adalimumab, infliximab, and others). mAb TNF inhibitors have been shown to be effective against two important JIA-associated conditions—anterior uveitis and IBD. Receptor fusion proteins have been shown to be far less effective in treating these conditions. The precise mechanism for these differences in treatment effectiveness is not clear but may be related to the ability of etanercept to bind lymphotoxin or the ability to bind surface membrane-bound TNF.[29]

Because the TNF inhibitors are large proteins, they must be administered parenterally, either by subcutaneous injection or intravenous infusion (see Table 53.2). TNF inhibitors are not generally associated with common medication adverse effects, such as headache or nausea, although they may result in injection site or infusion reactions. There appears to be a modest increase in the incidence of bacterial infections associated with TNF inhibitor use[30] and a significant risk of reactivation of latent tuberculosis.[29] Therefore individuals should be screened for tuberculosis infection prior to initiating treatment with TNF inhibitors.[24] The possible association between TNF inhibitors and an increased rate of malignancy in JIA remains an open question; however, currently, there is no convincing evidence of a strong increased risk of overall malignancy.[31]

In addition to TNF inhibitors, the T-cell costimulation modulator abatacept has been shown in a randomized clinical trial to be efficacious in the treatment of JIA.[1] Presumably, because of the effectiveness of TNF inhibitors, the use of abatacept has remained relatively limited in clinical practice.

The B-cell–depleting agent rituximab has been minimally studied in the treatment of JIA.[1] However, it appears effective in some instances, particularly in children who appear to have early-onset RA (teenagers with RF^+ and CCP-positive polyarthritis).

The IL-1 inhibitor anakinra has been shown in both uncontrolled and controlled studies to be particularly effective in the treatment of sJIA.[1,25] Similar to the experience in adults with RA, anakinra appears less effective in treating synovitis among children with the other categories of JIA. Additional IL-1 inhibitors (rilonacept, canakinumab) are also now commercially available and have been shown beneficial in clinical trials treating sJIA.[1]

Unlike other biological agents, the IL-6 inhibitor tocilizumab has been shown in randomized clinical trials to be efficacious in the treatment of both sJIA and polyarticular JIA.[1]

Treatment of Oligoarthritis (Arthritis of ≤4 Joints)

Because of fewer involved joints, oligoarthritis may generally be viewed as a milder form of JIA. However, significant disability can still occur, and children with this condition should not be assumed to have had good clinical outcomes without proper evaluation and therapy. The foundation of treatment for this JIA phenotype is intraarticular glucocorticoid injections. These injections may be administered in multiple joints concurrently and may be repeated, as needed.[28] A good response typically results in resolution of clinical signs and symptoms of arthritis for 4–12 months. DMARD therapy should be initiated in children who do not respond as desired to injections or who have more significant arthritis.[24] Methotrexate is typically the agent of first choice. Significant arthritis that does not respond adequately to methotrexate can be treated with TNF inhibitors.[24]

Treatment of Polyarthritis (Arthritis of ≥5 Joints)

Methotrexate is currently recommended for nearly all children with polyarthritis.[24] If a brief trial of methotrexate proves inadequate to control the arthritis, then TNF inhibitors are frequently recommended.[24] If an adequate response to the initial TNF inhibitor is not seen, then switching to another of the TNF inhibitors or switching to abatacept is currently recommended.[24] The most appropriate role of other effective biological DMARDs, such as tocilizumab (see Table 53.2), in the treatment of polyarticular JIA has yet to be clearly defined.

Treatment of Arthritis Involving Specific Joints

Arthritis involving the TMJ, hip, and sacroiliac joints may deserve special therapy. Destructive arthritis of the TMJ among children with JIA has been noted for decades. Clinical evaluation of this joint is particularly challenging, as symptoms are often absent initially, and physical examination findings may be normal. Accordingly, the optimal treatment for TMJ arthritis is unclear. Many authors recommend targeted therapy with intraarticular glucocorticoid injections, similar to the treatment of arthritis in other joints.[2] Increased systemic therapy is likely also appropriate, although TMJ arthritis has been known to demonstrate radiographic progression despite treatment with systemic TNF inhibitors and in the absence of signs of active synovitis of other joints.[2,22]

The presence of hip arthritis in JIA has been shown in several studies to portend a poor prognosis.[24] Accordingly, many authors advocate early intraarticular glucocorticoid injections and increased systemic therapy when active hip arthritis is identified.

Sacroiliac arthritis is strongly associated with the development of ankylosing spondylitis. Because axial arthritis has been shown to be less responsive to methotrexate therapy, current recommendations suggest a lower threshold for the initiation of TNF inhibitors when sacroiliac arthritis is present.[24] However, early treatment with TNF inhibitors, or perhaps even newer biological agents that inhibit IL-17 or IL-12/IL-23, may be optimal for the treatment of spondyloarthritis to help prevent progression of ankylosis in children.[27]

Treatment of Erosive Arthritis

It appears that not all children with JIA have the propensity to develop an erosive arthritis that is similar to that frequently seen in adults with RA. Children who develop erosions visualized on plain radiographs are considered to have a worse prognosis, and the current recommendation is to increase the intensity of their treatment accordingly.[24]

Treatment of Systemic Features of Systemic Arthritis

For decades, the mainstay of therapy for sJIA has been systemic glucocorticoids and NSAIDs. Nearly all children with sJIA will respond favorably to systemic glucocorticoids, if given in sufficient doses. However, often children become “steroid dependent,” and efforts to decrease the glucocorticoid burden to minimize adverse effects have been unsuccessful. Presumably because of its different pathogenesis, sJIA has not been shown to respond to TNF inhibitors as favorably as the other categories of JIA. Instead, IL-1 and IL-6 appear to be key cytokines in the disease process. Accordingly, IL-1 (anakinra, canakinumab) and IL-6 (tocilizumab) inhibitors are recommended as the first-line nonglucocorticoid treatments for sJIA.[25]

The appearance of clinically significant MAS generally requires directed therapy. The typical treatment approach involves increased systemic glucocorticoids. The calcineurin inhibitor cyclosporine is frequently added,[1,7] and some authors advocate IL-1 and IL-6 inhibitors for treatment of MAS.[1,7] In severe refractory cases, cytotoxic chemotherapeutic agents, such as cyclophosphamide or etoposide, may be warranted.

Treatment of Arthritis of Systemic Arthritis

Some children with sJIA will develop a chronic course of polyarthritis with a relative absence of concurrent systemic features. In general, it is recommended that these children be treated as those with polyarthritis who did not have systemic features at onset.[24] Based on clinical trial results, the IL-6 inhibitor tocilizumab may be the most effective treatment for these patients.

Treatment of Uveitis

JIA-associated anterior uveitis frequently requires directed therapy. Topical glucocorticoid eye drops, such as prednisolone acetate 1%, are frequently initiated at the time of diagnosis by the treating ophthalmologist.[14] Although effective in decreasing the inflammation of uveitis, glucocorticoid eye drops cannot be tolerated in high doses for extended periods because of the risk of cataracts and glaucoma.[14] For this reason, systemic medications are frequently employed in the treatment of JIA-associated uveitis. Methotrexate has been shown to be effective for uveitis and is the most commonly used systemic medication.[1,14] The mAb TNF inhibitors (infliximab, adalimumab, and others) have been shown highly effective in the treatment of anterior uveitis in uncontrolled studies, and they are a frequent choice for children in whom methotrexate proves inadequate to control the disease.[1,14] Other biological agents (*e.g.*, rituximab, abatacept, tocilizumab) appear to be effective in some children with refractory uveitis, but their overall role remains unclear.

Duration of Therapy

As stated, the current goal of therapy is the attainment of clinically inactive disease. However, once this goal is attained, the appropriate next steps in management are less clear. Although none of the currently available therapies is believed to be curative, many children are able to successfully decrease or discontinue therapies after attaining inactive disease status without immediate recurrence of active disease. Many pediatric rheumatologists will consider decreasing the level of therapy if inactive disease status is maintained for a prolonged period, such as 12 months, although this approach is arbitrary. Further study of the appropriate management of children who attain prolonged inactive disease status will be an important future focus.[32]

TRANSLATIONAL RESEARCH

ON THE HORIZON

- Novel criteria for identifying macrophage activation syndrome (MAS) among children with systemic juvenile idiopathic arthritis (sJIA) have been generated using real patient data that are diagnostically highly sensitive and specific.
- Genetic screening for mutations in cytolytic pathway genes will help identify those individuals at risk for developing MAS.
- Murine models are paving the way for a better understanding of MAS immunopathology and potential pathways for clinical intervention.
- Clinical trials involving treatment of MAS with inhibitors of proinflammatory cytokines are just underway.

The explosion in advances in immunology and genetics are leading to major breakthroughs in therapy for rheumatic diseases, including JIA. Challenges remain, however, in diagnosing and treating MAS complicating sJIA. To distinguish a sJIA disease flare-up from MAS, expert opinion and Delphi techniques are currently being used to explore novel criteria for diagnosing MAS in children with sJIA. Data collection regarding clinical, laboratory, and pathological features of children with sJIA, with and without MAS, have been used to develop new classification criteria.[33] Furthermore, genetic mutations and polymorphisms in genes linked to the defective cytolytic pathway in lymphocytes from patients with MAS are being explored to identify patients with sJIA with a propensity to develop MAS.[18] Mouse models of MAS are helping to better understand MAS immunopathology and the role of pro-inflammatory cytokines in the process.[34] Early recognition of MAS and a better understanding of the role of various cytokines in the pathogenesis of MAS will allow for improved targeted therapy for this often fatal condition.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Stoll ML, Cron RQ. Treatment of juvenile idiopathic arthritis: a revolution in care. Pediatr Rheumatol Online J 2014;12:13.
2. Stoll ML, Cron RQ. Temporomandibular joint arthritis in juvenile idiopathic arthritis: the last frontier. Int J Clin Rheumatol 2015;10(4):273–86.
3. Ravelli A, Varnier GC, Oliveira S, et al. Antinuclear antibody-positive patients should be grouped as a separate category in the classification of juvenile idiopathic arthritis. Arthritis Rheum 2011;63(1):267–75.
4. Eng SW, Duong TT, Rosenberg AM, et al. The biologic basis of clinical heterogeneity in juvenile idiopathic arthritis. Arthritis Rheum 2014;66(12):3463–75.
5. Hersh AO, Prahalad S. Immunogenetics of juvenile idiopathic arthritis: A comprehensive review. J Autoimmun 2015;64:113–24.
6. Holzinger D, Kessel C, Omenetti A, et al. From bench to bedside and back again: translational research in autoinflammation. Nat Rev Rheumatol 2015;11(10):573–85.
7. Cron RQ, Davi S, Minoia F, et al. Clinical features and correct diagnosis of macrophage activation syndrome. Expert Rev Clin Immunol 2015;11(9):1043–53.
8. Gmuca S, Weiss PF. Juvenile spondyloarthritis. Curr Opin Rheumatol 2015;27(4):364–72.

9. Li YR, Zhao SD, Li J, et al. Genetic sharing and heritability of paediatric age of onset autoimmune diseases. Nat Commun 2015;6:8442.
10. Meyer B, Chavez RA, Munro JE, et al. DNA methylation at IL32 in juvenile idiopathic arthritis. Sci Rep 2015;5:11063.
11. Zonneveld-Huijssoon E, Albani S, Prakken BJ, et al. Heat shock protein bystander antigens for peptide immunotherapy in autoimmune disease. Clin Exp Immunol 2013;171(1):20–9.
12. Stoll ML. Gut microbes, immunity, and spondyloarthritis. Clin Immunol 2015;159(2):134–42.
13. Arvikar SL, Steere AC. Diagnosis and treatment of Lyme arthritis. Infect Dis Clin North Am 2015;29(2):269–80.
14. Sen ES, Dick AD, Ramanan AV. Uveitis associated with juvenile idiopathic arthritis. Nat Rev Rheumatol 2015;11(6):338–48.
15. Wang Y, Pei F, Wang X, et al. Meta-analysis: diagnostic accuracy of anti-cyclic citrullinated peptide antibody for juvenile idiopathic arthritis. J Immunol Res 2015;2015:915276.
16. Miyara M, Ito Y, Sakaguchi S. TREG-cell therapies for autoimmune rheumatic diseases. Nat Rev Rheumatol 2014;10(9):543–51.
17. Cimaz R, Moretti D, Pagnini I, et al. What do cytokine profiles tell us about subsets of juvenile idiopathic arthritis? Curr Rheumatol Rep 2012;14(2):150–4.
18. Zhang M, Behrens EM, Atkinson TP, et al. Genetic defects in cytolysis in macrophage activation syndrome. Curr Rheumatol Rep 2014;16(9): 439.
19. Stoll ML, Punaro M. Psoriatic juvenile idiopathic arthritis: a tale of two subgroups. Curr Opin Rheumatol 2011;23(5):437–43.
20. Malleson PN, Mackinnon MJ, Sailer-Hoeck M, et al. Review for the generalist: The antinuclear antibody test in children - When to use it and what to do with a positive titer. Pediatr Rheumatol Online J 2010;8:27.
21. Ording Muller LS, Humphries P, Rosendahl K. The joints in juvenile idiopathic arthritis. Insights Imaging 2015;6(3):275–84.
22. Vaid YN, Dunnavant FD, Royal SA, et al. Imaging of the temporomandibular joint in juvenile idiopathic arthritis. Arthritis Care Res (Hoboken) 2014;66(1):47–54.
23. Cattalini M, Khubchandani R, Cimaz RJ. When flexibility is not necessarily a virtue: a review of hypermobility syndromes and chronic or recurrent musculoskeletal pain in children. Pediatr Rheumatol Online J 2015;13:40.
24. Beukelman T, Patkar NM, Saag KG, et al. 2011 American College of Rheumatology recommendations for the treatment of juvenile idiopathic arthritis: initiation and safety monitoring of therapeutic agents for the treatment of arthritis and systemic features. Arthritis Care Res (Hoboken) 2011;63(4):465–82.
25. Ringold S, Weiss PF, Beukelman T, et al. 2013 update of the 2011 American College of Rheumatology recommendations for the treatment of juvenile idiopathic arthritis: recommendations for the medical therapy of children with systemic juvenile idiopathic arthritis and tuberculosis screening among children receiving biologic medications. Arthritis Rheum 2013;65(10):2499–512.
26. Wallace CA, Giannini EH, Huang B, et al. American College of Rheumatology provisional criteria for defining clinical inactive disease in select categories of juvenile idiopathic arthritis. Arthritis Care Res (Hoboken) 2011;63(7):929–36.
27. Braun J, Kiltz U, Heldmann F, et al. Emerging drugs for the treatment of axial and peripheral spondyloarthritis. Expert Opin Emerg Drugs 2015;20(1):1–14.
28. Schiappapietra B, Varnier G, Rosina S, et al. Glucocorticoids in juvenile idiopathic arthritis. Neuroimmunomodulation 2015;22(1-2):112–18.
29. Jinesh S. Pharmaceutical aspects of anti-inflammatory TNF-blocking drugs. Inflammopharmacology 2015;23(2-3):71–7.
30. Hurd A, Beukelman T. Infectious complications in juvenile idiopathic arthritis. Curr Rheumatol Rep 2013;15(5):327.
31. Mannion ML, Beukelman T. Risk of malignancy associated with biologic agents in pediatric rheumatic disease. Curr Opin Rheumatol 2014;26(5):538–42.
32. Verazza S, Negro G, Marafon D, et al. Possible discontinuation of therapies after clinical remission in juvenile idiopathic arthritis. Clin Exp Rheumatol 2013;31(4 Suppl. 78):S98–101, quiz S102-7.
33. Ravelli A, Minoia F, Davì S, et al. 2016 Classification Criteria for Macrophage Activation Syndrome Complicating Systemic Juvenile Idiopathic Arthritis: A European League Against Rheumatism/American College of Rheumatology/Paediatric Rheumatology International Trials Organisation Collaborative Initiative. Arthritis Rheumatol 2016;42(7):566–76.
34. Weaver LK, Behrens EM. Hyperinflammation, rather than hemophagocytosis, is the common link between macrophage activation syndrome and hemophagocytic lymphohistiocytosis. Curr Opin Rheumatol 2014;26(5):562–9.

MULTIPLE-CHOICE QUESTIONS

1. Which of the following is NOT associated with increased risk of uveitis in children with juvenile idiopathic arthritis (JIA)?
   A. Female gender
   B. Antinuclear antibody (ANA) positivity
   C. Rheumatoid factor (RF) positivity
   D. Oligoarticular disease
   E. Young age at disease onset

2. Which of the following is NOT consistent with the diagnosis of macrophage activation syndrome in a child with systemic JIA?
   A. Thrombocytopenia
   B. Coagulopathy
   C. Hyperferritinemia
   D. Rising cross-reactive protein (CRP) level
   E. Rising erythrocyte sedimentation rate (ESR)

3. What is the most reliable finding denoting onset of temporomandibular joint (TMJ) arthritis in a child with JIA?
   A. TMJ pain
   B. Maximal incisal opening <5 cm
   C. Elevated ESR
   D. Synovial thickening on magnetic resonance imaging (MRI)
   E. TMJ popping

54

Sjögren Syndrome

Sarfaraz A. Hasni, Leyla Y. Teos, Ilias Alevizos

Sjögren syndrome (SS) is a chronic systemic autoimmune disease characterized by lachrymal and salivary gland dysfunction. It was named after the Swedish ophthalmologist Henrik Sjögren after he published a report of 19 cases of keratoconjunctivitis in 1933.[1] The hallmark feature of SS is deficient tear and saliva production as a result of lymphocytic infiltration of the salivary and lachrymal glands leading to *xerostomia* (dry mouth) and *xerophthalmia* (dry eyes). In addition, SS can involve any organ system and present with a wide spectrum of clinical features. The autoimmune process seems to primarily affect the lining epithelium of various organs; in fact, some experts have proposed the term "autoimmune epithelitis" to be used instead of SS.[2]

KEY CONCEPTS

- Sjögren syndrome (SS) is one of the most common autoimmune diseases, mainly affecting females. It targets the salivary and lachrymal glands, causing xerostomia and xerophthalmia.
- The diagnosis of SS is complex and requires a multidisciplinary team for the assessment of oral, ocular, and systemic manifestations.
- Anti-Ro/SSA and Anti-La/SSB autoantibodies are encountered in approximately 70% of patients with SS. Significantly increased incidence of congenital heart block is present in the offspring of Anti-Ro/Anti-La positive mothers.
- The risk of lymphoma is significantly higher in patients with SS, and this risk increases over time.
- Currently, there is only palliative treatment for alleviation of the symptoms of SS.

EPIDEMIOLOGY

SS predominantly affects females (female-to-male ratio 9 : 1) in their fourth and fifth decades of life. However, symptoms can be present for a much longer time, and there is usually a 5- to 10-year delay in the diagnosis of SS. The reported prevalence of SS varies widely from 0.1% to 4.8%.[3] This variation may result from the use of different classification criteria, geographical and environmental influences, and varying study sizes and target populations. When a more strict definition is applied to the available data, the true prevalence of SS is estimated to be around 1%,[1] making it the second most common systemic rheumatic disease after rheumatoid arthritis (RA).

IMMUNOPATHOGENESIS

The pathogenesis of SS is largely unknown. In a genetically predisposed individual, various environmental factors, such as viral infections, may lead to epithelial cell activation that triggers a protracted inflammatory response with features of autoimmunity. Autoreactive lymphocytes and autoantibodies are considered important in this process, although the pathogenic role of any particular autoantibody is still undefined.[4]

Immunogenetic Factors

SS is associated with certain human leukocyte antigen (HLA) class II haplotypes. Initial reports of HLA class II associations described enrichment of DR3 and DR2 in European populations of patients with primary SS. This genetic association predominantly involves antibody-positive but not antibody-negative patients with SS. Several genetic polymorphisms previously linked to other autoimmune diseases are also associated with SS. Among these, two transcription factors, STAT4 (signal transducer and activator of transcription 4) and IRF5 (interferon [IFN] regulatory factor 5), which were both independently associated with SS, showed an additive effect, increasing the risk of SS from around 1.6–1.9 for one risk allele to 6.7 when both risk alleles were present.[5]

Although the individual gene approach has successfully identified genes associated with SS, two large-scale genome-wide association studies have increased the number of genetic loci that function as potential risk factors.[6] A study by Lessard et al.[7] genotyped more than 10 000 patients with SS and control subjects of European descent and found seven genetic loci to be significantly associated with SS. These loci included the following genes: *MHC-II, IRF5, STAT4, IL-12A, BLK, CXCR5,* and *TNIP1.* A second large-scale study examined DNA variations from approximately 1700 patients with SS and controls subjects from Han, China, and identified a new gene associated with SS, the transcription factor GTF2I located in 7q21. However, the odds ratio for GTF2I was a moderate 2.2.[8]

Environmental Factors

The inciting etiopathogenic event in SS is not known, and it may not be a single event. The strong predominance of females suggests gender-specific predisposing factors. Although sex hormones are of obvious interest, there is no conclusive proof that the difference in the pathogenesis between males and females can be attributed to sex hormones alone. Overall, estrogens are considered contributors to autoimmunity, whereas androgens are thought to be protective.

Viral infections have also been proposed as inciting events. This theory is supported by the observation that chronic inflammation of the salivary glands occurs with chronic hepatitis C, human T-lymphotropic virus 1 (HTLV-1), and human immunodeficiency virus (HIV) infections, and such infections cause a disease with a

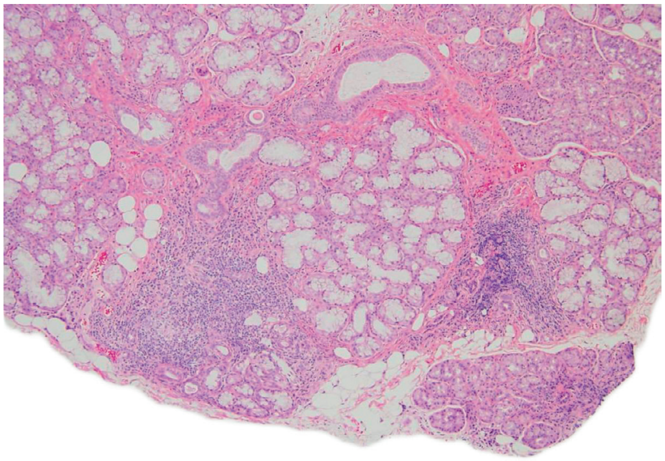


FIG 54.1 Minor salivary gland biopsy with characteristic periductal inflammation.

clinical spectrum similar to that of SS. The fact that some viruses, such as Epstein-Barr virus (EBV), replicate in oropharyngeal and lachrymal glands has led to the hypothesis that these viruses might contribute to the pathogenesis of SS. Other viruses, such as coxsackievirus or endogenous retroviruses, have also been proposed as agents of interest. However, to date, there is no proof that any of these viruses plays a pathogenic role in SS.

Epithelial Cell Activation and Chronic Inflammation

The histological hallmark of SS is a periductal mononuclear infiltrate in salivary and lachrymal glands (Fig. 54.1). The majority of the infiltrating cells are CD4 T lymphocytes, whereas CD8 cytotoxic T cells are found in smaller numbers. Activated B lymphocytes are also present, including autoantibody-secreting plasma cells.

Epithelial cell activation, signified by the expression of HLA class II molecules, is key to initiating recruitment of the inflammatory infiltrate. The expression of these molecules, along with upregulation of adhesion molecules and chemokines, contributes to the recruitment of inflammatory cells, such as T and B lymphocytes, macrophages, and dendritic cells (DCs), and suggests that epithelial cells may act as antigen-presenting cells (APCs), actively participating in lymphocyte activation.[9] The ensuing chronic inflammatory process is characterized by a complex interaction between activated epithelial cells and the innate and adaptive immune systems. In the most severe forms of inflammation, the characteristic periductal infiltrates can progress to the formation of germinal centers, which is associated with a higher risk of lymphoma development. Extraglandular manifestations occur as a result of similar lymphocytic infiltrations in other organs. This is described by some as *autoimmune epithelitis* to better reflect the systemic nature of the disease.[9] In most patients, only partial destruction of the glands is noted. Local production of cytokines, autoantibodies, metalloproteinases, and other inflammatory mediators may contribute to the dysfunction of the remaining epithelial cells. It is increasingly recognized that the innate immune system plays a crucial rule in the immunopathogenesis of SS.

Autoantibodies

Autoantibodies are the hallmark of systemic autoimmune diseases, including SS. The best-defined autoantibodies in SS are anti-Ro/SSA and anti-La/SSB antibodies.[10] Both autoantibodies are targeted against ribonucleoprotein antigens. Anti-Ro/SSA recognizes two RNA binding proteins (the 52-kilodalton [kDa] or the 60-kDa protein), whereas anti-La/SSB antibodies recognize RNA polymerase III. Anti-Ro/SSA antibodies are found in over 70% of patients with SS but are also frequently encountered in systemic lupus erythematosus (SLE) and other autoimmune diseases, even in the absence of oral or ocular dryness. Anti-La/SSB is more specific; it is present in 50% of patients with primary SS or SS/SLE but is rarely seen in other diseases. The pathogenic role of these antibodies has not yet been defined, except in infants born to women with anti-Ro/SSA and/or anti-La/SSB antibodies. These antibodies can cross the placenta and bind to Ro and La antigens located on the cell surface of fetal myocardial tissue, leading to fetal heart block. Other autoantibodies, such as antinuclear antibodies (ANAs) and rheumatoid factor, are frequently present in patients with both primary and secondary SS. Although these antibodies lack specificity, they are markers of a systemic autoimmune response and thus can help distinguish SS from other causes of salivary or lachrymal gland dysfunction.

AUTONOMIC NERVOUS SYSTEM ABNORMALITIES

Autonomic nervous system (ANS) abnormalities are common in SS and may play a critical role in its pathogenesis. Xerostomia and xerophthalmia, the cardinal manifestations of SS, are features of cholinergic parasympathetic ANS dysfunction, whereas sympathetic cholinergic failure results in xerosis and decreased sweating, which are frequently reported by patients with SS.

The complexity of the ANS, along with differences in methodology and studied populations, has resulted in variable results, but abnormalities in SS have been reported in both sympathetic and parasympathetic ANS domains.

CLINICAL MANIFESTATIONS

Clinical manifestations resulting from dysfunction of salivary and lachrymal glands are the dominant features of SS. Symptoms secondary to other exocrine gland dysfunction, such as skin and vaginal dryness, and chronic cough caused by tracheal dryness are frequently reported. Multiple extraglandular manifestations can also be present, reflecting the systemic nature of the autoimmune process.[4]

Constitutional Symptoms

One of the most common extraglandular manifestations of SS is excessive fatigue. Approximately 70% of patients with SS complain about disabling fatigue leading to reduced quality of life.[11] Attempts to identify a biological basis for fatigue have failed to reveal any correlations with levels of inflammatory markers, cytokines, or autoantibodies. This fatigue is commonly associated with arthralgia, malaise, and mental cloudiness (brain fog), resulting in a diagnosis of fibromyalgia or chronic fatigue syndrome. Less commonly, patients also experience low-grade fever and weight loss.

Ocular Involvement

Lachrymal gland dysfunction leads to dry, sandy eyes and sensation of foreign body in the eyes. Frequently, there is a history of intolerance to contact lenses. Patients often complain of sticky eyes and accumulation of thickened mucus as a result of loss of aqueous tear components.

CLINICAL PEARLS

- Chronic fatigue is a prominent presenting feature of Sjögren syndrome (SS).
- Autonomic nervous system involvement and peripheral nervous system involvement are often underrecognized.
- Mimics of SS include immunoglobulin G4 (IgG4)–related disease (Mikulicz disease), hepatitis C infection, sarcoidosis, and human T-lymphotropic virus (HTLV) infection.
- Presence of joint erosions and complement control protein (CCP) antibody is indicative of secondary SS caused by rheumatoid arthritis.
- Prolonged use of topical ophthalmic nonsteroidal antiinflammatory drugs (NSAIDs) and steroid preparations should be avoided to reduce the risk for local complications.
- Sudden normalization of previously elevated rheumatoid factor should prompt evaluation for development of lymphoma.
- Oral candidiasis should be identified and treated. Patients with SS are at a high risk for oral candidiasis, which can present as oral erythema and/or pain.
- Pediatric primary SS is rare and presents with variable, atypical features, most commonly recurrent tender parotid gland swelling.

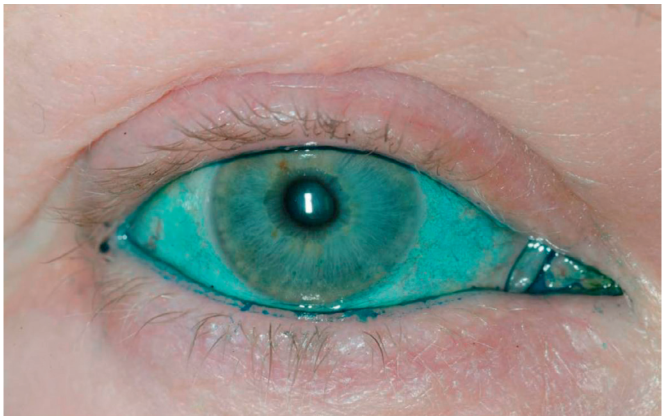


FIG 54.2 Punctate keratitis. Lissamine green stains dead or degenerated corneal epithelial cells green. (Courtesy of Manuel Datiles, MD.)

Enlarged lachrymal glands can be noted at the onset of disease. In the initial stages, injected conjunctivae with strands of thickened mucus at the inner canthus can be found. In more advanced stages, conjunctivae may lose their normal luster and become edematous. Corneal involvement results in punctuate keratitis as a result of epithelial cell death (Fig. 54.2). During early stages, the keratitis may be limited to the lower quadrants of the cornea to eventually involve the entire cornea. If left untreated, corneal ulcers can develop and result in rapid corneal thinning. In addition to reduced tear production, there is also evidence of inflammation of the ocular surface epithelium, as indicated by the presence of inflammatory cell infiltrate and elevated levels of inflammatory cytokines.

Oral Involvement

Saliva plays a critical role in maintaining oral health and comfort. It has antibacterial, lubricating, remineralizing, digestive, buffering, and cleansing properties. Therefore decreased salivary production or altered salivary composition can result in numerous conditions affecting oral health, comfort, and quality of life. Composed of 99% water, electrolytes, and proteins, the initial saliva is secreted by acinar cell clusters, which are innervated by both sympathetic and parasympathetic nerves. The sympathetic nerve activates alpha- and beta-adrenergic receptors at the plasma membrane, initializing a process that eventually catalyzes exocytosis of

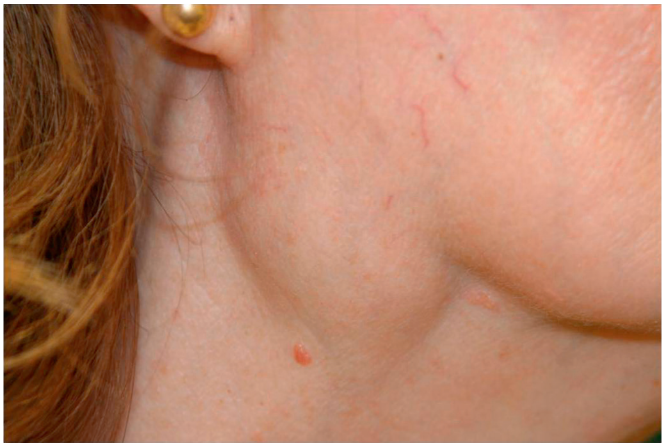


FIG 54.3 Enlargement of the parotid gland.

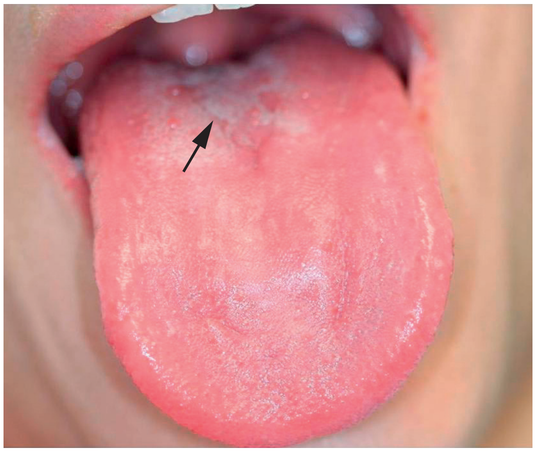


FIG 54.4 Papillary atrophy and candidiasis of the tongue. The loss of filiform papillae results in a smooth surface of the tongue. Oral candidiasis is common in Sjögren syndrome (SS) and can present as erythema only with minimal or no exudate (arrow).

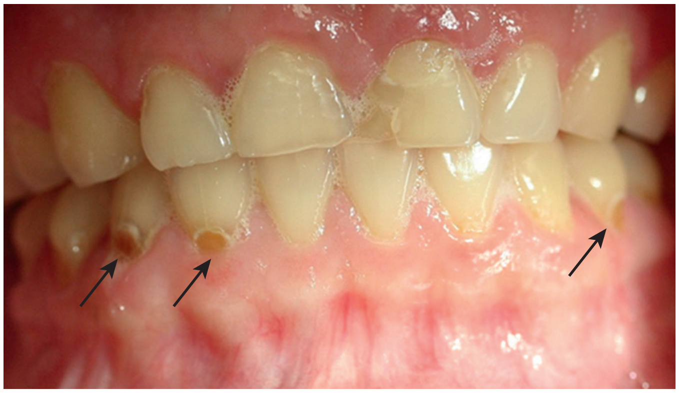


FIG 54.5 Root caries on anterior teeth caused by loss of the protective functions of saliva is a common finding. The cusp tips of the posterior teeth are also commonly affected.

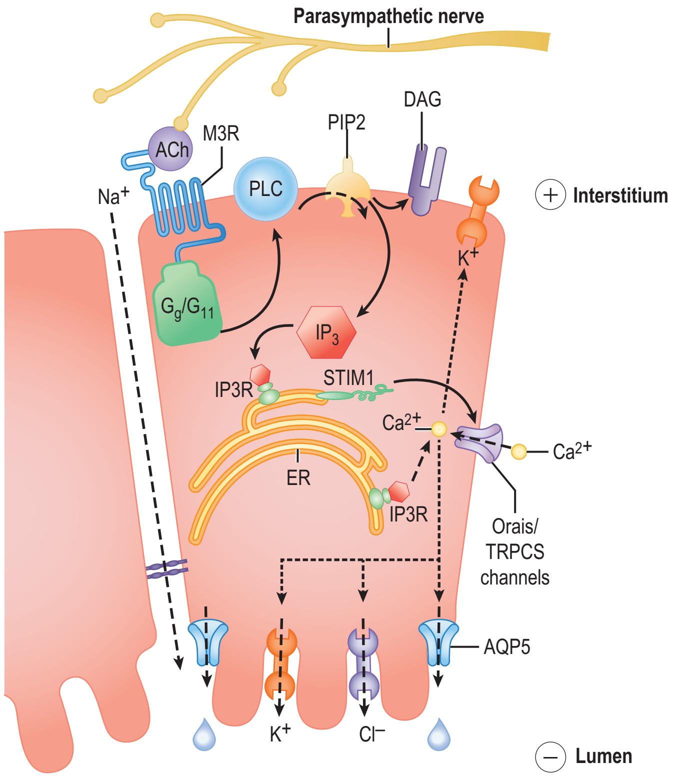


FIG 54.6 Acinar cells fluid and electrolyte regulation signaling pathway. M3 receptors coupled to G proteins activate phospholipase C (PLC); this results in hydrolysis of phosphatidylinositol 1,4 bisphosphate (PIP2) to yield diacylglycerol (DAG) and inositol 1,4,5-trisphosphate (IP3). IP3 diffuses in the cytosol and binds and activates the IP3 receptors positioned on the plasma membrane of the endoplasmic reticulum (ER).[19] This causes opening of the IP3 receptors releasing calcium from the stores. The decrease in calcium within the ER is sensed by the calcium sensing protein, stromal interaction molecule 1(STIM1). STIM1 translocates to the plasma membrane and activates the Calcium Release-Activated Calcium Modulator (ORAI) and Tansient Receptor Potential, Canonical (TRPCs) calcium channels. These channels are called *store-operated channels* (SOCs), which upon activation cause calcium influx. The calcium (Cl^-) entering the cells originating at the apical pole as a result of polarity activate potassium (K^+) and Cl^- channels releasing K^+ and Cl^- into the lumen. Subsequently, calcium propagates to the basolateral side of the cell where it activates K^+ channels releasing K^+ into the interstitium. This creates an electrochemical gradient that drives sodium (Na^+) across the paracellular pathway.[20] The ion accumulation in the lumen drives water out through the aquaporin 5 (AQP5) channels, causing cell shrinkage.

proteins. Meanwhile, the parasympathetic nerve activates muscarinic receptors (M3) by releasing acetylcholine, which induces a calcium-dependent cascade of acinar cell volume change (majority of salivary fluid) and secretion of electrolytes.[12-14]

Symptoms and signs of salivary gland dysfunction include oral dryness (xerostomia); oral pain and sensitivity; difficulty chewing, swallowing, and speaking; diminished taste; inflamed, morphologically altered, or infected mucosal tissues; and increased tooth decay.

Chronic xerostomia necessitates the use of fluids throughout the day and night. It is not uncommon for patients with SS to wake up at night to drink fluids to relieve dryness and also to require the use of fluids to be able to speak, chew, and swallow. A burning sensation along with taste alterations and decreased tolerance to spicy foods is described by a significant percentage of patients.

Head and neck examination often reveals dry and chapped lips, angular cheilitis, and swollen major salivary glands (Fig. 54.3), which may be recurrent or chronic. Intraoral findings include decreased salivary pooling under the tongue and a depapillated tongue surface often associated with erythematous candidiasis (Fig. 54.4). Vasculitis can also manifest in the oral cavity and lead to tissue breakdown. Teeth often exhibit an increased rate of caries with characteristic decay affecting the roots and cusp tips (Fig. 54.5).

A description of the water secretion molecular mechanism in salivary epithelial cells in shown in Fig. 54.6.

Musculoskeletal Involvement

Musculoskeletal complaints are present in the majority of patients with SS. The rate of incidence of arthralgias has been reported as 50–75% in various studies and precedes sicca symptoms in up to 30% of patients. Arthralgias are symmetrical and usually without evidence of synovitis or erosions. Large-joint and small-joint involvement is equally reported, and 10–20 % of patients have recurrent monoarthritis or oligoarthritis. From a clinical perspective, it is often difficult to differentiate between patients with primary SS and symmetrical joint involvement and those with RA presenting with sicca symptoms.

Neuropsychiatric Manifestations

Primary SS can result in multiple neurological manifestations involving both the central nervous system (CNS) and the peripheral nervous system (PNS). PNS manifestations are more common and are present in 5–20% of patients with SS. PNS involvement can present as sensory neuropathy, small-fiber neuropathy, cranial neuropathy, inflammatory myopathy, and autonomic neuropathy. Small-fiber neuropathy is increasingly being recognized as the most common PNS manifestation of SS. It presents as a subacute to chronic development of excruciating burning pain. There is selective impairment of pinprick and temperature sensation, with preserved vibratory sense and proprioception. Diagnosis is made with skin biopsy. Symptoms suggestive of autonomic dysfunction, such as orthostatic hypotension, temperature intolerance, constipation, urinary frequency or hesitancy, and delayed gastric emptying, are present in up to half the patients with SS.[15] CNS involvement is rare and can present as white matter lesions, difficult to distinguish from multiple sclerosis, myelopathy, and optic neuritis, as well as cognitive dysfunction, seizures, and encephalopathy.[15] Neuromyelitis optica spectrum disorders (NMOSDs), with the hallmark features of acute optic neuritis or transverse myelitis with a relapsing course, have been linked to the presence of antiaquaporin-4(AQP4) serum autoantibody. In 213 patients diagnosed with NMOSD, 78 (37%) were found to fulfill criteria for SS.[16-18] Anxiety and depression are commonly seen in patients with SS, perhaps because of the severe impact of the manifestations of disease on the patients' quality of life.

Dermatological Involvement

Skin is involved in about 50% of SS patients[19]; lesions can be vasculitic or nonvasculitic. The most common nonvasculitic involvement is dry, itchy skin (xerosis) and angular cheilitis. Other less common nonvasculitic lesions include alopecia, vitiligo, eyelid dermatitis, and erythema nodosum–like lesions. Annular erythema, a rash similar to subacute cutaneous lupus, has been reported in about 10% of patients with SS and is closely associated with the presence of anti-SSA and/or anti-SSB autoantibodies. However, the rash is less photosensitive and histologically has deeper perivascular and periappendageal lymphocytic infiltrates.

Vasculitic lesions include nonpalpable purpura caused by hypergammaglobulinemia, palpable purpura, and urticarial vasculitis.[19] Patients with vasculitis have higher prevalence of positive ANA, anti-Ro/SSA antibodies, and rheumatoid factor compared with those with nonvasculitic skin involvement. Palpable purpura is histologically characterized by leukocytoclastic vasculitis and may indicate presence of underlying cryoglobulinemia. Raynaud phenomenon (RP) is seen in 30–50% of SS patients and frequently precedes the development of sicca symptoms.

Gastrointestinal Involvement

Dysphagia is common as a result of dryness of the pharynx and esophageal dysmotility. Nausea, epigastric pain, and constipation are frequently encountered. Gastrointestinal (GI) tract dysmotility can be seen as a result of ANS involvement. Pancreatic dysfunction, with reduced production of amylase and lipase, is present in 25% of patients with SS.

Pulmonary Involvement

Cough is a common complaint in patients with SS and results from tracheobronchitis sicca (tracheal and bronchial dryness). Clinically significant pulmonary involvement is reported in 9–12% of SS patients.[20] Pulmonary manifestations include large airway disease and small airway disease, bronchial hyperresponsiveness, recurrent respiratory infections, interstitial lung disease, pulmonary hypertension, pleuritis, and pulmonary amyloidosis. Interstitial lung disease is increasingly recognized as an extraglandular manifestation of SS and can precede the diagnosis of SS.

Patients with primary SS presenting with unexplained dyspnea should be evaluated for pulmonary arterial hypertension (PAH); an unexpectedly high prevalence (14%) of pulmonary hypertension was found in a cohort of 1158 patients. The majority of patients with SS and PAH manifest symptoms early in the disease course.[21,22]

Cardiac Involvement

Clinically significant cardiac involvement is uncommon in primary SS. Patients may develop pericarditis, thickened pericardium, or left ventricle diastolic dysfunction. Rhythm abnormalities resulting from autonomic dysfunction may be encountered. Placental transfer of anti-SSA/Ro and anti-SSB/La antibodies from pregnant patients with SS can cause lesions in the cardiac conducting system of the fetus. Pregnant women with positive anti-SSA/Ro and anti-SSB/La antibodies have a 2.5% risk of delivering a baby with congenital heart block.

Like patients with other chronic inflammatory conditions, patients with SS are at an increased risk of premature atherosclerosis and cardiovascular events. Some traditional cardiovascular risk factors, such as hypertension and hypercholesterolemia, are more prevalent in patients with SS. However, disease-specific factors also contribute to the pathogenesis of premature atherosclerosis.[23-25]

Genitourinary and Renal Involvement

Renal involvement is frequently caused by tubulointerstitial nephritis often presenting as distal (type1) renal tubular acidosis, renal calculi, and hypokalemia. Glomerular disease has been rarely reported and, in most cases, is linked to cryoglobulinemic vasculitis. Inflammation of the urinary bladder epithelium can cause interstitial cystitis. Vaginal dryness leading to dyspareunia has been reported in 40% of women.

Clinical Manifestations in Children

Primary SS is rarely diagnosed in children, but many adult patients with SS recall the presence of symptoms since childhood. Clinical presentation in children is more insidious with predominant recurrent parotitis and, less commonly, sicca symptoms.[26]

Associated Autoimmune Conditions

Autoimmune thyroiditis is seen in more than one-third of patients with SS. Celiac disease is also increasingly recognized in these patients, with a prevalence 10 times higher than in the general population.

Lymphoma Associated With SS

SS is associated with increased risk of lymphoma, with about 5% of patients developing lymphoma. In a meta-analysis of 20 studies the standardized incidence rate of lymphoma was 18.9% (confidence interval [CI] 9.4–37.9).[27] Most of these are low-grade marginal zone lymphomas of the mucosa-associated lymphoreticular tissue (MALT) type. Only 10% of lymphomas in patients with SS evolve into a less differentiated cell (more aggressive) variety. Risk factors for development of lymphoma include persistent enlargement of the parotid glands, palpable purpura, leg ulcers secondary to vasculitis, mixed cryoglobulinemia, and low C4 levels.

DIAGNOSIS AND CLASSIFICATION CRITERIA

Patients with SS can present with heterogeneous combinations of sicca symptoms and extraglandular manifestations. Moreover, various clinical syndromes and medications can lead to similar symptoms and mimic SS. The diagnosis of SS is based on the combination of subjective symptoms of dryness (sicca symptoms), objective evidence of lachrymal or salivary gland hypofunction, and evidence of autoimmunity or salivary gland inflammation. Classification criteria were proposed by an American–European consensus group (AECG) in 2002,[28] primarily for classification of patients in clinical studies. These criteria are widely accepted and are frequently used as a guide for clinical diagnosis. Application of these criteria to classify patients with primary SS has a sensitivity of 89.5% and specificity of 95.2%, whereas for secondary SS, it has a sensitivity of 97.2% and specificity of 90.2%.[28] The AECG classification scheme is based on a set of two subjective criteria and four objective criteria.

Patients are classified as having primary SS in the absence of any other systemic autoimmune disease, whereas in secondary SS, sicca manifestations are present in conjunction with other underlying autoimmune diseases, such as SLE, RA, polymyositis/dermatomyositis, or systemic sclerosis. SS is excluded if patients use anticholinergic medications or have other conditions that can mimic SS, such as chronic viral infections, graft-versus-host disease (GvHD), or status post irradiation of the head and neck region.

Patients presenting with sicca symptoms suggestive of SS should undergo specialized testing to confirm the diagnosis. Lachrymal function can be assessed by using the Schirmer test to quantify the amount of tear production, whereas corneal and conjunctival damage caused by dryness can be measured by fluorescein and lissamine green staining. Salivary gland evaluation is done by collection of unstimulated saliva or by performing salivary scintigraphy. Laboratory workup should include testing for autoantibodies to SSA/Ro and SSB/La antigens, which are detected in sera of 45–70% and 20–50%, respectively, in patients with SS. However, the AECG criteria have been criticized for including subjective symptoms and physiological tests lacking specificity and for using alternative tests that are diagnostically nonequivalent. In addition, with the potential use of biological agents to treat SS, more standardized, clear, and highly specific criteria are needed.

To address these concerns, in 2012 the American College of Rheumatology (ACR) proposed new classification criteria for SS, which required the presence of two out of three objective criteria in individuals suspected of having SS based on their signs and symptoms (Table 54.1). These new criteria include no subjective ocular and oral symptoms and no functional tests for the salivary glands.

In subsequent studies comparing the AECG and ACR criteria, the degree of concordance ranged from modest ($\kappa = 0.53$) to high ($\kappa = 0.81$), but neither set of criteria has been shown to be clearly superior to the other. At present, it appears that the stricter ACR criteria are more suitable for high-risk clinical trials, such as those using biologicals, and the AECG criteria may have broader applicability in clinical research.[29-31]

Novel Biomarkers for Diagnosis and Management of Sjögren Syndrome

MicroRNAs (miRNAs) are small regulatory noncoding RNAs of 18–25 nucleotides in length. Over the last few years, interest in miRNAs as clinical biomarkers has been increasing steadily. The reasons for this increased interest mainly stems from recognition of the master regulatory roles they play in normal and altered physiological states and also their stability compared with mRNAs, a characteristic that allows them to be readily isolated from archived samples, such as formalin-fixed paraffin-embedded (FFPE) tissues or frozen saliva.[32-34]

TABLE 54.1 **Proposed American College of Rheumatology (ACR) Classification Criteria for Sjögren Syndrome**

The classification of Sjögren syndrome (SS), which applies to individuals with signs/symptoms that may be suggestive of SS, will be met in patients who have at least two of the following three objective features:

1. Positive serum anti-SSA (Ro) and/or anti-SSB (La) or (positive rheumatoid factor and antinuclear antibody [ANA] $\geq$ 1 : 320)
2. Labial salivary gland biopsy exhibiting focal lymphocytic sialadenitis with a focus score ≥ 1 focus/4 mm^2
3. Keratoconjunctivitis sicca with ocular staining score ≥ 3 (assuming that the patient is not currently using daily eye drops for glaucoma, and has not had corneal surgery or cosmetic eyelid surgery in the last 5 years)[3]

Prior diagnosis of any of the following conditions would exclude participation in SS studies or therapeutic trials because of overlapping clinical features or interference with criteria tests:

History of head and neck radiation treatment
Hepatitis C infection
Acquired immunodeficiency syndrome
Sarcoidosis
Amyloidosis
Graft-versus-host disease
Immunoglobulin G4 (IgG4)–related disease

The role of miRNAs in primary SS is being investigated by several groups worldwide not only for their effects on immune cells but also on the target organs, such as the salivary glands.[35-37]

Rheumatoid factor is commonly positive in patients with SS and is part of the new ACR criteria for the diagnosis of SS. Recently, anticentromere antibodies were found to be present in up to 13% of patients with SS.

TREATMENT

THERAPEUTIC PRINCIPLES

Management of Sjögren Syndrome (SS)

Ocular Manifestations

- Symptomatic relief with artificial tear drops and lubricating ointments
- Topical cyclosporine eye drops
- Blockage of nasolacrimal duct by punctual plug placement or surgical cautery

Oral Manifestations

- Symptomatic relief with water and artificial saliva preparations
- Physical measures to increase saliva production: sugar-free candies and gum
- Maintaining meticulous oral hygiene; use of fluoride toothpaste
- Sialogogues: pilocarpine or cevimeline

Systemic Manifestations

- Hydroxychloroquine most commonly used to alleviate fatigue and arthralgias
- Other immunosuppressive medications may be considered in more severe internal organ involvement
- Anti–tumor necrosis factor (TNF) blocking agents are not recommended
- Role of rituximab needs clarification

Patient Education

- Patients should receive the following instructions:
  - Avoid extreme dry environments, such as proximity to open fire places.
  - Wear protective eye glasses or goggles.
  - Use sugar-free sour candies or gum.
  - Maintain good ocular and oral hygiene.
  - Avoid use of medications and substances causing sicca symptoms.
- Women of child-bearing age with positive anti-SSA or anti-SSB antibodies should be counseled about the risk of congenital heart block.
- Educate patients about increased risk of lymphoma associated with SS.
- Safety precautions should be taken with regard to smoke and gas leakage in the environment of patients with loss of smell.
- Information regarding focus groups, such as Sjögren's Syndrome Foundation (http://www.sjogrens.com/) and the British Sjögren's Syndrome Association, should be provided to patients.

Symptomatic Treatment of Sicca Symptoms

The mainstay of treatment in SS is providing symptomatic relief. Use of artificial tears is often helpful. In patients with sticky mucus or strands over the eyes, use of mucolytic agents is indicated. Blockage of lachrymal ducts with punctual plugs or surgically can be beneficial.

Artificial saliva available as spray or lozenges provides only limited relief. Local activation of salivary glands (in patients with preserved function) with sugar-free sour candies or gums provides some relief of symptoms.

Cholinergic agonists, such as pilocarpine and cevimeline, provide effective symptom relief in selected patients. Cevimeline is a more selective muscarinic agonist predominantly affecting M1 and M3 receptors and hence is associated with fewer side effects compared with pilocarpine.[4]

The Sjögren's Syndrome Foundation's Clinical Practice Guidelines Committee recently recommended the use of topical fluoride to reduce dental caries in patients with SS who suffer from dry mouth.[38]

Immune-Modulating Medications

Cyclosporine eye drops are frequently prescribed to reduce the local immune response involving the conjunctiva and the cornea. Use of topical nonsteroidal antiinflammatory and steroid-based eye drops provide short-term benefits and should be used cautiously.

Hydroxychloroquine is commonly prescribed for patients with SS, even though some trials have shown significant laboratory improvement but no beneficial effects on symptoms.

Other immune-modulating agents, such as azathioprine, cyclosporine, methotrexate, and mycophenolate mofetil, have shown only limited benefits in treating sicca symptoms and are used primarily for extraglandular manifestations.[39]

Trials using biological agents, such as infliximab and etanercept, failed to show any significant improvement in the primary outcomes of treatment for oral and ocular dryness.[39]

Data on the use of rituximab are controversial, with some studies showing only modest benefit.[39]

An open-label trial using the anti–B cell–activating factor (BAFF) monoclonal antibody (mAb) belimumab revealed no major safety concerns. There were minor improvements in parotid gland swelling and lymphadenopathy but no significant improvement in salivary and lacrimal gland function or patient-reported outcomes. Despite the central role of autoantibodies in the pathogenesis of SS, targeting of B-cell function has been surprisingly unsuccessful in managing SS.

Treatment for Lymphoma Associated With SS

As discussed above, lymphoma associated with SS is usually low grade with 5-year survival rates of 86–100%. In a study of patients with SS, overall survival rates of patients with MALT lymphoma was similar in both the treated and the untreated groups.[24] However, patients with disseminated and more aggressive lymphoma have reduced overall survival. Recently, high levels of BAFF and Flt3 were shown to be highly predictive of lymphoma development in patients with SS.[40]

PATIENT EDUCATION

Because of its significant impact on quality of life, educating patients and their families is of utmost importance. Excellent resources for patient education are available through focus groups, such as the Sjögren's Syndrome Foundation (http://www.sjogrens.com/) and the British Sjögren's Syndrome Association.

In general, patients are advised to avoid dry environments, protect the eyes from bright sunlight, maintain good dental hygiene, and be aware of symptoms suggestive of lymphoma. They should avoid medications and substances that may worsen their sicca symptoms. Women of child-bearing age with positive anti-SSA/Ro and anti-SSB/La antibodies should be counseled regarding the risk of congenital heart block in the fetus. Patients with loss of smell associated with SS are at an increased risk of injury in the event of gas leakage. Alternative methods should be employed to detect gas leakage in the patient's environment.

LYMPHOMA IN SJÖGREN SYNDROME

Risk Factors

- Increase in risk with time: cumulative risk 3.4% at 5 years and 9.8% at 15 years after diagnosis of Sjögren syndrome (SS)
- Persistent enlargement of parotid glands
- Splenomegaly and lymphadenopathy
- Palpable purpura
- Leg ulcers secondary to vasculitis
- Mixed cryoglobulinemia
- Low complement levels
- CD4 lymphocytopenia

Staging and Management (in Collaboration With an Oncologist)

- Computed tomography (CT) of neck, thorax, and abdomen
- Laboratory tests: Lactate dehydrogenase (LDH), serum and urine electrophoresis and immunofixation, human immunodeficiency virus (HIV), and hepatitis C serology
- Bone marrow biopsy
- Localized low-grade mucosa-associated lymphoreticular tissue (MALT) lymphoma (most common): careful monitoring required
- Multiple extranodal site involvement: single agent chemotherapy
- High-grade transformation or aggressive lymphoma at presentation: rituximab and CHOP (cyclophosphamide, hydroxy doxorubicin, Oncovin, and prednisone)

TRANSLATIONAL RESEARCH AND FUTURE DIRECTIONS

ON THE HORIZON

- Whole transcriptome and exome studies will lead to better understanding of genetic risk factors.
- Further studies are needed to better understand the interaction between immune and nonimmune abnormalities leading to Sjögren syndrome.
- Pilot clinical studies are needed to identify potential therapeutic candidates for larger efficacy trials.
- There is a pressing need to identify and to validate clinically relevant biomarkers.

SS is the clinical manifestation of a complex interplay between genetic factors and environmental and stochastic events that involve innate and adaptive immunity, hormonal mechanisms, and the ANS. A better understanding of these elements is necessary to develop more effective treatments. Two large genome-wide association studies are currently underway to better delineate the genetic risk factors of SS. Pilot treatment trials targeting key cells (B and T lymphocytes) and mediators (Blyss, lymphotoxin, IFN) of autoimmune/inflammatory pathways are expected to identify the most promising molecule(s) that can be tested in larger efficacy studies. There is a clear and present need to identify and validate biomarkers that can be used in clinical trials as well as everyday clinical practice to improve the management of patients with SS.

As no effective treatment is currently available to treat patients with SS, several medications and other non–drug-based modalities are being studied. Currently, several clinical trials using agents targeting IFN, interleukin-6 (IL-6), Toll-like receptors (TLRs), and inhibition of T-cell stimulation are underway. Use of steroids locally in the salivary glands showed some promising initial results, which should be confirmed in independent studies.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Venables PJ. Sjogren's syndrome. Best Pract Res Clin Rheumatol 2004;18:313–29.
2. Skopouli FN, Moutsopoulos HM. Autoimmune epitheliitis: Sjogren's syndrome. Clin Exp Rheumatol 1994;12:S9–11.
3. Mavragani CP, Moutsopoulos HM. The geoepidemiology of Sjogren's syndrome. Autoimmun Rev 2010;9:A305–10.
4. Fox RI. Sjogren's syndrome. Lancet 2005;366:321–31.
5. Scofield RH. Genetics of systemic lupus erythematosus and Sjogren's syndrome. Curr Opin Rheumatol 2009;21:448–53.
6. Burbelo PD, Ambatipudi K, Alevizos I. Genome-wide association studies in Sjogren's syndrome: what do the genes tell us about disease pathogenesis? Autoimmun Rev 2014;13:756–61.
7. Lessard CJ, Adrianto I, Ice JA, et al. Variants at multiple loci implicated in both innate and adaptive immune responses are associated with Sjögren's syndrome. Nat Genet 2013;45:1284–92.
8. Li Y, Zhang K, Chen H, et al. A genome-wide association study in Han Chinese identifies a susceptibility locus for primary Sjogren's syndrome at 7q11.23. Nat Genet 2013;45:1361–5.
9. Voulgarelis M, Tzioufas AG. Pathogenetic mechanisms in the initiation and perpetuation of Sjogren's syndrome. Nat Rev Rheumatol 2010;6:529–37.
10. Hernandez-Molina G, Leal-Alegre G, Michel-Peregrina M. The meaning of anti-Ro and anti-La antibodies in primary Sjogren's syndrome. Autoimmun Rev 2011;10:123–5.
11. Bowman SJ. Patient-reported outcomes including fatigue in primary Sjogren's syndrome. Rheum Dis Clin North Am 2008;34:949–62, ix.
12. Aps JK, Martens LC. Review: the physiology of saliva and transfer of drugs into saliva. Forensic Sci Int 2005;150:119–31.
13. Humphrey SP, Williamson RT. A review of saliva: normal composition, flow, and function. J Prosthet Dent 2001;85162–9.
14. Castro I, Sepulveda D, Cortes J, et al. Oral dryness in Sjogren's syndrome patients. Not just a question of water. Autoimmun Rev 2013;12:567–74.
15. Chai J, Logigian EL. Neurological manifestations of primary Sjogren's syndrome. Curr Opin Neurol 2010;23:509–13.
16. Bejot Y, Osseby GV, Ben Salem D, et al. Bilateral optic neuropathy revealing Sjogren's syndrome. Rev Neurol (Paris) 2008;164:1044–7.
17. Park JH, Hwang J, Min JH, et al. Presence of anti-Ro/SSA antibody may be associated with anti-aquaporin-4 antibody positivity in neuromyelitis optica spectrum disorder. J Neurol Sci 2015;348:132–5.
18. Qiao L, Luo Y, Zhang LL, et al. [The clinical characteristics of primary Sjogren syndrome with neuromyelitis optica]. Zhonghua Nei Ke Za Zhi 2013;52:745–8.
19. Fox RI, Liu AY. Sjogren's syndrome in dermatology. Clin Dermatol 2006;24:393–413.
20. Hatron PY, et al. Pulmonary manifestations of Sjogren's syndrome. Presse Med 2011;40:e49–64.
21. Vassiliou VA, Moyssakis I, Boki KA, Moutsopoulos HM. Is the heart affected in primary Sjogren's syndrome? An echocardiographic study. Clin Exp Rheumatol 2008;26:109–12.
22. Lin DF, Yan SM, Zhao Y, et al. Clinical and prognostic characteristics of 573 cases of primary Sjogren's syndrome. Chin Med J 2010;123:3252–7.
23. Atzeni F, Sarsi-Puttini P, Signorelllo MC, et al. New parameters for identifying subclinical atherosclerosis in patients with primary Sjogren's syndrome: a pilot study. Clin Exp Rheumatol 2014;32:361–8.
24. Sezis Demirci M, Karabulut G, Gungor O, et al. Is There an Increased Arterial Stiffness in Patients with Primary Sjogren's Syndrome? Intern Med 2016;55:455–9.
25. Valim V, Gerdts E, Jonsson R, et al. Atherosclerosis in Sjogren's syndrome: evidence, possible mechanisms and knowledge gaps. Clin Exp Rheumatol 2016;34:133–42.
26. Houghton K, Maleson P, Cabral D, et al. Primary Sjogren's syndrome in children and adolescents: are proposed diagnostic criteria applicable? J Rheumatol 2005;32:2225–32.
27. Voulgarelis M, Moutsopoulos HM. Mucosa-associated lymphoid tissue lymphoma in Sjogren's syndrome: risks, management, and prognosis. Rheum Dis Clin North Am 2008;34:921–33, viii.
28. Vitali C, Bombardien S, Jonsson R, et al. Classification criteria for Sjogren's syndrome: a revised version of the European criteria proposed by the American-European Consensus Group. Ann Rheum Dis 2002;61:554–8.
29. Cornec D, Saraux A, Cochener B, et al. Level of agreement between 2002 American-European Consensus Group and 2012 American College of Rheumatology classification criteria for Sjogren's syndrome and reasons for discrepancies. Arthritis Res Ther 2014;16:R74.
30. Hernandez-Molina G, Avila-Casado C, Nunez-Alvarez C, et al. Utility of the American-European Consensus Group and American College of Rheumatology Classification Criteria for Sjogren's syndrome in patients with systemic autoimmune diseases in the clinical setting. Rheumatology (Oxford) 2015;54:441–8.
31. Rasmussen A, Ice JA, Li H, et al. Comparison of the American-European Consensus Group Sjogren's syndrome classification criteria to newly proposed American College of Rheumatology criteria in a large, carefully characterised sicca cohort. Ann Rheum Dis 2014;73:31–8.
32. Gallo A, Baldini C, Teos L, et al. Emerging trends in Sjogren's syndrome: basic and translational research. Clin Exp Rheumatol 2012;30:779–84.
33. Gallo A, Alevizos I. Isolation of circulating microRNA in saliva. Methods Mol Biol 2013;1024:183–90.
34. Michael A, Bajiracharya SD, Yuen PS, et al. Exosomes from human saliva as a source of microRNA biomarkers. Oral Dis 2010;16:34–8.
35. Willams AE, Choi K, Chan AL, et al. Sjögren's syndrome-associated microRNAs in CD14+ monocytes unveils targeted TGFβ signaling. Arthritis Res Ther 2016;18:1–13.
36. Alevizos I, Alexander S, Turner RJ, Illej GG. MicroRNA expression profiles as biomarkers of minor salivary gland inflammation and dysfunction in Sjogren's syndrome. Arthritis Rheum 2011;63:535–44.
37. Cotrim AP, Alevizos I. Human and viral microRNA expression in Sjogren syndrome. J Rheumatol 2014;41:2102–3.
38. Zero DT, Brennan MT, Daniels TE, et al. Clinical practice guidelines for oral management of Sjogren disease: dental caries prevention. J Am Dent Assoc 2016;147:295–305.
39. Ramos-Casals M, Tzioufas AG, Stone JH, et al. Treatment of primary Sjogren syndrome: a systematic review. JAMA 2010;304:452–60.
40. Papageorgiou A, Voulgarelis M, Tzioufas AG. Clinical picture, outcome and predictive factors of lymphoma in Sjögren syndrome. Autoimmun Rev 2015;14:641–9.

MULTIPLE-CHOICE QUESTIONS

1. What is the typical salivary gland histological feature in Sjögren syndrome (SS)?
   A. Diffuse inflammation throughout the salivary gland parenchyma
   B. Clusters of tightly packed lymphocytes surrounding ducts
   C. Increased proliferation of acinar cells
   D. Increased proliferation of ductal cells

2. Which of the following is NOT an exclusion criterion for the classification/diagnosis of SS?
   A. History of head and neck radiation treatment
   B. Hepatitis C infection
   C. Acquired immunodeficiency syndrome (AIDS)
   D. Mucosa-associated lymphoreticular tissue (MALT) lymphoma

3. Which of the following is NOT a risk factor for the development of lymphoma in SS?
   A. Increased CD4 counts
   B. Persistent enlargement of parotid glands
   C. Mixed cryoglobulinemia
   D. Low complement levels

Scleroderma–Systemic Sclerosis

John Varga, Fredrick M. Wigley

Systemic sclerosis (SSc) is a chronic multisystem connective tissue disease characterized by autoimmunity and inflammation, widespread functional and structural abnormalities in small blood vessels, and progressive fibrosis of the skin and visceral organs. Multiple cell types and their products interact to mediate the pathogenetic processes that underlie the diverse clinical manifestations of SSc.

PREVALENCE AND EPIDEMIOLOGY

SSc is a sporadic disease with worldwide distribution. Incidence estimates in the United States range from 9 to 19 cases per million per year, and prevalence rates range from 28 to 253 cases per million. According to the revised American College of Rheumatology classification criteria, which are more sensitive for identifying early stage of disease, the prevalence estimates are expected to be considerably higher.[1] Age, gender, and ethnicity are important factors that determine disease susceptibility.[2] Like other connective tissue diseases, SSc is more prevalent in women, with the most common age of onset in the range of 40–60 years. Disease onset tends to occur at an earlier age among patients of African descent than among whites. Furthermore, black patients are more likely to have diffuse skin involvement, digital ulcers, pulmonary hypertension (PH), and pulmonary fibrosis, and have a worse prognosis.

ETIOLOGY AND PATHOGENESIS

The pathogenesis of SSc involves dynamic interplay among inherited genetic risk factors, environmentally induced stable epigenetic modifications, and acute or subacute environmental exposures.

Genetic Factors

Systemic sclerosis is not inherited in a mendelian fashion, and disease concordance rates among both monozygotic and dizygotic twins are relatively low (<5%). Nonetheless, 1.6% of patients with SSc have a first-degree relative with the disease (relative risk of 13), indicating an important role for genetic background in SSc disease susceptibility. Specific human leukocyte antigen (HLA) haplotypes show strong association with distinct SSc-specific autoantibody responses. Genetic investigations in SSc have focused on candidate genes and genome-wide association studies (GWAS). Candidate genes shown to be associated with SSc include those implicated in interferon (IFN) signaling, T- and B-cell activation, DNA clearance, and innate immunity (Table 55.1). It is remarkable that a majority of these genes are shared with other autoimmune diseases, particularly systemic lupus erythematosus (SLE), and are involved in immune regulation and innate immune responses, highlighting the potential importance of immune dysregulation in the pathogenesis of SSc.[3]

Environmental Factors

Although the etiology if SSc is unknown, microbial exposures and exposure to environmental and occupational agents, dietary factors, and drugs have been implicated as potential triggering factors. Patients with SSc have increased serum levels of antibodies directed against the human cytomegalovirus (hCMV). Evidence for a potentially pathogenic role for CMV, Epstein-Barr virus (EBV), and parvovirus B19 infection or reactivation, as well as *Helicobacter pylori* infection, has also been presented. These studies need confirmation, and the etiological role of microbial pathogens in SSc remains a conjecture. Although reports of apparent geographical clustering of SSc cases suggest shared environmental exposures, careful investigations have generally not been able to substantiate these clusters. Epidemic outbreaks of apparently novel multisystemic illnesses with SSc-like features have been linked to environmental exposures, such as contaminated rapeseed cooking oils in Spain (the toxic oil syndrome) and L-tryptophan dietary supplements (eosinophilia–myalgia syndrome) in the United States.[4] The incidence of SSc is increased among miners and others with occupational exposure to silica. Additional occupational exposures linked to increased risk of SSc include polyvinyl chloride, trichloroethylene, and organic solvents. Drugs potentially implicated in SSc-like illnesses include bleomycin, taxane, pentazocine, cocaine, and anorexigens associated with pulmonary artery hypertension. The apparent association of SSc with silicone breast implants originally raised concern, but epidemiological investigations have not substantiated the hypothesis of an increased risk.[5]

Pathology

The hallmark pathological features of SSc are a noninflammatory obliterative microangiopathy in multiple vascular beds and fibrosis of the skin and internal organs. In early-stage disease, cellular infiltrates may be prominent in many organs. In skin, inflammatory cells are located predominantly around blood vessels. With disease progression, these infiltrates become sparse.

Vascular injury is the earliest and possibly primary event in the pathogenesis of SSc. Patients with established SSc show widespread vascular lesions characterized by bland intimal proliferation in the small- and medium-sized arteries (Fig. 55.1). In late stages, perivascular adventitial fibrosis and generalized capillary rarefaction are prominent.

TABLE 55.1 Genetic Associations With Systemic Sclerosis (SSc)

Gene Locus	Gene Function	Genetic Polymorphisms Associated With SSc
IRF5	Activation of interferon	rs2004640, others
IRF8	Monocyte differentiation	rs11642837
IRF7	Activation of interferon	rs4963128
IL12R	Interleukin-12/T-cell signaling	rs3790567
STAT4	T-cell signaling	rs7574865, others
DNASE1L3	DNA clearance	rs35677470
ATG	Autophagosome biogenesis	rs9373839
PPARG	Adipogenesis	rs310746
CD247	T-cell receptor signaling	rs2056626
CSK	Src family tyrosine kinase	rs1378942
PTPN22	T-cell signaling phosphatase	rs2476601
BANK1	B-cell receptor signaling	rs10516487, others
BLK	B-cell receptor signaling	rs2736340
TNFAIP3/A20	Negative regulation of nuclear factor (NF)-κB inflammation	rs5029939, others
TNIP1	Negative regulation of NF-κB inflammation	rs2233287, others
TNFSF4	T cell–antigen-presenting cell interaction	rs1234314, others

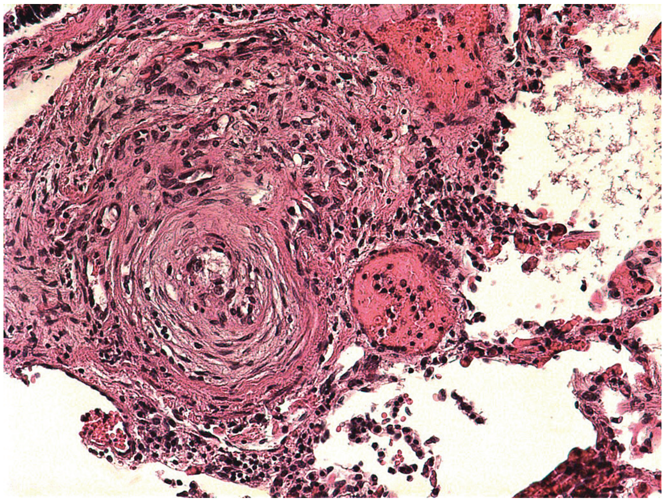


FIG 55.1 Pulmonary Arterial Involvement. Significant intimal layer hyperplasia is seen, leading to narrowing of the vascular lumen. (Courtesy of Dr. Anjana Yeldandi.)

Fibrotic changes are most prominent in skin, lungs, gastrointestinal (GI) tract, heart, tendon sheath, and perifascicular tissue surrounding skeletal muscle. Accumulation of collagen-rich connective tissue composed of fibulins; elastin; proteoglycans; matricellular proteins, such as tenascin-C and alternatively spliced fibronectin (EDA isoform), and other structural macromolecules in these organs leads to distortion of tissue architecture, resulting in progressive functional impairment. In skin, fibrosis causes massive dermal expansion with obliteration of hair follicles and sweat and sebaceous glands. Collagen deposition invades the subjacent adipose layer with entrapment and obliteration of adipocytes. Lungs show patchy infiltration of the alveolar walls with lymphocytes, plasma cells, macrophages, and eosinophils in early disease. In later stages, fibrosis and vascular damage predominate in the lungs, often coexisting in the same lesions. However, in patients with limited cutaneous SSc, the vascular lesions typically predominate. Intimal thickening of the pulmonary arteries, best seen with elastin stain, is a pathological hallmark of PH and at autopsy is often associated with multiple pulmonary emboli. Progressive fibrous thickening of the alveolar septae results in obliteration of the air spaces with a characteristic nonspecific interstitial pneumonia (NSIP) pattern and, less commonly, honeycombing, as well as loss of the pulmonary blood vessels. This process impairs gas exchange and contributes to worsening of PH. In the GI tract, pathological changes can occur at any level, from the mouth to the rectum. The esophagus shows atrophy of the lamina propria, submucosa, and muscular layers, with variable fibrosis. Replacement of the normal intestinal architecture leads to disordered peristaltic activity, resulting in gastroesophageal reflux, small bowel dysmotility, and bacterial overgrowth. Chronic gastroesophageal reflux leads to esophageal inflammation, ulcerations and stricture formation, and, in some cases, premalignant Barrett metaplasia.

The heart is frequently affected, with prominent myocardial contraction band necrosis, which reflects ischemia–reperfusion injury, and patchy areas of myocardial fibrosis. In the kidneys, vascular lesions predominate, and glomerulonephritis is rare. Chronic renal ischemia is associated with shrunken glomeruli and interstitial fibrosis. Scleroderma renal crisis (SRC) is associated with reduplication of elastic lamina, marked intimal proliferation, and narrowing of the lumen (onion skinning), often accompanied by microangiopathic hemolysis.

Pathogenesis

A comprehensive view of SSc pathogenesis must be capable of integrating the vasculopathy, immune dysregulation, and fibrosis of multiple organs, which are the hallmarks of the disease. As illustrated in Fig. 55.2, complex and dynamic interplay among these distinct pathomechanistic processes initiates, amplifies, and sustains tissue damage in SSc.[6] Animal models of disease (see Table 55.2) can be informative for identifying cell types, molecular mechanisms, and pathways contributing to SSc pathogenesis.

Microangiopathy

Evidence of vascular involvement is an early and widespread feature of SSc, and vascular damage has major impact on the course of the disease. Vascular endothelial cell injury is initially associated with largely functional and potentially reversible alterations. There is abnormal blood flow response to vasomotor or cold challenge and altered production of and responsiveness to factors mediating vasodilatation (nitric oxide and prostacyclins) and vasoconstriction (endothelins). Microvessels show loss of pericyte coverage, increased permeability, enhanced transendothelial leukocyte migration, activation of fibrinolytic cascades, and platelet aggregation culminating in thrombosis. Activated endothelial cells express elevated surface adhesion molecules and release endothelin-1, which further promotes leukocyte adhesion and smooth muscle cell proliferation, and might also drive endothelial–mesenchymal transition. Ensuing intimal and medial hypertrophy and fibrosis of the adventitial layers lead to vessel stiffening and luminal narrowing. Combined with endothelial cell apoptosis, the process culminates in the characteristically striking absence of blood vessels seen on angiograms of the

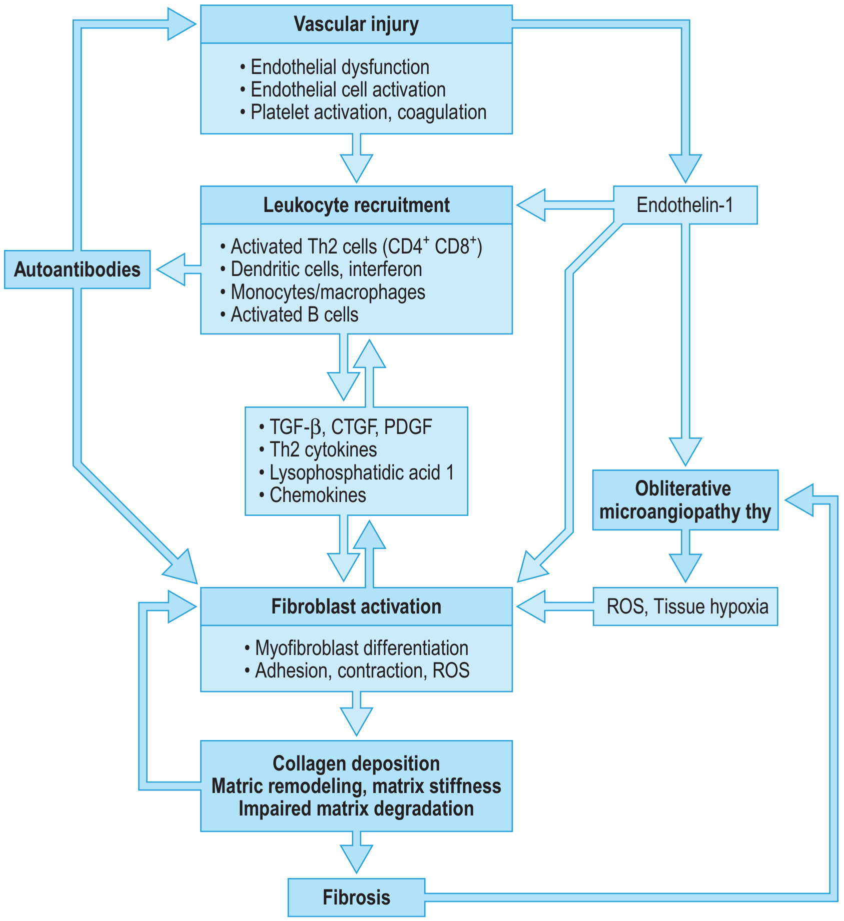

Vascular injury
• Endothelial dysfunction
• Endothelial cell activation
• Platelet activation, coagulation
Leukocyte recruitment
• Activated Th2 cells (CD4+ CD8+)
• Dendritic cells, interferon
• Monocytes/macrophages
• Activated B cells
Endothelin-1
Autoantibodies
• TGF-β, CTGF, PDGF
• Th2 cytokines
• Lysophosphatidic acid 1
• Chemokines
Obliterative microangiopathy thy
ROS, Tissue hypoxia
Fibroblast activation
• Myofibroblast differentiation
• Adhesion, contraction, ROS
Collagen deposition
Matric remodeling, matrix stiffness
Impaired matrix degradation
Fibrosis

FIG 55.2 **Pathogenesis of Systemic Sclerosis.** Interactions of cellular and molecular events triggered by injury that underlie the pathogenesis of vascular and immune dysfunction, culminating in fibrosis. TGF-β,transforming growth factor-β; CTGF, connective tissue growth factor; PDGF, platelet-derived growth factor; ROS, reactive oxygen species.

TABLE 55.2 Selected Animal Models of SSc

Animal Model	Fibrosis	Inflammation	Vascular	Affected Organ	Comment
Bleomycin (subcutaneous)	+	+	±	Skin, lungs	Fibrosis self-limited; localized
Graft-versus-host disease (GvHD)	+	+	–	Skin, lung, kidney	Complex procedure; radiation
Ang2-induced	+	+	+	Skin, lung	Skin localized
Fra2 tg	+	+	+	Skin, lung	Pulmonary hypertension (PH); early mortality
PDGFR tg	+	–	–	Multiorgan	
Wnt10b tg	+	–	–	Skin	Loss of dermal white adipose tissue (dWAT)
TSK1	+	–	–	Hypodermal lungs—emphysema	Fibrillin-1 mutation
TSK2	+	+	–	Skin	Collagen III mutation
Stiff skin syndrome	+	+	–	Skin	Fibrillin-1 mutation
Fli1-deleted	+	–	RVH	Skin, heart	
ROS	+	+	+	Skin, lung, kidney	Not widely used
Topo 1 immunization	+	+	–	Skin	Not well established
Chronic GvHD	+	+	±	Skin	Immune-driven

hands and kidneys in late-stage disease. Paradoxically, despite elevated levels of angiogenic signals, including vascular endothelial growth factor (VEGF), reflecting pervasive tissue hypoxia, the process of revascularization appears to be defective in SSc, possibly as a result of defects in the generation and maturation of bone marrow–derived $CD34^+$ endothelial progenitor cells.[7] Vascular injury in SSc is particularly prominent in the kidneys of patients with SRC, with evidence of complement activation and endothelin-1 deposition in the damaged small vessels.

Cellular and Humoral Immune Responses

In the early stages of SSc, activated T cells and monocytes/macrophages accumulate within lesional skin and the lungs, where they secrete proinflammatory and profibrotic mediators, including transforming growth factor-β (TGF-β), cytokines, and chemokines. These molecules can activate fibroblasts as well as endothelial cells, recruit further inflammatory cells, and initiate the fibrotic response. Because TGF-β, in particular, can induce its own production as well as that of other profibrotic growth factors, such as connective tissue growth factor (CTGF) and platelet-derived growth factor (PDGF), an initial cytokine burst could result in amplified cytokine production and sustained autocrine and paracrine signaling. Additionally, as the extracellular matrix (ECM) undergoes remodeling, it loses compliance and increases its stiffness, which, in itself, triggers biomechanical activation of resident stromal cells, in part by liberating and activating latent matrix-bound latent TGF-β. Alterations in the relative proportions and function of regulatory T cells (Tregs) and T-helper 17 (Th17) cells, and innate lymphoid cells (ILCs), have been documented and may play important roles in pathogenesis.

Virtually all patients with SSc have antinuclear and other autoantibodies in serum. These autoantibodies are generally highly specific for SSc and tend to be mutually exclusive (see below). Moreover, SSc-specific autoantibodies show strong association with individual disease subphenotypes, and their titers can fluctuate with disease activity. Multiple mechanisms have been proposed to account for autoantibody generation in SSc. In patients with SSc, B cells overexpress the surface receptor CD19, resulting in hyperresponsives.[8] Additionally, specific self antigens undergo posttranslational modifications, such as proteolytic cleavage, increased expression or misdirected subcellular localization that create neoepitopes recognized as nonself by the immune system. Although SSc-associated autoantibodies have well-established clinical utility as diagnostic and prognostic markers, their direct role in SSc-associated tissue damage remains uncertain.[9] Recent studies have indicated that patients with SSc also commonly have functional antibodies that are directed against fibroblasts and endothelial cells; the PDGF receptor; vascular receptors, such as endothelin-1 and angiotensin II receptors; and matrix metalloproteinases (MMPs).

Fibrosis: Cellular and Molecular Components

Interstitial and vascular fibrosis, the hallmarks of SSc, is characterized by replacement of normal tissue architecture with dense and noncompliant connective tissue. Although isolated organ fibrosis is a generalized and frequent sequel to any form of chronic or recurrent tissue injury, fibrosis concurrently affecting multiple organs is unique to SSc. Fibroblasts and related stromal cells of mesenchymal origin are key effector cells responsible for the development of fibrosis. In response to extracellular signals, such as TGF-β, CTGF, PDGF, Wnt, chemokines, endothelin-1, lysophosphatidic acid (LPA1), reactive oxygen species (ROS), and hypoxia, as well as integrin-mediated biomechanical signals from a stiff ECM, these cells proliferate; migrate; produce a broad array of matrix macromolecules; adhere to and remodel connective tissue; secrete growth factors, cytokines, and chemokines and express surface receptors for them; and undergo transdifferentiation into apoptosis-resistant contractile myofibroblasts. Under physiological conditions, the fibroblast repair program is tightly regulated and self-limiting to enable tissue regeneration, whereas under pathological conditions, fibroblast activation is sustained and amplified, resulting in exaggerated matrix deposition and disruption of tissue architecture. Uncontrolled activation of fibroblasts, combined with their relative resistance to apoptosis, is a fundamental pathogenetic alteration underlying fibrosis in SSc.

Fibroblasts explanted from lesional SSc tissues continue to display an abnormal phenotype when propagated *ex vivo*, with enhanced rate of synthesis of collagen and other ECM molecules, expression of α smooth muscle actin and cell surface adhesion molecules, and spontaneous generation of ROS through the nicotinamide adenine dinucleotide phosphate (NADPH) oxidase complex pathway as well as from mitochondrial respiration. Additionally, explanted SSc fibroblasts show hallmarks of senescence. The persistent activated scleroderma phenotype is likely to reflect epigenetic modifications caused by chromatin remodeling, DNA methylation, or altered expression of noncoding regulatory RNAs, including microRNA and long noncoding RNAs.

TGF-β is a pleiotropic cytokine and a pivotal regulator of tissue repair and fibrosis (Chapter 9). Multiple abnormalities in TGF-β pathways have been identified in SSc, indicating a key role in pathogenesis. Lesional fibroblasts secrete TGF-β and exhibit TGF-β hyperresponsiveness caused by elevated expression of TGF-β surface receptors, and integrin-mediated activation of latent TGF-β. Furthermore, SSc fibroblasts show evidence of constitutive phosphorylation of Smad transcription factors and other key intracellular signaling pathways that, together with deficiencies in endogenous antifibrotic mechanisms, might contribute to the persistence and progression of fibrosis.[10] Therapies that selectively or nonselectively block growth factor signaling and abrogate fibrogenic pathways triggered by growth factors, chemokines, PDGF, Wnt, and CTGF are currently under development.

CLINICAL FEATURES

Overview

The term “scleroderma” is used to describe both systemic sclerosis and localized scleroderma, a disorder generally limited to skin. SSc is highly variable in its clinical expression and thus represents a broad spectrum of disease. The disease process can target skin, blood vessels, and the lungs, heart, GI tract, kidneys, and musculoskeletal system, subsets of patients exist with unique clinical features and distinct clinical outcomes. Raynaud phenomenon is virtually universal in SSc, suggesting that perturbation of the terminal arteries of the circulation is a fundamental process that links the different subsets of the disease. Thickening of the skin distinguishes SSc from other rheumatic diseases (Table 55.3); scleroderma (hard skin) is the most specific and prominent physical finding. Patients vary in the expression of the skin changes and are classified into two major subsets defined by the degree of clinically involved skin. In the diffuse skin variant called *diffuse cutaneous SSc* (dcSSc), fibrosis of skin is widespread, and the illness is more violent in expression, with rapid onset and increased risk of serious and internal organ disease. In contrast, the limited skin variant called *limited cutaneous SSc* (lcSSc) presents with long-standing Raynaud phenomenon, skin fibrosis limited to the fingers or distal limbs, and a generally indolent disease course. Limited SSc is less likely to be associated with organ failure or shortened life expectancy. Predictors of elevated mortality rates among patients with SSc include diffuse skin disease with a rapid rate of skin progression and internal organ involvement (especially the lungs or the kidneys, resulting in an SRC), male gender, black race, and later age of disease onset.[11,12]

KEY CONCEPTS

Clinical Patterns of Systemic Sclerosis

Limited Cutaneous Scleroderma

- Distal skin sclerosis (sclerodactyly)
- Severe Raynaud phenomenon with digital ischemia
- Numerous telangiectasia
- Isolated pulmonary arterial hypertension

Diffuse Cutaneous Scleroderma

- Rapidly progressing widespread skin involvement
- Scleroderma renal crisis
- Interstitial lung disease
- Severe gastrointestinal dysmotility
- Cardiomyopathy

TABLE 55.3 Classification of Systemic Sclerosis (SSc)

Diffuse cutaneous scleroderma—skin thickening on the trunk in addition to the face, proximal and distal extremities

Limited cutaneous scleroderma—skin thickening limited distal to the elbow and knee; may also involve the face and neck

CREST syndrome—subcutaneous *c*alcinosis; *R*aynaud phenomenon; *e*sophageal dysmotility; *s*clerodactyly; *t*elangiectasia

Sine scleroderma—no apparent skin thickening but characteristic visceral organ involvement, vascular and serological features

Overlap syndrome—criteria for SSc, coexisting with features of systemic lupus erythematosus, rheumatoid arthritis, or inflammatory muscle disease

Mixed connective tissue disease—overlap syndrome with anti-U1 ribonucleoprotein (RNP) antibodies

Early disease—Raynaud phenomenon (RP) with clinical and/or laboratory features of SSc; specific autoantibodies, abnormal nailfold capillaroscopy, finger edema, and ischemic injury

Although classification of SSc as diffuse and limited cutaneous has utility, disease expression is far more complex, and several distinct phenotypes are recognized within each of the two subsets. For example, 10–15% of patients with lcSSc develop severe pulmonary arterial hypertension without significant lung fibrosis. Other patients develop systemic features of SSc without appreciable skin involvement, a phenotype that is termed *systemic sclerosis sine scleroderma.* Unique clinical phenotypes of SSc are associated with specific autoantibodies (Table 55.4). For example, anticentromere antibodies are associated with lcSSc, whereas antitopoisomerase-I antibodies are associated with the presence of interstitial lung disease. The presence of anti-RNA polymerase antibodies is associated with rapid and widespread skin disease and an increased risk of an SRC. In some patients with SSc who have "overlap" features, SSc coexists with clinical and laboratory evidence of another autoimmune disease, such as polymyositis, autoimmune thyroid disease, Sjögren syndrome, polyarthritis, autoimmune liver disease, or SLE (see Table 55.4).

Instead of SSc, the term " localized scleroderma" is now properly used to describe a group of patients with disease generally limited to fibrosing skin disorders that occurs with similar frequency in both children and adults (Table 55.5). Localized scleroderma is infrequently associated with significant systemic involvement. Morphea, a form of localized scleroderma, can occur as solitary or multiple lesions (generalized morphea) of expanding circular patches of thickened skin. The most common form of localized scleroderma in children is linear scleroderma. It presents as a streak of thickened skin, generally in one or both lower extremities. Linear scleroderma often involves the subcutaneous tissues with fibrosis and atrophy of supporting structures, muscle, and bone, and radiographs may show melorheostosis of long bones. In children with linear scleroderma, the growth of affected tissues, muscles, and long bones can be retarded. When the lesions cross joints, significant contractures of the affected joint can develop. Linear scleroderma can occur on the

TABLE 55.4 Frequency and Clinical Correlations of Systemic Sclerosis (SSc) Autoantibodies

	% Frequency in SSc	Disease Subtype	Clinical Associations	Prognosis
Anticentromere	20–38	lcSSc	Pulmonary arterial hypertension	Better prognosis
Antitopoisomerase I	15–42	dcSSc	Pulmonary fibrosis Heart involvement	Worse prognosis
Anti-RNA polymerase III	5–31	dcSSc	Renal crisis Tendon friction rubs, synovitis, myositis, Joint contractures. Increased risk of malignancy	Increased mortality
Anti-U3 RNP (fibrillarin)	4–10	dcSSc	Renal crisis and cardiac involvement	Poor prognosis especially in African Americans
Anti-Th/To	1–13	lcSSc	Pulmonary fibrosis and renal crisis	Poor prognosis
Anti-U11/U12 RNP	3.2	—	RP Gastrointestinal involvement Lung fibrosis	Increased mortality
Anti-U1 RNP	2–14	lcSSc	RP, puffy fingers, arthritis, myositis, overlap syndrome (*i.e.,* MCTD)	Better prognosis
Anti-PM-Scl	4–11	Overlap with polymyositis lcSSc	RP, arthritis, myositis pulmonary involvement, calcinosis, and sicca symptoms	Better prognosis
Anti-Ku	2–4	—	Myositis, arthritis, and joint contractures	—
Anti-hUBF (NOR 90)	<5	lcSSc	Mild internal organ involvement	Better prognosis
Anti-Ro52/TRIM21	15–20	Association with other autoimmune diseases	Older age onset, pulmonary fibrosis	—

dcSSc, diffuse cutaneous SSc; lcSSc, limited cutaneous SSc; MCTD, mixed connective tissue disease; RNP; ribonucleoprotein; RP, Raynaud phenomenon; TRIM, tripartite motif.
From Kayser C, Fritzler MJ. Autoantibodies in systemic sclerosis: unanswered questions. Frontiers Immunol 2015;6:167.

TABLE 55.5 Classification of Localized Scleroderma

- Plaque types of morphea
  - Circumscribed plaques
  - Guttate
  - Keloid/nodular
  - Bullous
- Generalized morphea
  - Three or more locations
  - Pansclerotic morphea
- Linear scleroderma
  - Frontoparietal linear morphea (*en coup de sabre*)
  - Parry-Romberg syndrome (progressive hemifacial atrophy)
  - Linear streaks on limbs or trunk
- Deep morphea

head leaving severe facial deformity. When this occurs, it is called *en coupe de sabre.* This should be distinguished from Perry-Rhomberg syndrome, or facial hemiatrophy, which also affects the face, usually unilaterally, with atrophy of skin rather than fibrosis. Both *en coupe de sabre* and Perry-Rhomberg syndrome can be associated with central nervous system (CNS) involvement, presenting as headaches, visual disturbance, or a seizure disorder. Uncommon forms of localized scleroderma include pansclerotic morphea, keloid morphea, deep morphea presenting as a fasciitis, and bullous morphea with blistering. Morphea lesions can be seen in patients with SSc, and often lichen sclerosus is seen in association with morphea.

Symptoms

Characteristically, the earliest symptoms of SSc are nonspecific and include fatigue, musculoskeletal distress (stiffness or pain), and feeling ill. Cold sensitivity and Raynaud phenomenon occur early in the disease and may be the only clinical clue to the presence of active disease. Symptoms of esophageal dysfunction with dysphagia, heartburn, and periodic GI reflux are common problems, which, along with Raynaud phenomenon, can precede other manifestations of SSc by years. Visible capillary abnormalities easily seen at the nailfold (dilatation and loss or dropout of capillaries) are almost a universal finding in SSc. These nailfold capillary changes provide a clinical clue to presence of the underlying vascular disease in that changes occur early in the disease course and progress in association with other systemic involvement. An individual presenting with Raynaud phenomenon, abnormal nailfold capillaries, and the presence of an SSc-specific autoantibody can be suspected of having SSc even before other more obvious manifestations are noted.[13]

CLINICAL PEARLS

Clinical Features of Early Systemic Sclerosis

- Definite Raynaud phenomenon
- Gastroesophageal reflux with heartburn
- Swelling of the fingers and hands
- Musculoskeletal pain and stiffness
- Dilated nailfold capillaries
- Hyper- or hypopigmentary changes of skin

Diffuse SSc

In general, patients with dcSSc have a short interval between the onset of Raynaud phenomenon and other signs and symptoms.

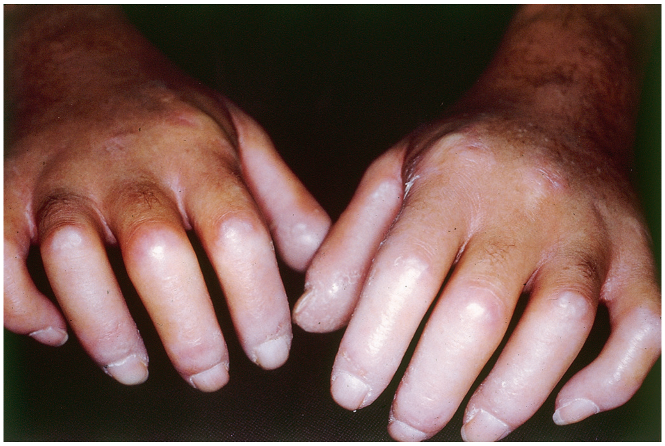


FIG 55.3 Skin Fibrosis in Systemic Sclerosis. Severe fibrosis of the skin of the hands and forearms causing joint contractures and skin ulcerations in a woman with diffuse cutaneous systemic sclerosis.

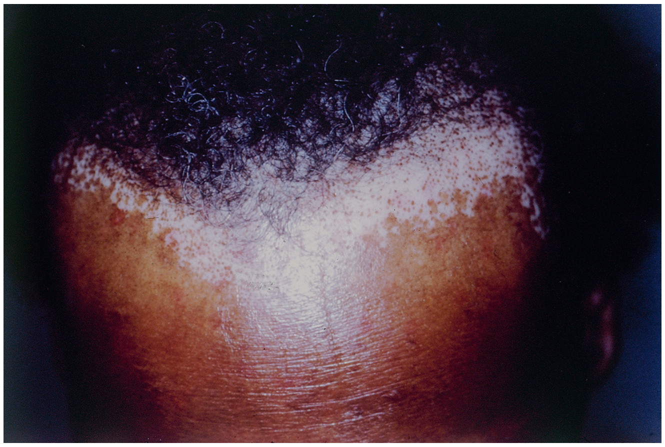


FIG 55.4 Pigmentation Changes in the Skin. Vitiligo (salt-and-pepper) appearance of the involved skin in a patient of African descent with diffuse cutaneous systemic sclerosis.

Soft tissue swelling, intense pruritus and burning, and nonpitting edema are signs of the early inflammatory or "edematous" phase of the diffuse cutaneous form of the disease. The skin of the fingers, hands, distal limbs, and face are usually affected first and more severely compared with other body areas (Fig. 55.3). Carpal tunnel syndrome can be present as a consequence of soft tissue inflammation around hands and wrists. Patients may note skin hyperpigmentation (patches or generalized tanning). Other early skin changes include vitiligo-like hypopigmentation ("salt and pepper" appearance), often on the chest or the back (Fig. 55.4). Escalating musculoskeletal symptoms are common and are associated with muscle weakness and decreased joint mobility. Significant finger and hand/wrist disease limits the patient from doing simple chores or performing self-care.

The early edematous phase of dcSSc has prominent inflammatory features, with significant skin edema and erythema, and is associated with inflammatory cell infiltration in the dermis. After a period of weeks to months, the inflammatory phase evolves into the "fibrotic" phase as increased collagen and extracellular material is deposited in skin, causing thickening and loss of flexibility. The fibrotic process starts in the dermis and is associated with loss of body hair, reduced production of skin oils, and a decline in sweating capacity as these cutaneous structures atrophy. Gradually, the subcutaneous tissue becomes affected, with atrophy of subcutaneous fat and fibrosis extending to the underlying fascia, muscle, and other soft tissue structures. During the fibrotic phase, patients note pain, progressive loss of flexibility, distressing disfigurement, and profound weight loss. Progressive flexion contractures of the finger joints ensue. Other joints, including wrist, elbow, shoulder, hip girdle, knees, and ankle, can also be affected as a result of fibrosis of the supporting joint structures (see Fig. 55.3). Thick ridges at the neck caused by firm adherence of skin to the underlying platysma muscle interfere with neck extension. Tendon friction rub is a prominent crepitation that can be felt or even heard over tendons of the lower and upper extremities. Tendon friction rubs are caused by fibrosis in the tissues surrounding the affected joints; they are associated with rapidly progressive skin disease and are linked to an overall poor prognosis. Fibrosis of the skin of the face yields a characteristic facial appearance with diminished oral aperture, loss of facial expression, vertical lines around the lips, loss of lip thickness, and protrusion of teeth.

The natural history of skin disease in dcSSc tends to be monophasic, and relapse is uncommon after the edematous and active fibrotic phase. The duration of active skin disease from the first signs of skin involvement to its maximal extent is characteristically less than 3 years; cutaneous inflammation and progressive fibrosis gradually subside, and regression of skin involvement begins. However, this time frame is highly variable from one patient to another, with many patients demonstrating rapid skin regression and others showing chronically active disease over years. In the late stages of dcSSc, skin remodeling can be dramatic, with return to normal-appearing skin in those areas spared from severe end-stage fibrosis. In more severely affected areas, the skin becomes thin or atrophic and can be bound down to underlying structures. Skin ulcerations or dystrophic calcification often complicate the fibrotic, atrophic, and avascular skin (Fig. 55.5).

Although the skin involvement is generally the most dramatic and visible manifestation of dcSSc, internal organ involvement occurs during the early active stage of advancing skin disease. Patients with dcSSc have a significant risk for interstitial lung disease, severe GI dysfunction, SRC, progressive heart disease, and recurrent digital ulcers during the initial active inflammatory and fibrotic phases of the skin disease. In practical terms, this means that in dcSSc, the initial 3–4 years is the period that the systemic process is most active; if organ failure does not occur during this period, the systemic process may stabilize without further progression.

Limited Cutaneous SSc

The disease course in the limited cutaneous variant of SSc is more indolent and often relatively benign. After the onset of Raynaud phenomenon, several years may pass before additional symptoms or signs are recognized. The most common non-Raynaud symptoms in patients with lcSSc are those of upper GI disease with dysphagia and gastroesophageal reflux. Dilated capillaries form visible erythematous vascular lesions (telangiectasia), seen most commonly on the fingertips, nailfold areas of the digits, palms, face, lips, and inside the oral cavity (Fig. 55.6). CREST (subcutaneous *c*alcinosis, *R*aynaud phenomenon, *e*sophageal dysfunction, *s*clerodactyly, and *t*elangiectasia) syndrome is a subtype of lcSSc, with distinctive features, an overall good prognosis, and an association with the presence of anti-centromere antibodies. Subcutaneous calcinosis caused by deposition of calcium hydroxyapatite crystals occurs commonly at sites of tissue ischemia and recurrent trauma, such as the fingertips, forearm, or elbow.

Although significant internal organ disease occurs with lower frequency in patients with lcSSc, isolated pulmonary arterial hypertension (PAH) develops in 10–15% of patients and may be life-threatening.[14] Significant interstitial lung disease can occur in about 20% of patients with lcSSc. Severe Raynaud phenomenon with macrovascular occlusive involvement occurs more frequently in lcSSc than in dcSSc and can be associated with critical digital ischemia, ischemic ulcerations, gangrene, and amputation. Overlap with other autoimmune syndromes, including the sicca complex, polyarthritis, cutaneous vasculitis, and autoimmune liver disease, such as biliary cirrhosis, is seen primarily in the lcSSc subset of SSc.

Raynaud Phenomenon

The skin has a unique vascular architecture that is designed to maintain a stable central body temperature. This system has superficial and deep vascular plexuses that allow shunting of

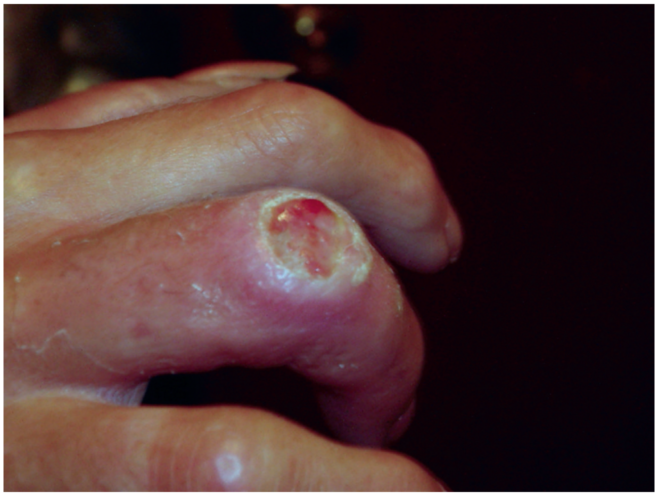


FIG 55.5 Traumatic Digital Ulcer. Ulceration of atrophic skin over the metacarpal joint of patient with diffuse scleroderma.

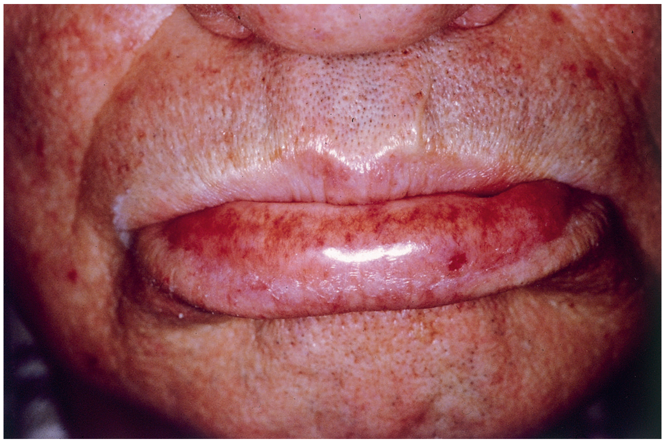


FIG 55.6 Telangiectasia. Characteristic telangiectasia on the lip in a woman with limited cutaneous systemic sclerosis.

blood away from the skin surface during cold exposure and increases surface blood flow during hot ambient temperatures. Arteriovenous anastomoses (AVAs) connecting arterioles and venules allow blood to bypass capillaries, thus allowing rapid and marked increases in cutaneous blood flow. Increased activity of the sympathetic nervous system via both the CNS (body chilling) and the peripheral nervous system (local cooling of skin) initiates vasoconstriction and heat conservation. Selective responses to cold by AVAs allow preservation of nutritional blood flow through skin capillaries. Raynaud phenomenon–associated increase in vascular tone results, in part, from a defect in the vascular response to normal sympathetic signals. Studies find that α_2 adrenergic receptors on the smooth muscle cells of cutaneous vessels mediate vasoconstriction in AVAs during a response to cold temperature. This response to sympathetic input is abnormally increased in patients with Raynaud phenomenon.[15]

Raynaud phenomenon is clinically defined as well-demarcated color changes of the digits induced by cold or emotional stress. Vascular constriction of AVSs, arterioles, and small arteries in the skin and tissues of the digits causes the pallor of the digits, the phase of complete loss of blood flow. The initial vasospasm and pallor is followed by cyanosis of the skin caused by venous pooling and low flow with deoxygenated blood. Finally, after rewarming, the recovery phase occurs with vasodilation manifested by hyperemia with blushed reddened skin as blood flow rebounds. Raynaud phenomenon affects as much as 3–5% of the general population. Raynaud phenomenon is considered to be primary when there is no other associated disease state. In these cases, it is thought to represent a genetic trait with cold sensitivity as a result of abnormal cutaneous vessel reactivity to environmental temperatures. In SSc, the disease process targets the peripheral vasculature, including AVAs or thermoregulatory blood vessels. This leads to not only the abnormal reactivity to ambient temperatures typical of Raynaud phenomenon but also significant structural disease of blood vessels, which causes tissue ischemia as a result of compromise to blood flow within capillaries or nutritional vessels. As a consequence of both an exaggerated response to cold and structural vascular disease, patients with SSc have severe Raynaud phenomenon with multiple, and often prolonged, daily episodes. The severe vascular disease can lead to critical ischemia with digital ulcerations or deep tissue infarction and gangrene. The vascular disease in SSc is not limited to skin but is systemic. There is evidence that abnormal vascular reactivity resulting in tissue injury occurs in the pulmonary, renal, GI, and coronary circulations.

Treatment of Raynaud phenomenon must be individualized and adjusted according to its severity. Initial therapy should include avoidance of cold exposure and use of methods of stress reduction. The best studied medications to treat Raynaud phenomenon are the calcium channel blockers (*e.g.*, nifedipine), which remain the recommended first choice of therapy. Other vasoactive drugs reported to be helpful include the phosphodiesterase inhibitors, such as sildenafil; serotonin reuptake inhibitors, such as fluoxetine; nitrates, such as topical nitroglycerine; and intravenous prostaglandins.[16] Bosentan, an endothelin-1 receptor inhibitor, was recently shown to reduce the incidence of new digital ulcers in patients with SSc but not to alter the frequency of Raynaud events.[17] Statins may also protect peripheral blood vessels and prevent ischemic digital ulcers, and antioxidants have the potential to reduce oxidative stress to tissues. Botulinum toxin injected locally at the base of the fingers is reported to be helpful, but a controlled clinical trial is needed to define its role. Digital sympathectomy is used when medical therapy has failed to reverse critical digital ischemia. In a few cases, macrovascular occlusive disease (particularly in the ulnar artery) can be surgically bypassed to improve distal blood flow.

Gastrointestinal Involvement

GI disease is a major cause of morbidity and mortality and accounts for one of the most frequent initial symptoms of scleroderma. Every part of the GI tract can be involved, including the oral pharynx, esophagus, stomach, and the small and large bowels. A reduced oral aperture (microstomia) is not only cosmetically unpleasant, but it also impairs daily hygiene and normal mastication. Periodontal disease is a consequence of several factors, including decrease salvia, difficulty in dental care, GI efflux disease, gum recession, and bone resorption. Caries and tooth loss can result. Initial symptoms are difficulty chewing and difficulty swallowing dry or a larger bolus of food.

Almost all patients with SSc demonstrate evidence of distal esophageal dysfunction. Clinically, this presents with dysphagia, heartburn, and regurgitation typical of gastroesophageal reflux disease (GERD). These symptoms are caused by peristaltic dysfunction of the distal esophagus, decreased lower esophageal sphincter (LES) pressure, and delayed gastric emptying, which is thought to be secondary to autonomic dysfunction. Pathological studies of the esophagus show that the smooth muscle (circular greater than longitudinal) of the bowel atrophies without significant fibrosis, vascular injury, or obvious inflammation.[18] Functional and pharmacological studies have shown that a neurogenic process precedes smooth muscle dysfunction. Autoantibodies directed against enteric neurons and antimuscarinic antibodies are found, suggesting that an immune process is involved in the pathogenesis of bowel dysmotility. Esophageal manometric evaluation reveals the characteristic triad of low or absent primary and secondary peristaltic activity in the distal esophagus, normal proximal esophageal (striated muscle) motility, and the loss of the LES tone. Complications include reflux esophagitis, esophageal strictures, mucosal erosions, and bleeding; Barrett metaplasia or, very rarely, adenocarcinoma may also develop.

Early satiety, bloating, nausea, periodic vomiting, and decreased appetite with weight loss are seen secondary to poor gastric emptying and retention of food and liquids in the stomach. Gastroparesis is commonly associated with esophageal dysfunction and aggravates GERD. Dilatation of mucosal capillaries in the GI tract is common, particularly in the gastric antrum. Gastric antral vascular ectasia (GAVE), also known as "watermelon stomach," is the major gastric manifestation, which may lead to occult bleeding and significant anemia. Intestinal dysmotility can involve the small and large bowels. Patients may note a change in bowel habits or present with episodes of pseudoobstruction with severe abdominal pain, bloating, abdominal distension, and vomiting. Persistent diarrhea may be a manifestation of malabsorption of fats as a result of bacterial overgrowth in the atonic small bowel and can be associated with dramatic weight loss and malnutrition. In the late stages, the bowel wall thins out and traps air (pneumatosis cystoides intestinalis). Wide-mouth diverticuli develop in the colon or the bowel and may rupture. Incontinence of stool can occur as a result of dysfunction of anal sphincters.

Pulmonary Involvement

Lung disease is now the leading cause of death in patients with SSc and accounts for significant lifetime morbidity. Lung involvement is often asymptomatic and nonprogressive, yet significant lung abnormalities can be detected in over 40% of patients with sensitive lung function testing or special imaging. In a subgroup of patients (10–20%), the disease process can lead to major pulmonary vascular disease or inflammation with progressive fibrosis and subsequent respiratory failure and death. A recent survey reported that 35% of all SSc-related deaths were directly attributable to pulmonary fibrosis or interstitial lung disease (ILD), and 26% of deaths resulted from PAH.[19] Additional pulmonary manifestations of SSc include aspiration pneumonitis, pulmonary hemorrhage, bronchiectasis, bronchiolitis obliterans with organizing pneumonia, pleural reactions, and pneumothorax. The risk of lung cancer is increased in patients with SSc.

Interstitial Lung Disease

Nonspecific interstitial pneumonia (NSIP) is the most common histopathological pattern in patients with SSc-related ILD (SSc-ILD). The prevalence of SSc-ILD is found in 50% of those with dcSSc compared with 35% in those with lcSSc. The presence of topoisomerase-I autoantibodies is strongly linked to the development of ILD, and the worst outcome is often seen in African Americans and Native Americans. Inflammatory alveolitis and subsequent tissue fibrosis causes restrictive ventilatory defect and an abnormal gas exchange. Patients with ILD usually experience rapid initial decline in lung function during the early period of active disease. About 20% of patients who eventually developed severe ILD (FVC ≤50%) have a fall in forced vital capacity (FVC) within 4 years of the onset of symptoms. If there is only minor impairment in FVC on lung function after more than 5 years of disease duration, then is it unlikely that the patient will ever develop severe lung fibrosis. In early disease, most patients have no respiratory symptoms. The most sensitive methods to detect early lung disease are pulmonary function testing and imaging with high-resolution computed tomography (HRCT). The presence of fibrosis and ground-glass opacities on HRCT is an indication of the activity of the disease process and correlates with subsequent decline in FVC. When the extent of disease defined by HRCT is <20%, the 10-year survival is reported to be 67%, whereas in the group with HRCT-defined disease extent of >20%, 10-year survival is considerably poorer at 43%.[20] Likewise, FVC of <70% is associated with a poor prognosis. Biological biomarkers that can detect active alveolitis and predict disease progression are needed. The presence of neutrophilia in fluid obtained from bronchoalveolar lavage (BAL) was thought to predict outcome, but neutrophilia does not clearly define prognosis or response to therapy. Both Krebs von den Lungen-6 (KL-6), a lung glycoprotein, and surfactant protein-D (SP-D) levels are elevated in patients with SSc compared with controls and higher in patients with alveolitis compared with those without alveolitis; thus SP-D levels are thought to be useful biomarkers of disease progression. Two multicenter clinical trials have been completed; one compared cyclophosphamide to placebo, and the other compared cyclophosphamide to mycophenolate.[21,22] Both studies demonstrated statistically significant stabilization or improvement of lung function along with improvement in scleroderma skin score and a variety of clinical measures. These studies suggest that immunosuppressive therapy can control active lung disease. A variety of other agents are now being tested in SSc-ILD, including rituximab (monoclonal antibody [mAb] to CD20 on B lymphocytes), pirfenidone (an antifibrotic/antiinflammatory agent used in idiopathic pulmonary fibrosis), and nintedanib (a tyrosine kinase inhibitor).[23,24]

Pulmonary Arterial Hypertension

PAH is detected in 8–15% of patients with SSc and may occur as the sole pulmonary manifestation of the disease or in association with ILD.[25] The natural history of PAH is highly variable. In many patients, it is clinically silent for years before it presents with exertional dyspnea and hypoxia secondary to altered gas exchange and right heart failure. It is most often a late (9–12 years after onset of Raynaud phenomenon [RP]) complication in patients with limited scleroderma; but PAH can occur within 5 years from the disease onset, particularly in patients older at the time of SSc diagnosis. Risk factors for PAH in SSc include disease onset at a later age, history of severe RP, numerous skin telangiectasias, abnormal nailfold capillaries, the presence of anticentromere, anti-U1 ribonucleoprotein (RNP), anti–U3 RNP (fibrillarin), or anti-B23 antibodies. Untreated, the estimated 3-year survival is approximately 50%. Although responses to current therapy are not as good among patients with SSc and PAH compared with patients with idiopathic PAH, there is evidence of improved survival on modern therapy compared with historical controls. Prognosis is poor among patients with PAH associated with ILD, and an even worse prognosis is associated with increased blood brain natriuretic protein (BNP), echocardiographic signs of right heart disease, and decreased diffusion capacity of the lungs for carbon monoxide (DLCO) on lung testing. PH can occur in SSc in the absence of PAH as a result of left heart disease or secondary to chronic hypoxia from severe ILD. Therefore patients need to undergo a right heart cauterization to confirm the presence of PAH or PH and characterize its severity. Current therapy for SSc-associated PAH is focused on supportive care, reduction of cardiac work load, and vasoactive drugs.[26] Short-acting prostacyclin analogues, including intravenous (epoprostenol or treprostinil) or inhaled (iloprost) prostaglandins, are used for treating severe disease. Selective and nonselective endothelin-1 receptor inhibitors (bosentan, ambrisentan, and macitentan), a phosphodiesterase type 5 inhibitor (sildenafil, tadalafil, and vardenafil), and a soluble guanylate–cyclase stimulator (riociguat) are helpful for treating milder disease. Early intervention and combination therapy may help achieve improved outcomes. In relatively short-term clinical trials, each of these treatments has been shown to improve exercise tolerance and hemodynamics, with variable benefit on disease progression. Early detection of PAH is important as it provides an opportunity for timely intervention. A systematic detection program in which periodic screening was performed in every patient with SSc detected less severe pulmonary vascular disease and showed better survival on therapy compared with that in patients managed by usual clinical practice.[27] Because of the complexity of life-threatening PAH, it is recommended that patients be referred to experts in pulmonary medicine, especially in PAH, to manage these cases. Lung transplantation remains an option for carefully selected patients with SSc, with survival following lung transplantation being similar to that of non-SSc patients with other lung diseases.

Cardiac Involvement

Cardiac involvement is common, but often there are no signs or symptoms until the heart has been severely affected. The reported prevalence of heart disease varies, depending on method of detection. Clinical criteria are insensitive, but modern sensitive tools, such as echocardiography, tissue Doppler imaging, speckle-tracking echocardiography, cardiac magnetic resonance imaging, or thallium scanning, can detect disease early and suggest that 40–60% of patients have heart disease. When heart disease is symptomatic, the prognosis is poor. SSc can affect virtually any cardiac structure, including the myocardium with fibrosis or inflammatory myocarditis, the myocardial microvasculature, or the pericardium with asymptomatic small effusions or large effusions leading to cardiac tamponade. As a consequence of cardiac tissue injury, left ventricular (LV) systolic dysfunction and/or LV diastolic dysfunction and/or right ventricular (RV) failure and/or conduction abnormalities with complex arrhythmias can occur, but valvular heart disease is unusual. Cardiac dysfunction can be primary to the disease process or secondary to other organ diseases, including PAH, ILD, or SRC. Cardiac disease occurs in both the limited and the diffuse subtypes but is more common in the patients with rapidly progressive diffuse skin disease, in those with anti-U3 RNP antibodies, and in association with peripheral muscle disease.[28] Heart disease is reported to be a cause of death in 20–30% of patients and is an independent risk factor for death. Optimal screening for heart disease is not well defined, but periodic clinical assessment coupled with echocardiography and blood levels of BNP is recommended. Modern techniques for detecting occult heart disease provide opportunities for early intervention and should be used for thorough evaluation of the patient when cardiac involvement is suspected. The ideal therapy is defined by the underlying situation (*e.g.,* vasodilator therapy to improve heart perfusion or antiarrhythmia therapy for a complex arrhythmia). A disease-modifying drug that specifically targets scleroderma heart disease has not yet been discovered.

Renal Involvement

Renal disease is an important clinical problem with both direct involvement and kidney compromise secondary to other organ dysfunction. The most dreaded renal complication is an SRC, which is defined as the new onset of accelerated arterial hypertension and/or rapidly progressive oliguric renal failure during the course of the disease. It is recognized that other causes of renal disease, including interstitial nephritis, glomerulonephritis, and antineutrophil cytoplasmic antibody (ANCA)–associated vasculitis, can occur. SRC develops in 10–15% of patients with dcSSc, but in less than 5% of patients with lcSSc. It is generally a relatively early (<3–4 years) complication of the disease and is more likely to occur in males.[29] The presence of rapid progression of diffuse skin disease; tendon friction rubs; cardiac disease with clinically silent pericardial effusion, arrhythmias, or congestive heart failure; new unexplained anemia; or the use of corticosteroids are all associated with an increased risk of SRC. The autoantibody status is also helpful in identifying patients at risk for SRC. In approximately 30% of patients with anti-RNA polymerase III, seen almost exclusively in dcSSc, SRC develops; however, SRC occurs in only 10% of patients with antitopoisomerase antibodies and only rarely in patients with anticentromere antibody. SRC characteristically presents as sudden onset of malignant hypertension with rapidly progressive oliguric renal failure. Proteinuria, microhematuria, and evidence of a microangiopathic hemolytic process with thrombocytopenia are present. The exact mechanism triggering acute SRC is not known, but an autoimmune insult to the underlying SSc microvascular disease is thought to initiate a self-perpetuating process, which leads to glomerular hypoperfusion and a high renin state.

Patients with SRC who present with a creatinine >3 mg/dL have a poor outcome, with increased likelihood of permanent hemodialysis, need for renal transplantation, and/or high mortality. Prompt and aggressive intervention with angiotensin-converting enzyme (ACE) inhibitors to achieve blood pressure control before evidence of renal dysfunction is seen significantly improves the prognosis. Introduction of ACE inhibitors dramatically improved survival from a 1-year survival rate of 10% to a 5-year survival rate of 60%. All patients with SRC should be treated with an ACE inhibitor, but other vasodilators, including calcium channel blockers, endothelin receptor inhibitors, and prostaglandins, can be helpful. About 45% of patients whose blood pressure is normalized do not require dialysis. Reasonable kidney function can resume months after renal failure and the need for dialysis support. Despite ideal therapy, approximately 40% of patients still need long-term kidney support or have a poor survival. Overall survival on chronic dialysis is worse in patients with SSc compared with those without SSc, but renal transplantation is still an option with a long-term outcome similar to that of other causes of renal failure.

CLINICAL PEARLS

Recommended Approach to Systemic Sclerosis

- Determine scleroderma skin score to define clinical subtype
- Frequent clinical reassessment to define disease activity
- Careful clinical history to evaluate for gastrointestinal dysmotility
- Monitor for new onset hypertension: prevent renal crisis
- Pulmonary function test: early detection of interstitial lung disease
- Doppler echocardiography: screen for pulmonary arterial hypertension
- Obtain serology profile to help predict clinical outcome

Musculoskeletal Complications

Musculoskeletal impairment is a leading cause of disability and poor quality of life among patients with SSc. It is seen in the majority of patients, occurs early in the disease, and commonly has a dominant effect on hand function. There is a variety of associated clinical symptoms, including arthralgias, myalgias, stiffness, pain and loss of function as a result of joint contractures, inflammatory arthritis, myopathy, nerve entrapment (*e.g.,* carpal tunnel syndrome), fibrosis of tendons, and deconditioning from disuse. Loss of joint function is less likely to be secondary to synovitis but is usually caused by fibrosis of the overlying skin, the supporting joint structures, and the joint capsule itself. Flexion contractures associated with "friction rubs" most prominently felt over the ankles, knees, shoulders, and wrist are typical of advanced dcSSc. Muscle weakness is a common comorbid condition caused by deconditioning, muscle disuse, or malnutrition associated with weight loss. A scleroderma-related myopathy defined by weakness with elevated serum creatine kinase (CK) level; myopathic features on electromyography (EMG); and/or muscle biopsy is reported to occur in about 20% of patients. Most patients with skeletal myopathy have proximal muscle weakness on examination, and about 40% have one of the scleroderma/myopathy–associated autoantibodies (*i.e.,* anti-PM/Scl-75, anti-PM/Scl-100, or anti-CD1). The histopathological findings are highly variable, with fibrosis alone seen only in a small subset. When inflammation and/or necrosis are seen, immunosuppressive therapy has the potential to control the muscle disease.

Emotional Aspects

SSc impacts every aspect of a patient's life and is a disfiguring disease; therefore, it is not surprising that significant depression is common among patients with SSc.[30] Patients are frequently distressed by their appearance and avoid social interaction. Chronic pain and lack of social support are major causes of depression; therefore, family education, emotional support, pain control, and management of depressive symptoms are most important.

Treatment

THERAPEUTIC PRINCIPLES

Treatment of Systemic Sclerosis

Organ-Specific Therapy

- Vasodilator therapy for Raynaud phenomenon
- Proton pump inhibitor for gastroesophageal reflux
- Angiotensin-converting enzyme (ACE) inhibitor for scleroderma renal crisis
- Vasodilator therapy for pulmonary arterial hypertension
- Antiinflammatory therapy for arthritis
- Immunosuppressive therapy for interstitial lung disease

Disease-Modifying Agents

- No ideal treatment available; immunomodulatory, antiinflammatory, vasoactive, and antifibrotic agents are used and/or under study.

To date, no drug or intervention has been proven to be effective and safe for modifying the overall disease course. Most successful SSc treatments are directed to the specific organ(s) involved. Therapy is focused on an organ-specific disease process (*e.g.*, ACE inhibitor in SRC) or to aid a failing organ (*e.g.*, proton pump inhibitors for GERD). Therefore it is important to carefully characterize the patient's clinical phenotype, specific organ involvement, and level of disease activity before deciding on therapy. For example, a patient who has lcSSc but no evidence of visceral organ involvement is likely to have a benign course that requires only symptomatic therapy (*e.g.*, treatment of GERD and RP); use of systemic disease-modifying drugs would not be indicated. It is also important to distinguish disease activity from severity or advanced cumulative organ damage. For example, a patient with late-stage dcSSc with irreversible organ damage is unlikely to benefit from aggressive immunosuppressive or antiinflammatory therapy. Intervention with currently available drugs should be initiated early, ideally during the edematous active inflammatory phase of the disease. It is during the early stage of the disease that immunosuppression and antiinflammatory and antifibrotic agents have the greatest potential to control disease progression. Clinical experience teaches that once the edematous phase of the disease shifts to the more indolent fibrotic phase, current treatments are less likely to control the progression of disease and tissue damage. The primary outcome measure in clinical trials that focused on aggressive skin disease has been the modified Rodnan skin score; in current studies, however, composite scores that tally the burden of disease (activity and severity) as well as patient reported distress are included. Few well-designed controlled studies have been carried out for the treatment of early active disease; most reports are of anecdotal, uncontrolled experiences, complicated by investigator bias and the highly variable natural course of SSc.

Evidence-based management of SSc is helpful[31] while innovative therapies are being further studied in clinical trials.[32] Scleroderma is a multifaceted disease that includes active autoimmunity, vascular injury, and fibrosis and/or epithelial damage. Thus the major targets for therapy include regulating the immune system, controlling tissue fibrosis, and protection or prevention of progressive vascular disease. The majority of therapeutic interventions for scleroderma attempt to modify the disease process by immunomodulation in *early active* dcSSc because of the evidence that the SSc is an autoimmune disease initiated and/or propagated by an immune process. Drugs currently used for dcSSc include methotrexate, cyclophosphamide, and mycophenolate mofetil. Targeting specific immune cells is now a novel strategy that includes eliminating B cells with rituximab and suppressing activation of T cells with abatacept. In cases of severe life-threatening disease in highly selected cases not responding to low-dose immunosuppression, immunoablative therapy with high-dose cyclophosphamide alone or with radiation, followed by stem cell rescue, may be considered. Studies to evaluate the safety and efficacy of stem cell therapies are in progress.[33] Inhibiting proinflammatory and/or profibrotic cytokines or signaling pathways that may block tissue injury and fibrosis are being explored. These include antichemokines (CCR2) or inhibitors of CTGFs, recombinant humanized mAb to block TGF-β1 activity, modulation of interleukin-6 (IL-6) signaling, antagonists of the bioactive phospholipid LPA, and small-molecule inhibitors, such as tyrosine kinase inhibitors or pirfenidone. It is recognized that vascular disease also plays a fundamental role in the morbidity and mortality seen in scleroderma. Novel drugs targeting vascular disease, including endothelin receptor antagonists, agents that enhance nitric oxide production, prostacyclin analogs, antiplatelet agents, and statins, are currently being used.

ON THE HORIZON

Novel Therapeutic Approaches to Systemic Sclerosis (SSc) Management in Clinical Trials

- Targeted therapies to block transforming growth factor-β (TGF-β)
- Development of antifibrotic agents, including new approaches to blocking of fibrogenic pathways
- Use of biological agents, including anticytokines, antichemokines, and growth factor inhibitors

Other Fibrosing Diseases

Several disorders can cause skin fibrosis and mimic SSc.[34] SSc can be distinguished from these SSc-like fibrosing conditions by its characteristic clinical, pathological, and laboratory features. The most distinguishing clinical features of SSc include the presence of RP, nailfold capillary changes, the characteristic distribution, and characteristic skin changes. The skin pattern in SSc is noted for severe finger, hand, and distal limb involvement and sparing of the skin of the back of the trunk. SSc-specific serum autoantibodies also help define the presence of the SSc disease process.

Conditions that mimic SSC include localized forms of scleroderma (see Table 55.5), eosinophilic fasciitis, nephrogenic systemic fibrosis, scleromyxedema (papular mucinosis), scleredema, graft-versus-host disease (GvHD), toxic oil syndrome, and eosinophilia–myalgia syndrome (Table 55.6). Eosinophilic fasciitis presents with induration of subcutaneous tissues, particularly in the proximal limbs, without RP or systemic organ involvement. A peripheral eosinophilia with radiological and

TABLE 55.6 Differential Diagnosis of Systemic Sclerosis (SSc) and Scleroderma

Disorders Characterized by Similar Presentations

- Systemic lupus erythematosus
- Sjögren syndrome
- Rheumatoid arthritis
- Polymyositis/dermatomyositis
- Primary Raynaud phenomenon

Disorders Characterized by Similar Visceral Features

- Primary pulmonary hypertension
- Primary biliary cirrhosis
- Idiopathic intestinal hypomotility
- Idiopathic pulmonary fibrosis
- Malignant hypertension

Disorders Characterized by Skin Thickening

- Scleromyxedema
- Scleredema (of Buschke), diabetic scleredema
- Nephrogenic fibrosing dermatopathy
- Eosinophilic fasciitis/diffuse fasciitis with eosinophilia
- Eosinophilia–myalgia syndrome
- Generalized morphea
- Chronic graft versus host disease
- POEMS syndrome (polyneuropathy, organomegaly, endocrinopathy, monoclonal protein, skin changes)
- Amyloidosis
- Carcinoid syndrome
- Pentazocine-induced scleroderma
- Diabetic digital sclerosis
- Vinyl chloride disease
- Toxic oil syndrome
- Bleomycin exposure
- Werner syndrome
- Phenylketonuria
- Porphyria cutanea tarda
- Vibration white finger syndrome
- Chronic reflex sympathetic dystrophy

histological evidence of an inflammatory process in the fasciae is present. Nephrogenic systemic fibrosis is described in patients with end-stage renal disease who were exposed to gadolinium-based contrast agents.[35] Scleromyxedema, a condition of mucinous deposition in skin, is associated with the presence of monoclonal gammopathy that is characterized by a diffuse, papular skin eruption. Scleredema occurs with deposition of collagen and mucin in the dermis and primarily involves the upper back, neck, and face. It occurs in association with diabetes, a monoclonal gammopathy, or postinfection pharyngitis, particularly streptococcal pharyngitis. GvHD occurs after bone marrow or stem cell transplantation and is characterized by patchy areas of inflammatory and thickened skin with liver and GI involvement. Toxin exposure was linked to the toxic oil syndrome (adulterated rapeseed oil) and eosinophilia–myalgia syndrome (related to the presence of an impurity in the production process for L-tryptophan, a sleep aid). These syndromes have largely subsided by banning the agents involved, but sporadic cases of eosinophilia–myalgia syndrome continue to occur.[36]

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Maricq HR, Weinrich MC, Keil JE, et al. Prevalence of scleroderma spectrum disorders in the general population of South Carolina. Arthritis Rheum 1989;32:998–1006.
2. Mayes MD, Lacey JV Jr, Beebe-Dimmer J, et al. Prevalence, incidence, survival, and disease characteristics of systemic sclerosis in a large US population. Arthritis Rheum 2003;48:2246–55.
3. Bossini-Castillo L, López-Isac E, Martín J. Immunogenetics of systemic sclerosis: defining heritability, functional variants and shared-autoimmunity pathways. J Autoimmun 2015;64:53–65.
4. Hummers LK. The importance of recognizing scleroderma-type disorders in clinical practice. Nat Clin Pract Rheumatol 2008;4:638–40.
5. Janowsky EC, Kupper LL, Hulka BS. Meta-analyses of the relation between silicone breast implants and the risk of connective-tissue diseases. N Engl J Med 2000;342:781–90.
6. Bhattacharyya S, Wei J, Varga J. Understanding fibrosis in systemic sclerosis: shifting paradigms, emerging opportunities. Nat Rev Rheumatol 2011;8:42–54.
7. Kuwana M, Okazaki Y, Yasuoka H, et al. Defective vasculogenesis in systemic sclerosis. Lancet 2004;364:603–10.
8. Sato S, Fujimoto M, Hasegawa M, et al. Altered blood B lymphocyte homeostasis in systemic sclerosis: expanded naive B cells and diminished but activated memory B cells. Arthritis Rheum 2004;50:1918–27.
9. Harris ML, Rosen A. Autoimmunity in scleroderma: the origin, pathogenetic role, and clinical significance of autoantibodies. Curr Opin Rheumatol 2003;15:778–84.
10. Mori Y, Chen SJ, Varga J. Expression and regulation of intracellular SMAD signaling in scleroderma skin fibroblasts. Arthritis Rheum 2003;48:1964–78.
11. Domsic RT, Rodriguez-Reyna T, Lucas M, et al. Skin thickness progression rate: a predictor of mortality and early internal organ involvement in diffuse scleroderma. Ann Rheum Dis 2011;70:104–9.
12. Mendoza F, Derk CT. Systemic sclerosis mortality in the United States: 1999-2002 implications for patient care. J Clin Rheumatol 2007;13:187–92.
13. LeRoy EC, Medsger TA Jr. Criteria for the classification of early systemic sclerosis. J Rheumatol 2001;28:1573–6.
14. Chung L, Domsic RT, Lingala B, et al. Survival and predictors of mortality in systemic sclerosis-associated pulmonary arterial hypertension: outcomes from the pulmonary hypertension assessment and recognition of outcomes in scleroderma registry. Arthritis Care Res 2014;66:489–95.
15. Flavahan NA. A vascular mechanistic approach to understanding Raynaud phenomenon. Nature 2015;11:146–58.
16. Cappelli L, Wigley FM. Management of Raynaud phenomenon and digital ulcers in scleroderma. Rheum Dis Clin North Am 2015;41: 419–38.
17. Korn JH, Mayes M, Matucci Cerinic M, et al. Digital ulcers in systemic sclerosis: prevention by treatment with bosentan, an oral endothelin receptor antagonist. Arthritis Rheum 2004;50:3985–93.
18. Roberts CG, Hummers LK, Ravich WJ, et al. A case-controlled study of the pathology of oesophageal disease in systemic sclerosis (scleroderma). Gut 2006;55:1697–703.
19. Tyndall AJ, Bannert B, Vonk M, et al. Causes and risk factors for death in systemic sclerosis: a study from the EULAR Scleroderma Trials and Research (EUSTAR) database. Ann Rheum Dis 2010;69:1809–15.
20. Goh NS, Desai SR, Veeraraghavan S, et al. Interstitial lung disease in systemic sclerosis: a simple staging system. Am J Respir Crit Care Med 2008;177:1248–54.
21. Tashkin DP, Elashoff R, Clements PJ, et al. Scleroderma Lung Study Research Group. Cyclophosphamide versus placebo in scleroderma lung disease. N Engl J Med 2006;354:2655–66.
22. Clements PJ, Tashkin D, Roth M, et al. The Scleroderma Lung Study II (SLS II) shows that both oral cyclophosphamide (CYC) and mycophenolate mofetil (MMF) are efficacious in treating progressive interstitial lung disease (ILD) in patients with systemic sclerosis (SSc). Arthritis Rheumatol 2015;67(Suppl. 10). http://acrabstracts.org/abstract/

the-scleroderma-lung-study-ii-sls-ii-shows-that-bothoral-cyclophosphamide-cyc-and-mycophenolate-mofitil-mmf-are-efficacious-in-treatingprogressive-interstitial-lung-disease-ild-in-patients-w/.
23. Silver KC, Silver RM. Management of systemic-sclerosis-associated interstitial lung disease. Rheum Dis Clin North Am 2015;41:439–57.
24. Lota HK, Wells AU. The evolving pharmacotherapy of pulmonary fibrosis. Expert Opin Pharmacother 2013;14:79–89.
25. Le Pavec JL, Humbert M, Mouthon L, et al. Systemic sclerosis-associated pulmonary arterial hypertension. Am J Respir Crit Care Med 2010;181:1285–93.
26. Mathai SC, Hassoun PM. Therapy for pulmonary arterial hypertension associated with systemic sclerosis. Curr Opin Rheumatol 2009;21:642–8.
27. Humbert M, Yaici A, de Groote P, et al. Screening for pulmonary arterial hypertension in patients with systemic sclerosis: clinical characteristics at diagnosis and long-term survival. Arthritis Rheum 2011;63:3522–30.
28. Follansbee WP, Zerbe TR, Medsger TA Jr. Cardiac and skeletal muscle disease in systemic sclerosis (scleroderma): a high risk association. Am Heart J 1993;125:194–203.
29. Denton CP, Lapadula G, Mouthon L, et al. Renal complications and scleroderma renal crisis. Rheumatology (Oxford) 2009;48(Suppl.iii):32–5.
30. Haythornthwaite JA, Heinberg LJ, McGuire L. Psychologic factors in scleroderma. Rheum Dis Clin North Am 2003;29:427–39.
31. Khanna D, Denton CP. Evidence-based management of rapidly progressing systemic sclerosis. Best Pract Res Clin Rheum 2010;24:387–400.
32. Denton CP. Advances in pathogenesis and treatment of systemic sclerosis. Clin Med (Lond) 2016;16:55–60.
33. Van Laar JM, Naraghi K, Tyndall A. Haematopoietic stem cell transplantation for poor-prognosis systemic sclerosis. Rheumatology (Oxford) 2015;54:2126–33.
34. Boin F, Hummers LK. Scleroderma-like fibrosing disorders. Rheum Dis Clin North Am 2008;34:199–220.
35. Todd DJ, Kay J. Gadolinium-induced fibrosis. Annu Rev Med 2016;67:273–91.
36. Allen J, Peterson A, Sufit R, et al. Post-epidemic eosinophilia myalgia syndrome associated with L-tryptophan. Arthritis Rheum 2011;63:3633–9.

MULTIPLE-CHOICE QUESTIONS

1. Which of the following complications is MOST frequent in systemic sclerosis (SSc)?
   A. Glomerulonephritis
   B. Scleritis
   C. Renal crisis
   D. Tendon rupture
   E. Pulmonary arterial hypertension

2. Characteristic thoracic involvement in SSc include all of the following EXCEPT:
   A. Pulmonary fibrosis
   B. Apical nodules
   C. Peripheral reticulations
   D. Enlarged pulmonary artery
   E. Patulous esophagus

3. Which of the scleroderma related autoantibodies are associated with the highest risk for developing a scleroderma renal crisis (SRC)?
   A. Anticentromere
   B. Antitopoisomerase
   C. Anti-RNA polymerase III
   D. Anti-U1 ribonucleoprotein (RNP)
   E. Anti-U3 RNP
   F. Anti-U11/U12 RNP

4. What is the major cause of mortality in scleroderma in the modern era?
   A. Scleroderma renal crisis
   B. Severe skin disease
   C. Scleroderma lung disease
   D. Gastrointestinal failure
   E. Complications of medications

56

Inflammatory Muscle Diseases

Arash H. Lahouti, Lisa Christopher-Stine

The idiopathic inflammatory myopathies (IIMs) are a group of rare systemic diseases. They consist of polymyositis (PM), dermatomyositis (DM), inclusion body myositis (IBM), and more recently described immune-mediated necrotizing myopathy (IMNM). Traditionally, they are believed to be autoimmune diseases, although recent studies suggest that there is a close resemblance between IBM and other neurodegenerative diseases. Similar to patients with other autoimmune diseases, those with IIM often have serum autoantibodies. Some of these antibodies are specific for myositis and are not seen in other rheumatic disorders. Our knowledge of the myositis-specific antibodies (MSAs) is incomplete. There has been considerable effort to better characterize antibodies associated with myositis and to discover new antibodies. Many MSAs are closely linked to a unique clinical phenotype. Thus they are useful in classifying the IIM. They can forewarn of particular extramuscular manifestations and guide appropriate treatment strategies.

IIM patients present with muscle weakness and elevated muscle enzymes. In patients with DM, skin manifestations may be the initial presentation. The differential diagnosis of muscle disease is broad and includes drug-induced myopathies, neuromuscular disorders, muscular dystrophies, and metabolic and endocrine myopathies. To add to the complexity, some DM patients do not develop muscle weakness, referred to as clinically amyopathic DM. The patient's age and gender, the pattern of the weakness, the severity of manifestations, and associated symptoms are commonly helpful in leading to the correct diagnosis. For example, IBM is distinguished from other IIMs by a characteristic involvement of the finger flexor and knee extensor muscles, which is often asymmetrical. IBM is more common in older men, whereas PM and DM are commonly seen in young to middle-aged women and children. DM is often associated with characteristic skin findings, which are not a feature of PM and IBM. However, in many patients differentiation between the IIMs cannot be made on clinical grounds, and further diagnostic testing is required. For example, both PM and IMNM present in the same fashion with a predominantly symmetrical proximal muscle weakness and elevated muscle enzymes. They can be distinguished only by pathological examination. On muscle biopsy, IMNM is associated with necrosis and regeneration of muscle fibers and a characteristic sparse inflammatory infiltrate. In contrast, PM is associated with the presence of cytotoxic inflammatory cells surrounding and invading muscle fibers. Electromyography is a valuable tool for differentiating between weakness originating from muscle rather than peripheral nerves. Magnetic resonance imaging (MRI) can be extremely helpful in identifying inflammatory changes in patients with subtle clinical muscle involvement. Moreover, MRI can help to better estimate the burden of disease and differentiate acute changes (edema) from chronic changes (atrophy). The role of new imaging modalities such as whole-body MRI needs to be further investigated.

Interstitial lung disease, gastrointestinal involvement, and arthritis are among the most common extramuscular manifestations of the IIM. Interstitial lung disease commonly occurs as part of the antisynthetase syndrome in a subset of patients who have antisynthetase autoantibodies. Gastrointestinal manifestations include dysphagia and aspiration pneumonia. Dysphagia is particularly common and can be seen in all forms of the IIMs. Certain forms of IIM, particularly DM, can be a paraneoplastic phenomenon. The most common cancers associated with myositis include gynecological (ovarian), pulmonary, gastrointestinal (pancreatic, stomach, and colorectal), and non-Hodgkin lymphoma. Also, the IIMs may be associated with other autoimmune diseases such as systemic lupus erythematosus (SLE), Sjögren syndrome, and systemic sclerosis.

Muscle biopsy is critical for the diagnosis of the IIMs. The presence of perifascicular atrophy is strongly suggestive of DM, whereas the finding of rimmed vacuoles in the appropriate context suggests IBM. Muscle biopsy also differentiates between the IIMs and other forms of myopathy such as drug-induced myopathies and muscular dystrophies.

Treatment of the IIMs is based largely on experience. Corticosteroids remain the mainstay of therapy. Steroid-sparing agents such as azathioprine, methotrexate, mycophenolate mofetil, and hydroxychloroquine are frequently initiated at presentation while a steroid taper is attempted. In patients with refractory disease, rituximab, intravenous immunoglobulin, and biological antirheumatic medications may be tried. IBM is the most resistant subset of the IIMs.

DM, PM, IBM, and IMNM constitute the largest subgroup of the acquired myopathies. They are a heterogeneous group and are rare among immunological illnesses. However, the IIMs share many clinical features and laboratory abnormalities, including related autoantibodies, and are closely related to major autoimmune diseases. Because their presentation and major clinical manifestations are weakness and/or rash, the differential diagnosis includes many more common diseases familiar to neurologists and dermatologists.

CLINICAL FEATURES

The clinical hallmark of PM, DM, and IMNM is the gradual onset of symmetrical proximal muscle weakness over weeks to months. In some cases, myalgia may be the presenting or most bothersome symptom, but more often the patient is evaluated for the physical limitations imposed by weakness: difficulty arising from a low chair or bed, or combing and brushing hair. Rash is the first feature in a considerable proportion of patients who have DM, but muscle weakness usually follows within a few months. A subset of patients with DM, clinically amyopathic dermatomyositis (C-ADM), may present only with rash in the absence of muscle weakness throughout the course of their illness. These patients are also at risk for pulmonary involvement, as are those with classic DM. The prevalence of interstitial pneumonitis in C-ADM can approach 5–10%, compared with 40% of patients with classic DM.[1] Arthritis, Raynaud phenomenon, fever, or lung disease presenting as cough or dyspnea may dominate the clinical picture. Cardiac and gastrointestinal symptoms, other than dysphagia in severe cases, are rarely early manifestations. Renal and central nervous system (CNS) involvement are almost never a part of the IIMs.

KEY CONCEPTS

Definition and Incidence of Idiopathic Inflammatory Myopathies (IIMs)

- Polymyositis (PM), dermatomyositis (DM), and related inflammatory muscle diseases are called IIM.
- Indistinguishable muscle inflammation may accompany other autoimmune connective tissue diseases or limb-girdle muscular dystrophies.
- The annual incidence in the US is 5–10 cases per million. DM and PM are more common in women than in men in all age groups; inclusion body myositis (IBM) is more common in men.

Some of the rashes of DM are virtually pathognomonic; others are not disease-specific (Fig. 56.1). The heliotrope rash, a violaceous discoloration of the eyelids, is sometimes no more than a line along the margin of the upper lid, but it may also affect both upper and lower lids completely and can be associated with edema mimicking thyroid disease. A reddish, sometimes raised and/or scaly, eruption over the metacarpophalangeal joints is known as Gottron papules. In some cases, the metatarsophalangeal joints, elbows, knees, and malleoli show a similar rash. Both heliotrope and Gottron rashes can occur rarely in cases of frank SLE without muscle involvement. Other common rashes include a flat, red blanching eruption of the upper chest (often in a V distribution), the upper back (where a shawl would touch), and sometimes the extensor surfaces of the upper arms and thighs. Another rash that mimics the malar rash of lupus on the face may be present; however, in contrast to lupus, it does not spare the nasolabial folds. Although found on sun-exposed parts of the body, these rashes are often not photosensitive in nature. As in other connective tissue diseases, nailfold capillary dilatation, infarcts, and cuticular overgrowth occur. Mechanic's hands, a roughening and cracking of the radial sides of the fingers and the palm, resembling a condition found in people who labor with their hands, is characteristic of a subset of myositis patients with the "antisynthetase syndrome" and can also be seen in patients with PM-Scl and U1-RNP autoantibodies.

CLASSIFICATION

In the past 40 years, several investigators have proposed diagnostic classification criteria for IIM. The criteria proposed by Bohan and Peter four decades ago remain the most familiar and accepted definitions of PM and DM (Table 56.1).[2,3] They combine clinical, laboratory, electrodiagnostic, and pathological features. These criteria currently serve as the gold standard for clinical diagnosis and for inclusion in clinical trials. However, they are limited by their poor specificity in distinguishing PM from other entities, including late-onset muscular dystrophies. The resultant misclassification limits the homogeneity of the patients included in previous observational and interventional studies. Additionally, the Bohan and Peter criteria completely omit the diagnosis of IBM, the most frequent type of IIM in patients over 50 years of age.

In 2004, more comprehensive criteria of IIM were proposed and approved by a group of international experts (Table 56.2). Unlike the previous criteria, these new criteria offer the advantage of classifying two rare forms of autoimmune myositis, *i.e.,* IMNM

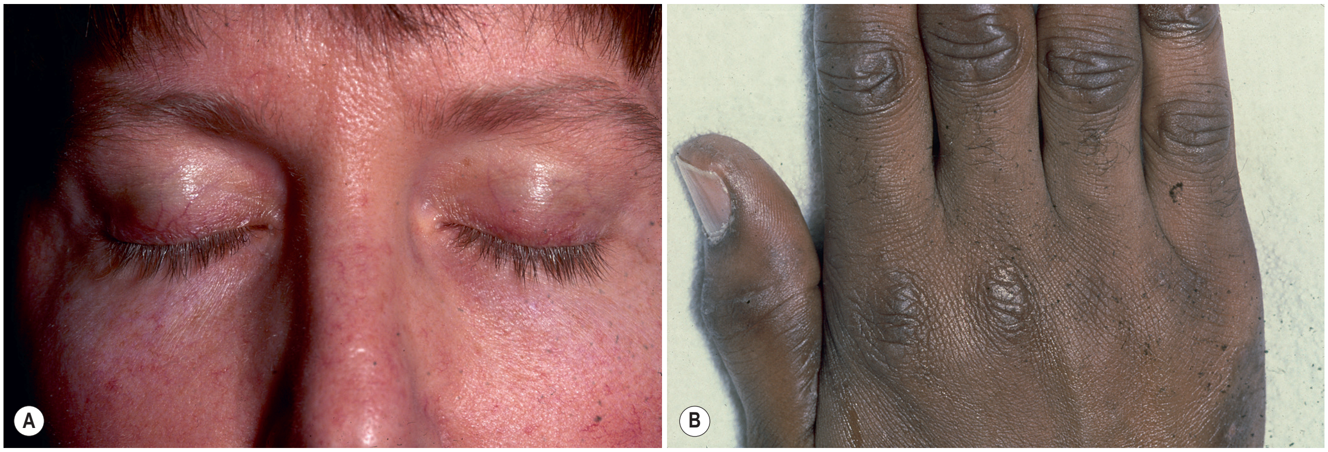


FIG 56.1 Dermatomyositis Rash. (A) In addition to the heliotrope rash on the eyelids of this patient with dermatomyositis, there is a flat, red rash on the nose and cheeks. (B) A raised, shiny-red rash—Gottron papules—is apparent on the interphalangeal and the second and third metacarpophalangeal joints of this man with dermatomyositis.

TABLE 56.1 Idiopathic Inflammatory Myopathy: Diagnostic Criteria

Bohan and Peter Criteria

1. Symmetrical proximal muscle weakness
2. Skeletal muscle enzyme elevation
3. Abnormal EMG[a]
4. Muscle biopsy abnormalities
5. Typical skin rash of DM[b]

[a] Polyphasic, short, small motor-unit potentials; fibrillation, positive sharp waves, increased insertional irritability; bizarre, high-frequency, repetitive discharges.

[b] Gottron sign, heliotrope rash.

Possible PM = any two of the first four criteria; possible DM = criterion 5 (rash) + any two criteria. Probable PM = any three of the first four criteria; probable DM = criterion 5 (rash) + any three criteria. Definite PM = all four of the first four criteria; definite DM = criterion 5 (rash) + all four other criteria.

DM, dermatomyositis; *EMG,* electromyography; *PM,* polymyositis.

and DM *sine* dermatitis, as separate categories (Table 56.2).[4] Separate classification criteria systems for IBM have been devised. According to most of the previous IBM criteria, characteristic muscle biopsy changes were necessary for a definite diagnosis of IBM. However, because selective involvement of finger flexors and knee extensors is almost exclusive to IBM among other forms of myopathy, the newly proposed *ENMC IBM Research Diagnostic Criteria 2011* entails a separate category for clinically defined IBM (Table 56.3).[5]

It has been useful for some purposes to divide cases into groups: PM, DM, juvenile myositis, myositis associated with another connective tissue disease (usually systemic sclerosis, SLE, or Sjögren syndrome), cancer-associated myositis (usually cases in which the diagnoses are made within 6–12 months of one another), IBM, and a miscellaneous group that includes such rare entities as eosinophilic myositis (Table 56.4). This classification has allowed recognition of unique clinical and pathogenetic features and response to therapy. In the case of cancer-associated myositis, a more rational approach to workup based on recognition of groups at risk is now possible.

KEY CONCEPTS

Characteristic Hallmarks of Inflammatory Myositis

- The clinical hallmark is proximal limb and neck weakness, rarely associated with muscle pain.
- The laboratory hallmarks are elevated serum levels of creatine kinase (CK), aldolase, lactic dehydrogenase, and the transaminases; a characteristic pattern ("irritable myopathy") is seen on electromyography (EMG). Elevated serum levels of autoantibodies are common.
- The pathological hallmarks are focal muscle necrosis, degeneration, regeneration, and inflammation.

Several autoantibodies, called "myositis-specific autoantibodies," are unique to myositis. These have allowed a useful alternative classification (Table 56.5). For example, patients with antibodies to the aminoacyl-tRNA (transfer RNA) synthetases, of which Jo-1 is the best known, have a characteristic syndrome called antisynthetase syndrome, which usually includes interstitial lung disease, nondeforming inflammatory arthritis, fevers, mechanic's hands, and Raynaud phenomenon, in addition to myositis. Those with antibodies to the signal recognition particle (anti-SRP) have an IMNM manifested by severe disease of abrupt onset, often in the autumn. According to some studies, cardiac involvement is less common and survival is better in patients with anti-SRP than has previously been reported.[6] Those with antibodies to the nuclear antigen Mi-2 almost always have the V and shawl rashes and cuticular overgrowth in addition to myositis. A recently discovered antibody, anti-3-hydroxy-3-methylglutaryl-coenzyme A reductase (anti-HMGCR), is tightly linked to IMNM.[7,8] Interestingly, the HMGCR is the pharmacological target of statin drugs, such as atorvastatin, and a subset of patients with anti-HMGCR antibodies develop this antibody in the context of previous statin exposure, although it is also seen in statin-naïve individuals, albeit to a lesser extent. Patients with anti-HMGCR antibodies present with profound proximal muscle weakness and a marked creatine kinase (CK) elevation. Further, compared with statin-exposed individuals, statin-naïve anti-HMGCR positive patients tend to be younger, have higher CK levels, and have a blunted response to immunosuppressive medications (Table 56.5).[2,8]

IBM is different from other inflammatory myopathies. Patients with IBM rarely improve in strength with immunosuppressive therapy. They tend to be older, and in contrast to patients with PM and DM, who are predominantly women, patients with IBM are more commonly men. They have gradual, painless, asymmetrical, progressive weakness and focal atrophy that develop over years, and they may complain of frequent falls. Two decades after onset, they are frequently wheelchair-bound. The forearms of these patients exhibit a scalloped appearance, attributed to muscle atrophy. Difficulty with swallowing can occur with any inflammatory myopathy and is frequently a major problem in patients with IBM. The CK and other skeletal muscle–associated serum enzymes are normal in about one-quarter of patients with IBM and only moderately elevated in the remainder. The electromyogram in IBM frequently demonstrates both myogenic and neurogenic features secondary to the effective denervation of some muscle cells by inflammation and necrosis.

Proposed criteria for the diagnosis of IBM rely on both pathological and clinical features. In the clinical setting of an inflammatory myopathy, the presence of the characteristic inclusions or rimmed vacuoles is diagnostic. Among the inflammatory myopathies, IBM is distinguished by substantial numbers of rimmed cytoplasmic vacuoles with tubulofilamentous material within myofibers. A variety of proteins have been found by immunohistochemistry in the muscle cells in IBM, including ubiquitin, β-amyloid precursor protein, and the transcription factor nuclear factor (NF)-κB. Many of the proteins accumulating in the IBM muscle are also involved in other neurodegenerative diseases. For example, aggregation of p62, an autophagy-related protein, has been shown in IBM muscle, in Lewy bodies in Parkinson disease, and in neurofibrillary tangles in Alzheimer disease, suggesting the possibility of a degenerative process in IBM. More recently, an autoantibody to cytosolic 5′-nucleotidase 1a (c5N1A) was discovered. Although this antibody has a good specificity to distinguish IBM against other forms of autoimmune myopathy, it is roughly positive in 60% of IBM patients.[9] In addition, anti-c5N1A can be detected in other connective tissue

TABLE 56.2 Proposed Diagnostic Criteria for Inclusion Body Myositis

The European Neuromuscular Centre (ENMC) criteria for the idiopathic inflammatory myopathies[4]

1. Clinical criteria
   a. Subacute onset
   b. Age >18 years (onset may be in childhood in DM and nonspecific myositis)
   c. Symmetrical proximal weakness
   d. Typical DM rash (including heliotrope, Gottron papules, Gottron sign, V sign, and shawl sign)
   e. Lack of features suggestive of IBM (asymmetry, finger flexor ≥ deltoid weakness, and knee extensors/ankle dorsiflexors ≥ hip flexors weakness), toxic myopathies, endocrine myopathies, amyloidosis, family history of muscular dystrophy or proximal motor neuropathies (*e.g.,* SMA)
2. Elevated serum creatine kinase level
3. Other laboratory criteria
   a. Abnormal electromyography
      Inclusion criteria
      I. Increased insertional and spontaneous activity in the form of fibrillation potentials, positive sharp waves, or complex repetitive discharges
      II. Morphometric analysis reveals the presence of short duration, small amplitude, polyphasic motor-unit action potentials (MUAPs)
      Exclusion criteria
      I. Myotonic discharges that would suggest proximal myotonic dystrophy or other channelopathy
      II. Morphometric analysis reveals predominantly long duration, large-amplitude MUAPs
      III. Decreased recruitment pattern of MUAPs
   b. MRI: diffuse or patchy increased signal (edema) within muscle tissue on Short T1 Inversion Recovery (STIR) images
   c. Myositis-specific antibodies detected in serum
4. Abnormal muscle biopsy
   a. Endomysial inflammatory cell infiltrate surrounding and invading nonnecrotic muscle fibers
   b. Endomysial CD8 T cells surrounding, but not definitely invading, nonnecrotic muscle fibers; or ubiquitous MHC-1 expression
   c. Perifascicular atrophy
   d. Membrane attack complex (MAC) depositions on small blood vessels, or reduced capillary density, or tubuloreticular inclusions in endothelial cells on EM, or MHC-1 expression of perifascicular fibers
   e. Perivascular, perimysial inflammatory cell infiltrate
   f. Scattered endomysial CD8 T-cell infiltrate that does not clearly surround or invade muscle fibers
   g. Many necrotic muscle fibers as the predominant abnormal histological feature. Inflammatory cells are sparse or only slight perivascular; perimysial infiltrate is not evident
   h. Rimmed vacuoles, ragged red fibers, cytochrome oxidase–negative fibers that would suggest IBM
   i. MAC deposition on the sarcolemma of nonnecrotic fibers and other indications of muscular dystrophies with immunopathology

Polymyositis–*Definite polymyositis:* 1. All clinical criteria with the exception of rash, 2. Elevated serum creatine kinase (CK), 3. Muscle biopsy criteria include a and exclude c, d, h, i; *Probable polymyositis:* 1. All clinical criteria with the exception of rash, 2. Elevated serum CK, 3. Other laboratory criteria (1 of 3), 4. Muscle biopsy criteria include b and exclude c, d, g, h, i. **Dermatomyositis–***Definite dermatomyositis:* 1. All clinical criteria, 2. Muscle biopsy criteria include c; *Probable dermatomyositis:* 1. All clinical criteria, 2. Muscle biopsy criteria include d or e, or elevated serum CK, or other laboratory criteria (1 of 3). **Amyopathic dermatomyositis–**1. Rash typical of DM: heliotrope, periorbital edema, Gottron papules/sign, V sign, shawl sign, holster sign, 2. Skin biopsy demonstrates a reduced capillary density, deposition of MAC on small blood vessels along the dermal-epidermal junction, and variable keratinocyte decoration for MAC, 3. No objective weakness, 4. Normal serum CK, 5. Normal electromyogram, 6. Muscle biopsy, if done, does not reveal features compatible with definite or probable DM. **Possible dermatomyositis sine dermatitis–**1. All clinical criteria with the exception of rash, 2. Elevated serum CK, 3. Other laboratory criteria (1 of 3), 4. Muscle biopsy criteria include c or d. **Nonspecific myositis–**1. All clinical criteria with the exception of rash, 2. Elevated serum CK, 3. Other laboratory criteria (1 of 3), 4. Muscle biopsy criteria include e or f and exclude all others. **Immune-mediated necrotizing myopathy–**1. All clinical criteria with the exception of rash, 2. Elevated serum CK, 3. Other laboratory criteria (1 of 3), 4. Muscle biopsy criteria include g and exclude all others.

TABLE 56.3 ENMC IBM Research Diagnostic Criteria 2011[5]

Clinical and Laboratory Features	Pathological Features
a. Duration >12 months	I. Endomysial inflammatory infiltrate
b. Age at onset >45 years	
c. CK no greater than 15 × ULN	II. Rimmed vacuoles
d1. Knee extension weakness > hip flexion weakness	III. Protein accumulation* or 15–18-nm filaments
d2. Finger flexion weakness > shoulder abduction weakness	IV. Upregulation of MHC class I

*Demonstration of amyloid or other protein accumulation by established methods (*e.g.*, for amyloid Congo red, crystal violet, thioflavin T/S, for other proteins p62, SMI-31, TDP-43). Current evidence favors p62 in terms of sensitivity and specificity, but the literature is limited, and further work is required.

Clinicopathologically defined IBM: clinical criteria include a–c and at least one of the d criteria plus the first three pathological features. **Clinically defined IBM**: All clinical criteria plus one or more pathological features. **Probable IBM**: clinical criteria include a–c and at least one of the d criteria plus one or more pathological features.

CK, creatine kinase; *IBM,* inclusion body myositis; *ULN,* upper limit of normal.

TABLE 56.4 Traditional Classification of Idiopathic Inflammatory Myopathies

Type I	Primary idiopathic polymyositis
Type II	Primary idiopathic dermatomyositis
Type III	Dermatomyositis or polymyositis associated with malignancy
Type IV	Childhood dermatomyositis or polymyositis
Type V	Myositis associated with another connective tissue disease
Type VI	Inclusion body myositis
Type VII	Miscellaneous: eosinophilic myositis, localized nodular myositis, etc.

TABLE 56.5 Clinical Features Associated With Myositis-Specific Autoantibodies

Autoantibodies	Association	Characteristic Clinical Features
Anti-Jo-1 and other antisynthetases	PM, DM, IMNM	Relatively acute onset of myositis, frequent interstitial lung disease, fever, Raynaud phenomenon, arthritis, mechanic's hands, moderate response to therapy, persistent disease. Patients sometimes meet criteria for SLE or RA, but muscle disease or lung disease dominate the clinical picture and prognosis.
Anti-SRP	IMNM	Very acute onset of myositis, often in autumn, severe weakness, no rash, palpitations, females predominate, poor response to therapy
Anti-Mi-2	DM	Relatively acute onset of myositis, classic dermatomyositis rashes with V sign and shawl sign, cuticular overgrowth, good response to therapy
Anti-HMGCR	IMNM	Necrotizing myopathy, may be preceded by statin therapy, very high CK levels, minimal muscle wasting
Anti-MDA5	DM, Clinically amyopathic DM	Clinically amyopathic dermatomyositis with rapidly progressive interstitial lung disease; characteristic skin findings include cutaneous ulcers and palmar papules
Anti-TIF1-γ (p155/140)	DM	Juvenile dermatomyositis and cancer-associated dermatomyositis
Anti-SAE	DM	Severe skin manifestations and dysphagia
Anti-NXP-2	DM	Associated with calcinosis and muscle contractures in children

DM, dermatomyositis; *IMNM,* immune-mediated necrotizing myopathy; *MDA5,* melanoma differentiation–associated protein 5; *NXP-2,* nuclear matrix protein; *PM,* polymyositis; *RA,* rheumatoid arthritis; *SAE,* small ubiquitin-like modifier activating enzyme; *SLE,* systemic lupus erythematosus; *SRP,* signal recognition particle; *TIF1-γ,* transcription intermediary factor 1-γ.

diseases (such as lupus, DM, and Sjögren syndrome) as well as in healthy subjects.[10]

IBM must be distinguished from other chronic myopathies. These include acquired myopathies, such as those caused by toxins, and genetically determined myopathies, such as some muscular dystrophies and the metabolic myopathies. There are several important differences between IBM and these other myopathies. The distinction, however, is not as well defined as might be expected. Although some of the familial forms of IBM have a distinctive clinical presentation, often early in life, there have been several families with the typical late onset and inflammatory picture of the presumed sporadic cases. Several genetic loci have been identified in familial IBM, so it will be important to assess any identified mutations in familial IBM–associated genes in sporadic cases.[11]

ETIOLOGY

Immunological Clues to Origin

The implications of myositis-specific autoantibodies that bind to and inhibit the function of native human enzymes involved in the formation of new proteins are tantalizing and probably significant. These autoantibodies appear to develop before symptomatic weakness or serum CK elevation, suggesting a close link to an initiating factor. Some of the clinical features, such as fever, arthritis, and lung disease, and the apparent seasonality in onset of disease in patients with anti-SRP, are reminiscent of some viral infections. In patients with autoantibodies against one of the aminoacyl-tRNA synthetases, such as Jo-1 (histidyl-tRNA synthetase), the possible connection to a viral inciting agent appeared compelling, as certain picornaviruses, which are closely related to viruses long suspected of causing myositis, such as coxsackie viruses, can mimic tRNA in acting as a substrate for an aminoacyl-tRNA synthetase.[12] Direct proof connecting picornaviruses to human myositis, however, has not been obtained.

In individuals infected with HIV or human T-cell lymphotropic virus (HTLV)-1, the development of IBM has been noted. However, it is not always the first manifestation in these cases. There is no evidence of viral replication within the muscles, but instead the chronic infection triggers an inflammatory response.[13]

Drugs and Toxins

A large number of environmental agents have been associated with myopathies. Drug-induced myopathy should be considered particularly in cases where no other cause has been identified. Sometimes the illness strikingly resembles the spontaneous disease. D-penicillamine, for example, can induce a variety of autoimmune phenomena, including an inflammatory myopathy that closely resembles PM. A number of drugs can also produce myopathies that can be clinically confused with IIM but are histologically distinct. This large group includes 3-hydroxy-3-methylglutaryl-coenzyme A (HMG-coA) reductase inhibitors, corticosteroids, colchicine, and zidovudine (AZT).

In corticosteroid-induced myopathy, type II muscle fiber atrophy is prominent on muscle biopsy, and weakness improves when the dose is lowered. Colchicine can cause myopathy and painful neuromyopathy. The CYP3A4 system metabolizes colchicine, and taking another drug metabolized by the same pathway can result in myopathy.[14] Muscle biopsy shows autophagic vacuoles that stain for acid phosphatase. Discontinuation of colchicine usually results in improvement. AZT produces a characteristic mitochondrial myopathy. The myopathy associated with HIV infection can be distinguished from that caused by AZT by muscle biopsy, as characteristic mitochondrial abnormalities are found in the latter. Amiodarone rarely causes proximal and distal muscle weakness, or the accompanying tremor or distal sensory loss. Muscle biopsy reveals autophagic vacuoles with myeloid inclusions and debris. This is seen more commonly in patients with chronic kidney disease. The antimalarial drugs chloroquine and hydroxychloroquine produce a vacuolar myopathy, possibly by raising the intralysosomal pH so that the acid cathepsins that digest waste products in the lysosome are inoperable, so waste products accumulate in vacuoles.

Bacterial and Parasitic Diseases

Certain parasitic diseases can produce an illness by direct invasion of muscle. Weakness, fever, and eosinophilia are usually present. Bacterial pyomyositis is uncommon in North America but occurs more commonly in other parts of the world. It is attended by the signs of local infection and is often asymmetrical. A recent report, however, details two cases of pyomyositis initially diagnosed as PM. Patients presenting with disseminated pyomyositis may be difficult to distinguish clinically from those with IIM, especially in immunosuppressed individuals, who may not mount a systemic response.[15]

PATHOGENESIS

The precise mechanisms of cell damage and death in the IIMs are still unknown. T lymphocytes invading muscle fibers are predominantly lacking the costimulatory receptor CD28, indicative that they are chronically stimulated and end-differentiated. CD28^{null} lymphocytes are found in both muscle tissue and peripheral blood. They produce proinflammatory cytokines and release perforin and granzyme into targeted muscle cells. B cells and plasma cells are also present in myositis muscle infiltrates. The differentiation of B cells into antibody-producing plasma cells occurs within myositis muscle and is regulated by B cell–activating factor (BAFF). These antibodies are likely directed against autoantigens or cross-reacting viral antigens.

In addition to adaptive immune mechanisms, innate and nonimmune mechanisms also may be responsible for muscle cell damage and death. For example, Toll-like receptor (TLR)-3 and TLR-7 are expressed by infiltrating inflammatory cells and immature myoblasts in myositis muscle. TLR activation leads to production of inflammatory cytokines such as interleukin (IL)-6, thereby triggering or perpetuating inflammatory responses. Classic apoptosis is absent in IIM biopsies, but growing evidence suggests that autophagy could be responsible for myofiber death.

After damage, a muscle fiber, which is a syncytium, can regenerate. Initially, on light microscopy one can appreciate loss of striations leading to a homogeneous appearance, fiber size variation, atrophy, and centralization of the nuclei. In the inflammatory myopathies, cell-mediated cytotoxicity is a primary method of destruction. Macrophages and cytotoxic T cells invade the myofibers. The myocyte cytoplasm close to the invaginated cells appears vacuolated and swollen. Other regions of the same cell may show intense regeneration, as seen histologically by aggregates of nuclei with prominent nucleoli and fiber splitting. Thus degeneration and regeneration can coexist in the same fiber.

Cytokines and chemokines (Chapters 9 and 10) , which form an integrated network regulating inflammation, are produced by muscle fibers, immune cells, and endothelial cells. As an example, IL-1 is consistently found in myositis muscle. IL-1 facilitates transmigration of leukocytes by increasing the expression of adhesion molecules on endothelial cells. Further, IL-1 can decrease the proliferation of myoblasts. In some myositis patients, treatment with anakinra, an IL-1 receptor antagonist, results in improvement. This finding further supports a pathogenic role of IL-1. IL-1 acts in synergy with the other proinflammatory cytokines, tumor necrosis factor (TNF)-α and IL-17, to induce the production of IL-6 and CCL20 by myoblasts. The chemokine CCL20 recruits Th17 lymphocytes and immature dendritic cells into the muscle. IL-6 perpetuates inflammation by suppressing FOXP3$^+$ regulatory T cells. More recently, IL-15 has been reported to have a pathogenic role. IL-15 is expressed in muscle cells and myoblasts derived from PM and DM patients. IL-15 regulates T-cell activation and proliferation, and a higher expression of this cytokine at baseline may be associated with a poor response to immunosuppressive therapy.[16]

The role of type I interferons (IFN-Is) in the pathogenesis of DM has been extensively studied. The IFN-Is are a family of highly related glycoproteins, including IFN-α and IFN-β, that act on the type 1 IFN receptor. In DM muscle, the IFN-Is likely are produced by plasmacytoid dendritic cells (pDCs).[17] After activation of the IFN receptor, a cluster of IFN-I–inducible genes is transcribed, resulting in upregulation of several proteins, also known as the IFN signature, in the muscle and peripheral blood of patients with DM. Well known is the myxovirus resistance protein A (MxA), which, similar to other IFN-I–inducible proteins, has antiviral properties. This protein is overexpressed in both the muscle and skin in DM but not in PM or IBM. Further, decreased expression of MxA mRNA in peripheral blood mononuclear cells may correlate with decreasing muscle symptoms in juvenile DM patients. Indeed, some of the IFN-induced proteins may sustain autoimmunity in DM. For example, retinoic acid-inducible gene 1 (RIG-I) and melanoma differentiation-associated gene 5 (MDA-5) are IFN-I–induced proteins that induce IFN-I production in a positive feedback fashion. Both proteins contribute to IFN induction in response to measles virus infection,[18] nurturing the suspicion for a role of viral infections in myositis. Engel and colleagues[19] have demonstrated several important pathological distinctions between DM and PM and have suggested that in DM, the capillaries are the initial point of injury. A model of disease has been proposed in which primarily capillaries are damaged and myocytes are secondarily involved. Complement and immunoglobulins may be found in the capillary walls even where the remainder of the muscle is normal. The membrane attack complex of complement deposits there, and endothelial cells are swollen and pale. Even in unaffected regions of muscle, special staining reveals a marked decrease in capillary numbers. More advanced changes include microtubular endothelial inclusions and microvacuoles. Evaluation of lymphocyte populations in DM shows a high percentage of B cells and macrophages in the perivascular regions and an increasing frequency of CD4 T cells and pDCs toward the perimysium and endomysium. In DM, especially the juvenile-onset form, inflammation and necrosis followed by atrophy appear in a perifascicular pattern. Perivascular lymphocytic infiltrates, typical of later disease, have not been described as an early change. Consistent with the importance of immune complexes, histological abnormalities in the DM skin are indistinguishable from the changes in lupus.

In contrast to DM, PM and IBM do not demonstrate marked capillary changes, the perivascular infiltrates are less pronounced, and T-cell infiltrates in the perimysial and endomysial regions are more pronounced. Nonnecrotic fibers may be surrounded by T cells and macrophages. T cells are enriched for the CD8 subset. There is negligible evidence of natural killer (NK) cell presence. Should cytotoxic T cells prove to be the major effector cells, cloned T cells could lead to relevant antigenic targets. CD8 T cells recognize their antigenic peptide in association with major histocompatibility complex (MHC) class I molecules. Although resting normal muscle has very low class I expression, it is upregulated in regenerating and degenerating fibers found in both inflammatory and noninflammatory myopathies. Interestingly, in DM, class I expression is upregulated predominantly in the perifascicular regions, around sites of atrophy, and near sites of cellular invasion. In contrast, in PM, class I expression may be diffusely upregulated even where there is no cellular infiltrate. IBM shows a more focal class I distribution in regions of T-cell invasion. The presence of focal regions of MHC class I expression in nonnecrotic fibers at the site of activated CD8 T cells is compatible with cytotoxicity as a prime mechanism of myocyte necrosis in IBM and PM. A pivotal study using transgenic mice demonstrated that abnormal accumulation of MHC class I molecules in the endoplasmic reticulum (ER) of muscle may initiate the ER stress response.[20]

KEY CONCEPTS

Differential Histological Features of Myositis

- In dermatomyositis the earliest changes involve vessel walls, and B cells and CD4 T cells predominate in the muscle biopsy.
- In polymyositis and inclusion body myositis, the dominant pathological feature is targeting and invasion of muscle cells by CD8 cytotoxic T cells.

Proposed pathogenic mechanisms for the development of the sporadic forms of IBM are complex and include both autoimmunity and degeneration. Many proteins also found in other neurodegenerative diseases have been shown to accumulate in IBM muscle. In addition, IBM patients typically do not respond to immunosuppressive medications. These findings likely suggest that IBM is a degenerative disease. However, IBM is sometimes associated with other autoimmune diseases. In IBM biopsy specimens, inflammatory cells are predominantly composed of CD8 T cells, and a majority of patients express a circulating antibody to cytosolic 5′-nucleotidase 1A, supporting the role of autoimmunity in the pathogenesis of IBM.

In necrotizing autoimmune myositis (also known as IMNM), there is necrosis of muscle fibers with myophagocytosis and regeneration and paucity of T-cell infiltration. Complement deposition on blood vessels has been reported. In muscle biopsies from patients with statin-associated necrotizing autoimmune myositis, MHC class-I upregulation is frequently seen.[13]

The pathogenic role of autoantibodies found in IIM patients remains uncertain. MSAs are found in 60–80% of patients and appear to delineate specific clinical entities; each group has a strong but not absolute human leukocyte antigen (HLA) association. In a patient with myositis and antihistidyl-tRNA synthetase (Jo-1) autoantibodies, sera available from long before the onset of symptoms or biochemical damage to muscle tissue contained the autoantibodies, suggesting that the autoantibodies were not merely a response to release of tissue antigens. The extraordinary specificity of MSAs for IIM and the lack of evidence for strong polyclonal stimulation in these diseases suggest that MSAs are related to the fundamental causative process in IIM.

The structures bound by MSAs are mostly intracellular ribonucleoproteins involved in protein synthesis, such as the aminoacyl-tRNA synthetases and the SRP. These autoantigens are found in every nucleated cell. In general, the antibodies bind to conformational epitopes and, at least in the case of the antisynthetases, block enzymatic activity. It is possible that a structural property of muscle allows these particular proteins to be presented to the immune system when the cells are damaged; alternatively, the capacity of muscle fibers to degenerate alongside intense regeneration within the same fiber may allow these proteins to be efficiently displayed.[21] Experiments have suggested that some aminoacyl-tRNA synthetases have a direct proinflammatory role through a subsidiary chemokine-like action.

A landmark study determined that cultured myoblasts express high levels of autoantigens, which are strikingly downregulated as cells differentiate into myotubes *in vitro.* These data strongly associate regenerating rather than mature muscle cells as the source of continuous autoantigen supply in autoimmune myositis.[21]

GENETICS

The IIMs do not exhibit a simple mode of inheritance, and the rare familial cases mostly reflect IBM of early onset. As noted above, there are HLA associations for particular MSAs. Specifically, HLA-DR52 has a strong association (90%) with antisynthetase-positive myositis in people of both European and African descent.[22] HLA DRB1*11:01 was recently shown to be associated with an increased risk of anti-HMGCR myopathy.[23]

NATURAL HISTORY

The prognosis for patients with IIM varies greatly with clinical type, autoantibodies, extraskeletal muscle involvement, and the interval between diagnosis and the start of treatment.

Patients with DM or myositis accompanying another connective tissue disease are likely to recover most of their strength with prompt and adequate therapy. Although recurrences are common, persistent profound weakness does not usually occur. Most patients with anti-Mi-2 autoantibodies also usually respond well to therapy. Strength usually recovers well in patients whose myositis is cancer related, but overall mortality due to the tumor is high. Indeed, an accompanying tumor remains one of the most frequent causes of death in patients with an IIM. Among the MSAs, anti-TIF-1γ and anti–NXP-2 antibodies are found with increased frequency in patients with cancer-associated DM.[24]

Patients with PM fare less well, even when those with IBM are rigorously excluded. A return to normal strength is very unusual, and each recurrence is likely to be followed by greater residual weakness, even if inflammation is fully controlled. IBM has a poorer prognosis, but it is possible that the gradual decline in strength can be halted for long periods by corticosteroid and/or cytotoxic therapy if continuing inflammation is present. Severe muscle weakness and atrophy and very high CK levels are prominent features in patients with anti-SRP autoantibodies. Patients with anti-MDA5 antibodies are at increased risk for developing progressive interstitial lung disease. Those with anti-Jo-1 autoantibodies or antibodies to another synthetase are likely to respond to therapy initially, but they typically require continuing immunosuppression to treat frequent recurrences. In this group morbidity and mortality are heavily influenced by the progression of lung involvement. Longitudinal outcome studies in DM and PM are few. Cardiac involvement, respiratory involvement, and cancer were the main causes of death in several cohort analyses. Disease course is monocyclic in approximately 20% of patients, polycyclic in 20%, and chronic in the remainder. Relapses have been noted in the initial years of therapy and after prolonged disease-free intervals; therefore periodic surveillance is warranted for at least 2 years after remission.

PATIENT MANAGEMENT

The treatment of myositis is based on controlling skeletal muscle inflammation and damage. Immunosuppressive therapy is used in the initial stages of the disease to reduce inflammation and muscle damage. There are very few randomized controlled trials of any of the immunosuppressive agents used; thus therapeutic regimens and responses have remained largely anecdotal. After the initial inflammation is controlled, strengthening exercises are useful in improving functional capabilities.

Corticosteroids

Corticosteroids are the main immunosuppressive agents used in myositis treatment. An initial course of pulses of methylprednisolone may be helpful, particularly in disease of acute onset, and may also be helpful in managing disease flares. If active muscle inflammation persists or the side effects of corticosteroids are severe, other immunosuppressive treatments are employed.

Second-Line and Third-Line Immunosuppressive Therapies

The most frequently used second-line agents in the treatment of myositis are azathioprine and methotrexate. Azathioprine reduces long-term disability. Methotrexate is useful in patients with little or no response to corticosteroid therapy. Combination therapies, such as methotrexate with azathioprine, are useful even if patients have failed to respond to one of the agents alone. High-dose intravenous immunoglobulin (IVIG) is of proven benefit in DM.[25] Whereas patients with statin-associated anti-HMGCR myopathy may be particularly responsive to IVIG,[26] its usefulness in PM is less predictable. Apheresis proved ineffective in a controlled blinded study.[27] Both cyclosporine and tacrolimus have been effective in some cases, as have cyclophosphamide and chlorambucil. The most recent therapeutic options include mycophenolate and rituximab. Recently a large randomized placebo-phase clinical trial was conducted to elucidate the role of a 44-week course of rituximab therapy in adults and children with refractory PM and DM. In general, 83% of refractory myositis patients met the Definition of Improvement (DOI), and the presence of antisynthetase and anti–Mi-2 autoantibodies was associated with a shorter time to improvement.[28,29] In refractory patients who have not responded to the above-mentioned treatments, biological agents such as etanercept, adalimumab, anakinra, abatacept, and high-dose cyclophosphamide have been tried with variable success.[16,30–33] Thus far, there has been no effective therapeutic regimen for IBM. However, IVIG has been reported to provide a transient response.[13]

Monitoring Disease Activity

Improvement in strength and normalization of serum CK activity are the best indirect measures of disease activity. A decrease in serum CK activity may herald clinical improvement, but corticosteroid treatment alone can reduce CK activity without associated clinical improvement. A lack of improvement in strength in a corticosteroid-treated patient may be due to the resistance of the inflammatory process, the presence of a corticosteroid-induced myopathy, muscle atrophy, and/or misdiagnosis. A diagnostic and therapeutic taper of the corticosteroids may then be warranted. If inflammation is present concurrently, other immunosuppressive agents are useful as the dosage of corticosteroids is lowered. If the CK value begins to rise, even if it is still within the normal range, and the symptoms of myositis are worsening in a patient whose disease has previously been controlled with corticosteroids, an increase in the dose may be warranted.

Treatment-Resistant Myositis

Some treatment-resistant PM patients have another disease. In such cases IBM or a liED-girdle muscular dystrophy should be suspected. Unlike other myositis patients, those with IBM rarely, if ever, improve in strength with immunosuppressive therapy, but stabilization of strength can be achieved in some IBM patients with immunosuppressive agents.[34,35] Patients with limb-girdle muscular dystrophies may mimic PM clinically. They may have inflammation on muscle biopsy and may occasionally have associated autoantibodies. Thus patients with a suspected IIM who do not respond to immunosuppressive therapy should undergo further evaluation, including genetic testing, to search for a limb-girdle muscular dystrophy.

Nonskeletal Muscle Involvement

Other organs frequently involved in myositis include the skin, lungs, and joints. Such organ involvement and the systemic features of myositis (fever and weight loss) usually improve with immunosuppressive therapy that controls inflammation in the skeletal muscle. Hydroxychloroquine and other antimalarials are useful in controlling the rashes associated with myositis.

DIAGNOSTIC TOOLS, EVALUATION, AND DIFFERENTIAL DIAGNOSIS

CLINICAL PEARLS

Clinical Features That Suggest a Non-Idiopathic Inflammatory Myopathy Diagnosis

- Family history of a similar illness
- Weakness related to exercise, eating, or fasting
- Sensory, reflex, or other neurological signs
- Cranial nerve involvement
- Fasciculations
- Muscle cramping (severe)
- Myasthenia (increasing weakness with repeated contractions)
- Myotonia (difficulty relaxing a contracted muscle)
- Significant atrophy or hypertrophy early in the illness
- Marked asymmetry
- Dyspnea due to diaphragmatic weakness with normal chest x-ray

Clinical, laboratory, pathological, and electrodiagnostic findings contribute to the proper diagnosis of IIM. Even in individuals with typical clinical features, it is essential to exclude other diseases that may have similar symptoms and signs (Table 56.6). Certain clinical features should suggest a different diagnosis. These include a family history of a similar illness; sensory, reflex, or other neurological changes; fasciculations; a relationship of the weakness to exercise, food intake, or fasting; major muscle cramping, myotonia (difficulty relaxing a contracted muscle), or myasthenia (increasing weakness with repeated contractions); significant early muscle atrophy or hypertrophy; marked asymmetry; weakness in the distribution of the cranial nerves; and dyspnea due to diaphragmatic weakness rather than lung fibrosis.

The single most useful laboratory feature of muscle destruction is elevation of the serum CK, although this is nonspecific, and a small proportion of patients—probably <5%—have a bonafide inflammatory muscle disease without ever having an elevated CK. Elevations of the serum levels of aldolase, serum glutamic-oxaloacetic transaminase (SGOT), serum glutamate pyruvate transaminase (SGPT), and lactate dehydrogenase (LDH) are as frequent but less specific for muscle disease. Unlike inflammatory markers for other autoimmune inflammatory diseases, those such as the erythrocyte sedimentation rate (ESR) and C-reactive protein are often not elevated. Although some studies have shown ESR to be elevated in 50% of patients, most experts find a substantially lower proportion of IIM patients to have an elevated

TABLE 56.6 Differential Diagnosis of Idiopathic Inflammatory Myopathy

Neuromuscular Disorders

Genetic muscular dystrophies
Metabolic myopathies
Disorders of carbohydrate metabolism: McArdle disease, phosphofructokinase deficiency, adult acid maltase deficiency, and others
Disorders of lipid metabolism: carnitine deficiency, carnitine palmitoyl transferase deficiency
Disorders of purine metabolism: myoadenylate deaminase deficiency
Mitochondrial myopathies
Spinal muscular atrophies
Neuropathies: Guillain–Barré and other autoimmune polyneuropathies, diabetes mellitus, porphyria
Myasthenia gravis and Eaton–Lambert syndrome
Amyotrophic lateral sclerosis
Myotonic dystrophy and other myotonias
Familial periodic paralysis

Endocrine and Electrolyte Disorders

Hypokalemia, hypercalcemia, hypocalcemia, hypomagnesemia
Hypothyroidism, hyperthyroidism
Cushing syndrome, Addison disease

Toxic Myopathies (Partial List)

Alcohol
Amiodarone
Chloroquine and hydroxychloroquine
Cocaine
Colchicine
Corticosteroids
D-penicillamine
Ipecac
Statins and other lipid-lowering agents
Zidovudine (AZT)

Infections

Viral: HIV, human T-lymphotropic virus 1 (HTLV-1), influenza
Bacterial: *Staphylococcus, Streptococcus, Clostridia*
Parasitic: toxoplasmosis, trichinosis, schistosomiasis, cysticercosis

Miscellaneous

Polymyalgia rheumatica
Vasculitis
Eosinophilia myalgia syndrome
Paraneoplastic syndromes

ESR, even with active disease.[36] Likewise, hematological abnormalities, including anemia, are uncommon and rarely related to the underlying myopathy. If a significant abnormality is found, the physician should be alert to another cause for it.

Electromyographic (EMG) abnormalities are frequently present. Although the test is useful for excluding some neurological diseases that resemble IIM, it is painful for many patients and not useful for following the course of the illness.

MRI, especially a combination of the T1 and the fat-suppressed T2 (STIR) sequences, is remarkably useful in defining the extent of involvement and in planning a biopsy (Fig. 56.2). Whole-body MRI has been shown to facilitate the characterization of inflammatory myopathy, as certain patterns of muscle and subcutaneous tissue inflammation were predictive of the IIM subset (DM, PM, or IBM). In a recent study in juvenile DM, whole-body MRI was able to reveal inflammatory changes of subcutaneous and fascial tissue that was undetectable on clinical examination.[37] MRI can also help differentiate active disease from chronic disease, with active myositis being notable for changes consistent with muscle edema on T2-weighted images, and chronic myositis revealing a decrease in muscle bulk and replacement by adipose on T1-weighted images.[38] Although not specific, the changes of inflammatory myopathy on imaging can provide considerable assistance in confusing cases, as well as help in choosing a site to biopsy.

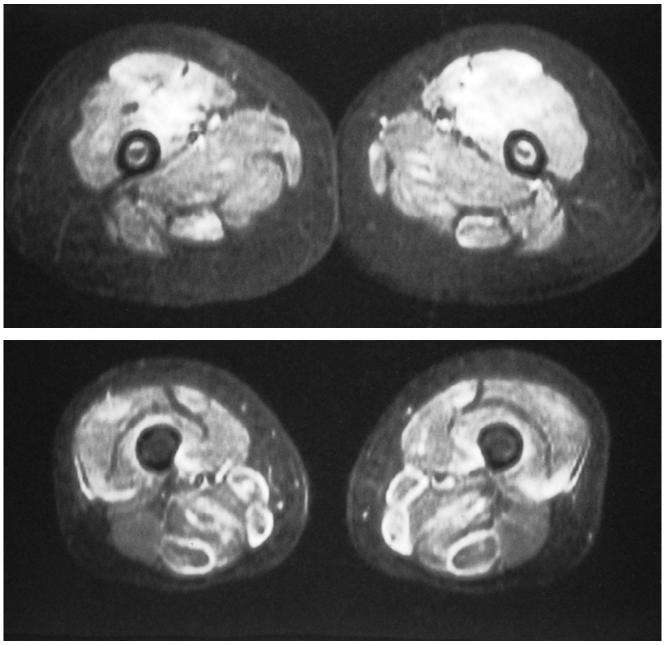


FIG 56.2 Magnetic Resonance Images of the Upper and Lower Thighs of a Patient With Dermatomyositis Using the Fat-Suppressed T2 (STIR) Technique. With this technique inflammation appears as a bright signal; normal muscle is gray; bone, fat, fascia, and normal skin are dark. Blood vessels may appear as bright spots. Note the remarkable symmetry of the inflammation. In this patient most of the involvement is in the quadriceps in the upper thighs and around the periphery of the hamstring muscle group.

A muscle biopsy should be performed in every suspected case of myositis (Fig. 56.3). Although the patchy involvement means that the biopsy can occasionally miss inflammation, certain confounding diagnoses—for example amyloidosis, eosinophilic myositis, dystrophy, or some metabolic myopathies as well as the important variant IBM—can be diagnosed definitively only by biopsy. The identification of autoantibodies, particularly the myositis-specific autoantibodies, has distinct clinical and prognostic use.

ON THE HORIZON

- Development of revised and updated classification criteria for inflammatory myopathies is needed.
- Improved diagnostic classification of cases through novel autoantibodies as well as immunohistochemistry.
- Long-term studies are needed to better characterize the prognosis of idiopathic inflammatory myopathies, particularly in relation to recently discovered autoantibodies.

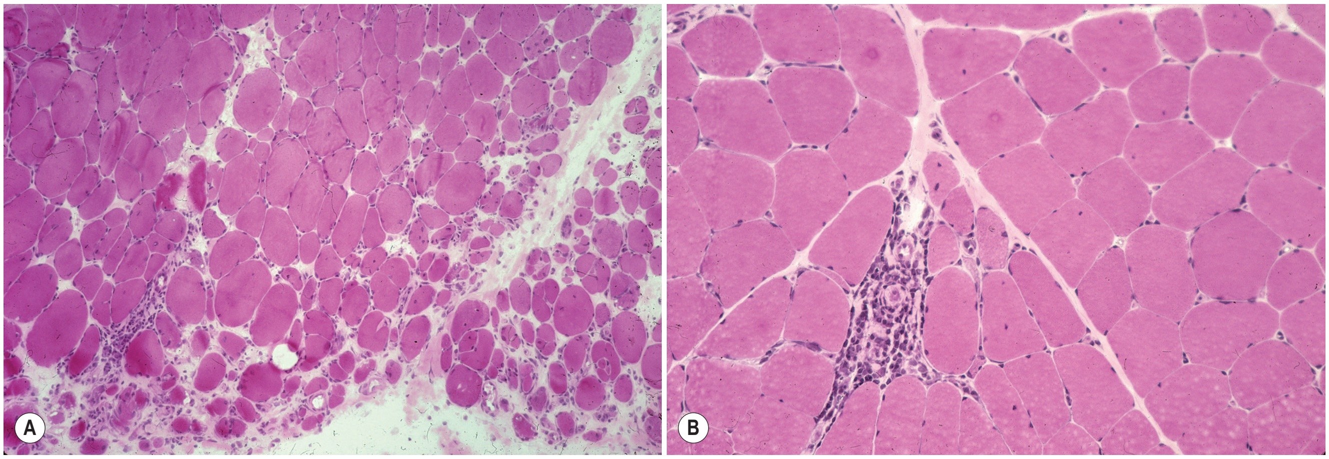


FIG 56.3 Biopsy in Dermatomyositis. (A) Low-power (original magnification ×100) view of a muscle biopsy from a patient with dermatomyositis. Note the marked variation in fiber size and the large number of atrophic myocytes, particularly at the periphery of the fascicles. (B) High-power (original magnification ×200) view of inflammation around the vessels in the muscle biopsy of a patient with dermatomyositis. There are nearby atrophic cells and cells whose nuclei have moved away from the periphery of the cell (centralized nuclei).

PITFALLS

It is increasingly apparent that the boundary between IIM and some genetically determined myopathies cannot be cleanly drawn. Recently dystrophies with an extraordinary variety of clinical manifestations (with regard to age and distribution of weakness) have also been described.[39] Not only can inflammation be seen on biopsy in some patients, but a partial clinical response to corticosteroids also can occur. Furthermore, it is increasingly recognized that mitochondrial abnormalities can be limited to groups of skeletal muscles, leading to confusion with IIM. Toxic myopathies, of course, will continue to occur with the release of new drugs and thus will continue to be a possible source of diagnostic confusion.

Thus not only must the history, physical examination, and biopsy be performed and interpreted with compulsiveness and care, but molecular diagnostic techniques also must be employed by clinicians in pursuit of an accurate diagnosis and appropriate therapy. The correct response to disease that persists in the face of powerful immunosuppressive therapy is a careful rethinking of the diagnosis, including, on occasion, rebiopsy and molecular consultation.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Sontheimer RD. Dermatomyositis: an overview of recent progress with emphasis on dermatologic aspects. Dermatol Clin 2002;20:387–408.
2. Bohan A, Peter JB. Polymyositis and dermatomyositis (first of two parts). N Engl J Med 1975;292:344–7.
3. Bohan A, Peter JB. Polymyositis and dermatomyositis (second of two parts). N Engl J Med 1975;292:403–7.
4. Hoogendijk JE, Amato AA, Lecky BR, et al. 119th ENMC international workshop: trial design in adult idiopathic inflammatory myopathies, with the exception of inclusion body myositis, 10–12 October 2003, Naarden, The Netherlands. Neuromuscul Disord 2004;14:337–45.
5. Rose MR. 188th ENMC International Workshop: Inclusion Body Myositis, 2–4 December 2011, Naarden, The Netherlands. Neuromuscul Disord 2013;23:1044–55.
6. Kao AH, Lacomis D, Lucas M, et al. Anti-signal recognition particle autoantibody in patients with and patients without idiopathic inflammatory myopathy. Arthritis Rheum 2004;50:209–15.
7. Christopher-Stine L, Casciola-Rosen LA, Hong G, et al. A novel autoantibody recognizing 200-kd and 100-kd proteins is associated with an immune-mediated necrotizing myopathy. Arthritis Rheum 2010;62:2757–66.
8. Mammen AL, Chung T, Christopher-Stine L, et al. Autoantibodies against 3-hydroxy-3-methylglutaryl-coenzyme A reductase in patients with statin-associated autoimmune myopathy. Arthritis Rheum 2011;63:713–21.
9. Lahouti AH, Amato AA, Christopher-Stine L. Inclusion body myositis: update. Curr Opin Rheumatol 2014;26:690–6.
10. Lloyd TE, Christopher-Stine L, Pinal-Fernandez I, et al. Cytosolic 5’-nucleotidase 1A as a target of circulating autoantibodies in autoimmune diseases. Arthritis Care Res (Hoboken) 2016;68:66–71.
11. Argov Z, Eisenberg I, Mitrani-Rosenbaum S. Genetics of inclusion body myopathies. Curr Opin Rheumatol 1998;10:543–7.
12. Mathews MB, Bernstein RM. Myositis autoantibody inhibits histidyl-tRNA synthetase: a model for autoimmunity. Nature 1983;304:177–9.
13. Dalakas MC. Pathophysiology of inflammatory and autoimmune myopathies. Presse Med 2011;40:e237–47.
14. Mor A, Wortmann RL, Mitnick HJ, et al. Drugs causing muscle disease. Rheum Dis Clin North Am 2011;37:219–31, vi.
15. Walji S, Rubenstein J, Shannon P, et al. Disseminated pyomyositis mimicking idiopathic inflammatory myopathy. J Rheumatol 2005;32:184–7.
16. Zong M, Loell I, Lindroos E, et al. Effects of immunosuppressive treatment on interleukin-15 and interleukin-15 receptor alpha expression in muscle tissue of patients with polymyositis or dermatomyositis. Ann Rheum Dis 2012;71:1055–63.
17. Greenberg SA, Pinkus JL, Pinkus GS, et al. Interferon-alpha/beta-mediated innate immune mechanisms in dermatomyositis. Ann Neurol 2005;57:664–78.
18. Ikegame S, Takeda M, Ohno S, et al. Both RIG-I and MDA5 RNA helicases contribute to the induction of alpha/beta interferon in measles virus-infected human cells. J Virol 2010;84:372–9.

19. Engel AG, Arahata K, Emslie-Smith A. Immune effector mechanisms in inflammatory myopathies. Res Publ Assoc Res Nerv Ment Dis 1990;68:141–57.
20. Nagaraju K, Casciola-Rosen L, Lundberg I, et al. Activation of the endoplasmic reticulum stress response in autoimmune myositis: potential role in muscle fiber damage and dysfunction. Arthritis Rheum 2005;52:1824–35.
21. Casciola-Rosen L, Nagaraju K, Plotz P, et al. Enhanced autoantigen expression in regenerating muscle cells in idiopathic inflammatory myopathy. J Exp Med 2005;201:591–601.
22. Goldstein R, Duvic M, Targoff IN, et al. HLA-D region genes associated with autoantibody responses to histidyl-transfer RNA synthetase (Jo-1) and other translation-related factors in myositis. Arthritis Rheum 1990;33:1240–8.
23. Mammen AL, Gaudet D, Brisson D, et al. Increased frequency of DRB1*11:01 in anti-hydroxymethylglutaryl-coenzyme A reductase-associated autoimmune myopathy. Arthritis Care Res (Hoboken) 2012;64:1233–7.
24. Fiorentino DF, Chung LS, Christopher-Stine L, et al. Most patients with cancer-associated dermatomyositis have antibodies to nuclear matrix protein NXP-2 or transcription intermediary factor 1gamma. Arthritis Rheum 2013;65:2954–62.
25. Dalakas MC, Illa I, Dambrosia JM, et al. A controlled trial of high-dose intravenous immune globulin infusions as treatment for dermatomyositis. N Engl J Med 1993;329:1993–2000.
26. Mammen AL, Tiniakou E. Intravenous immune globulin for statin-triggered autoimmune myopathy. N Engl J Med 2015;373:1680–2.
27. Miller FW, Leitman SF, Cronin ME, et al. Controlled trial of plasma exchange and leukapheresis in polymyositis and dermatomyositis. N Engl J Med 1992;326:1380–4.
28. Oddis CV, Reed AM, Aggarwal R, et al. Rituximab in the treatment of refractory adult and juvenile dermatomyositis and adult polymyositis: a randomized, placebo-phase trial. Arthritis Rheum 2013;65:314–24.
29. Aggarwal R, Bandos A, Reed AM, et al. Predictors of clinical improvement in rituximab-treated refractory adult and juvenile dermatomyositis and adult polymyositis. Arthritis Rheumatol 2014;66:740–9.
30. The Muscle Study Group. A randomized, pilot trial of etanercept in dermatomyositis. Ann Neurol 2011;70:427–36.
31. Park J-K, Yoo H-G, Ahn D-S, et al. Successful treatment for conventional treatment-resistant dermatomyositis-associated interstitial lung disease with adalimumab. Rheumatol Int 2012;32:3587–90.
32. Lahouti AH, Brodsky RA, Christopher-Stine L. Idiopathic inflammatory myopathy treated with high-dose immunoablative cyclophosphamide—A long-term follow-up study. JAMA Neurol 2015;72:1205–6.
33. Kerola AM, Kauppi MJ. Abatacept as a successful therapy for myositis—a case-based review. Clin Rheumatol 2015;34:609–12.
34. Leff RL, Miller FW, Hicks J, et al. The treatment of inclusion body myositis: a retrospective review and a randomized, prospective trial of immunosuppressive therapy. Medicine (Baltimore) 1993;72:225–35.
35. Sayers ME, Chou SM, Calabrese LH. Inclusion body myositis: analysis of 32 cases. J Rheumatol 1992;19:1385–9.
36. Rider LG, Miller FW. Laboratory evaluation of the inflammatory myopathies. Clin Diagn Lab Immunol 1995;2:1–9.
37. Malattia C, Damasio MB, Madeo A, et al. Whole-body MRI in the assessment of disease activity in juvenile dermatomyositis. Ann Rheum Dis 2014;73:1083–90.
38. Cantwell C, Ryan M, O'Connell M, et al. A comparison of inflammatory myopathies at whole-body turbo STIR MRI. Clin Radiol 2005;60:261–7.
39. Emery AE. The muscular dystrophies. BMJ 1998;317:991–5.

MULTIPLE-CHOICE QUESTIONS

1. A 42-year-old Asian woman comes to your office complaining of a rash on her face and upper back that started 2 months ago. Around the same time, she developed joint pain and swelling in both hands. Approximately a week ago, she noticed mouth sores as well as weakness, with difficulties walking upstairs. Other medical problems include an episode of pulmonary collapse resulting in intensive care unit admission and requiring intubation and treatment with antibiotics and steroids 3 years ago, but she had full recovery after this. Otherwise, she has been healthy. She has never smoked and does not drink alcohol. Her mother has rheumatoid arthritis. She takes no medications. Physical examination reveals heliotrope rash, Gottron sign, shawl sign, nail bed erythema, dilated capillary loops, and swelling of the hands. There is extensive ulceration and necrosis of the fourth digit. Neurological examination demonstrates symmetrically reduced hip flexion (4/5) muscle strength. Further evaluation reveals a serum CK level of 100 U/L (normal range: 24–170 U/L) and an aldolase level of 15.6 U/L (normal = <8.1 U/L). Which of the following antibodies is most likely to be detected in this patient's serum?
   A. Anti-Jo-1
   B. Anti-SRP
   C. Anti-RNP
   D. Anti-MDA5
   E. Anti-La

2. A 58-year-old Caucasian man began noticing progressive weakness with bilateral foot slapping 3 years ago. He began struggling with stairs and progressive weakness in his handgrip about 1 year later. He also began having difficulties walking on uneven ground. He has not had any difficulty swallowing. His past medical history is significant for elevated serum CK levels of 1400 (normal range: 24–170 U/L) 5 years ago and hypercholesterolemia. His family history includes breast cancer in his mother. He has a 33-pack-year smoking history, but he quit 2 years ago. He has no other medical problems. Neurological examination reveals reduced triceps strength (right: 4+/5, left: 4/5), finger flexion (2/5 bilaterally), knee extension (4+/5 bilaterally), and left wrist flexion (4+/5) muscle strength. The patient's laboratory tests show an elevated CK level of 1030 U/L. Electromyography reveals positive waves, fibrillations, and mixed neurogenic and myopathic motor units in the involved muscles. A quadriceps muscle biopsy is performed. Which of the following findings is most likely present on biopsy?
   A. Many necrotic muscle fibers and sparse inflammatory cells
   B. Perifascicular atrophy
   C. Endomysial inflammatory infiltrate and rimmed vacuoles
   D. MAC depositions on small blood vessels

3. Which of the following antibodies is associated with the disease presented in question 2?
   A. Anti-HMGCR
   B. Anti-SAE
   C. Anti-c5N1A
   D. Anti-Mi-2
   E. Anti-SRP
   F. No recognized antibodies are associated with this condition.

57

Spondyloarthritis

John D. Reveille

The term *spondyloarthritis* (SpA) (otherwise known as *spondyloarthropathy*) encompasses a heterogeneous group of inflammatory diseases characterized by spinal and peripheral joint arthritis, inflammation of the attachments of ligaments and tendons to bones (enthesitis), and, at times, mucocutaneous, ocular, and/or cardiac manifestations. These disorders show familial aggregation and are typically associated with genes of the major histocompatibility complex (MHC), particularly human leukocyte antigen (HLA)-B27 (Chapter 5). The SpA spectrum includes *(i)* axial spondyloarthritis (AxSpA), including ankylosing spondylitis (AS); *(ii)* reactive arthritis (ReA), known previously as *Reiter syndrome; (iii)* psoriatic arthritis (PsA) and/or spondylitis; *(iv)* enteropathic arthritis and/or spondylitis associated with the inflammatory bowel diseases (IBD), ulcerative colitis (UC), or Crohn disease (CD); and *(v)* undifferentiated SpA, which encompasses patients expressing elements of, but failing to fulfill, accepted criteria for one of the above diseases. In addition, isolated acute anterior uveitis (AAU)[1] and spondylitic heart disease (complete heart block and/or lone aortic regurgitation)[2] associated with HLA-B27 may also be classified within the spectrum of SpA.

CLASSIFICATION OF SPONDYLOARTHRITIS

There are no diagnostic criteria for any of the SpA types. Classification criteria have been developed to provide greater specificity in clinical studies. The European Spondyloarthropathy Study Group (ESSG) criteria for SpA were developed in 1991, have been validated in numerous population groups, and remain the basis of many clinical and epidemiological studies (Table 57.1).[3] The modified New York criteria, developed in 1966 and modified in 1984, remain the "gold standard" for AS.[4] However, this depends on the presence of radiographic sacroiliitis, which takes up to 10 years to develop after the onset of inflammatory back pain. For PsA, the classification criteria for psoriatic arthritis (CASPAR) (see Table 57.1)[5] have stemmed from the perception that the ESSG criteria were not "sensitive" enough.

The International League Against Rheumatism (ILAR) juvenile idiopathic arthritis (JIA) classification criteria for enthesitis-related arthritis (ERA) have been proposed for juvenile SpA (JSpA; see Table 57.1). No criteria have been validated for enteropathic arthritis.

New criteria have been developed by the Assessment of SpondyloArthritis International Society (ASAS) for axial SpA based on the presence of inflammatory back pain that take advantage of imaging innovations, such as magnetic resonance imaging (MRI) and HLA-B27 typing.[6] More recently, criteria for peripheral spondyloarthritis based on the presence of peripheral arthritis, enthesitis, or dactylitis have also been developed utilizing these innovations.[7] The result of this is a reclassification of all the various spondyloarthritides under the diagnoses of axial and peripheral SpA.

EPIDEMIOLOGY

The frequency of SpA in general and AS in particular varies in different populations and parallels the frequency of HLA-B27. The prevalence of AS varies from 0.2% to 0.7% among people of European ancestry[8-10] and has been reported in similar frequencies in eastern Asia (Table 57.2). Higher rates of prevalence have been reported in Eskimo groups from Siberia and Alaska as well as in groups from Scandinavia and as high as 4.3% in male Haida Native Canadian groups, where the frequency of HLA-B27 is 50%.[8] AS is much less frequent in persons of African and Japanese descents, in whom HLA-B27 is rare.

The prevalence of ReA is unknown and probably varies over time, depending on endemic rates of the sexually acquired (*Chlamydia*) and enteric (*Shigella, Salmonella, Campylobacter*) infections that trigger it. It appears that the frequency of ReA has dramatically declined following the adoption of safer sexual practices and better sanitation, at least in Western countries, in the wake of the human immunodeficiency virus (HIV) epidemic.

Psoriasis affects 1–3% of the general population.[9] The frequency of PsA is less clear and is higher in those with more severe disease; population studies among Caucasians estimate a frequency of approximately 0.1%.[8]

The prevalence of IBD is 100–200 per 100 000 among Caucasians, with an equal male-to-female ratio.[8,10] It is rare in people of African and Asian descents. The risk of spondylitis and peripheral arthritis varies in different reports, perhaps reflecting the subspecialty of the observer. Spondylitis occurs in as many as 15–20% of those with IBD. In general, peripheral arthritis occurs less frequently in those with UC (up to 10%) than in those with CD (up to 20%), although its frequency tends to be higher in series where the assessor was a rheumatologist.[8,10]

PATHOGENESIS

Genetics of Spondyloarthritis

Familial Aggregation

Susceptibility to AS is clearly attributable to genetic factors, with a sibling recurrence risk ratio as high as 82 and twin-based studies

TABLE 57.1 **Current Classification Criteria for Spondyloarthritis**

A. European Spondyloarthropathy Study Group (ESSG) Criteria for Spondyloarthritis

1. Inflammatory back pain or synovitis (asymmetrical, lower extremity) *plus* one of the following:
   - (a) Alternating buttock pain
   - (b) Sacroiliitis
   - (c) Heel pain (enthesitis)
   - (d) Positive family history
   - (e) Psoriasis
   - (f) Crohn disease, ulcerative colitis
   - (g) Urethritis or cervicitis or acute diarrhea in the preceding 4 weeks

B. The Modified New York Criteria for Ankylosing Spondylitis[4]

1. Clinical criteria:
   - (a) Low-back pain and stiffness for more than 3 months which improves with exercise, but is not relieved by rest
   - (b) Limitation of motion of the lumbar spine in both the sagittal and frontal planes
   - (c) Limitation of chest expansion relative to normal values correlated for age and gender
2. Radiological criterion:
   - (a) Sacroiliitis grade 2 bilaterally or grade 3–4 unilaterally
   - (b) *Definite ankylosing spondylitis* if the radiological criterion is associated with at least one clinical criterion

C. The Classification Criteria for Psoriatic Arthritis (CASPAR) Criteria for Psoriatic Arthritis

1. Inflammatory joint disease plus at least three points from the following features:
   - (a) Current psoriasis (assigned a score of 2; all others assigned a score of 1)
   - (b) History of psoriasis
   - (c) Family history of psoriasis
   - (d) Dactylitis
   - (e) Juxtaarticular new bone formation
   - (f) Rheumatoid factor seronegativity
   - (g) Nail dystrophy

D. ASAS Criteria for Axial Spondyloarthritis in Patients With Chronic Low Back Pain for at Least 3 Months

1. Sacroiliitis on imaging (either MRI findings of inflammatory disease in the sacroiliac joints or sacroiliitis on standard pelvic X-rays by New York criteria) plus ONE of the following OR HLA-B27 positivity plus TWO of the following:
   - (a) Inflammatory low back pain
   - (b) Arthritis
   - (c) Enthesitis
   - (d) Dactylitis
   - (e) Psoriasis
   - (f) Crohn disease or ulcerative colitis
   - (g) Good response to nonsteroidal antiinflammatory agents
   - (h) Positive family history of SpA
   - (i) Presence of human leukocyte antigen (HLA)-B27
   - (j) Elevated cross-reactive protein

E. ASAS Criteria for Peripheral Spondyloarthritis—Arthritis or Enthesitis or Dactylitis With at Least One of:

a. Uveitis
b. Psoriasis
c. Crohn disease or ulcerative colitis
d. Previous infection
e. Presence of HLA-B27
f. Sacroiliitis on imaging
OR
with at least two of:
a. Inflammatory low back pain
b. Arthritis
c. Enthesitis
d. Dactylitis
e. Positive family history of SpA

F. The International League Against Rheumatism (ILAR) Juvenile Idiopathic Arthritis Classification Criteria for Enthesitis-Related Arthritis (ERA)

Arthritis and enthesitis
OR
Arthritis or enthesitis with at least two of:

1. Sacroiliac joint tenderness and/or inflammatory spinal pain
2. Presence of HLA-B27
3. Family history in at least one first- or second-degree relative of medically confirmed HLA-B27-associated disease
4. Anterior uveitis that is usually associated with pain, redness, or photophobia
5. Onset of arthritis in a boy after 8 years of age

Exclusions

- Psoriasis confirmed by a dermatologist in at least one first- or second-degree relative
- Presence of systemic arthritis

estimating disease heritability to exceed 90%.[11] The concordance rate for AS in identical twins has been reported to be as high as 63% compared with 23% in nonidentical twins.[11] The concurrence rate for psoriasis in monozygotic twins is 70 versus 15–30% in dizygotic twins. Recurrence risk for parents and sibs of patients with CD is 4.8% and 7%, respectively, and for UC 0.9 and 1.2%, respectively.[11]

HLA-B27 and Spondyloarthritis

HLA-B27, which is encoded in the MHC class I region (Chapter 5), confers the greatest known risk for AS and is found in up to 90% of AS patients of European ancestry[11-13] (Table 57.3), as opposed to 6–8% of healthy Caucasians.[12,14] The estimated prevalence of HLA-B27 in the United States is 6.1% overall, is highest in younger individuals (7.5% before age 50 years), and falls rapidly over age 50 years (3.3%).[14]

KEY CONCEPTS

The Genetic Basis of Spondyloarthritis

- Human leukocyte antigen (HLA)-B27 comprises nearly 80% of the overall genetic susceptibility to ankylosing spondylitis (AS) and contributes heavily to susceptibility to reactive arthritis, psoriatic arthritis, and enteropathic spondylitis.
- Additional influences seem to come from other major histocompatibility complex (MHC) genes, including *HLA-A*02, B*40* for AS and *C*0602* for psoriasis, whose identification has been confounded by linkage to HLA-B27.
- Genome-wide association studies (GWAS) utilizing dense single nucleotide polymorphism (SNP) mapping have located at least 70 genes or genetic regions in AS susceptibility.
- Over 60 genes have been identified thus far in psoriasis pathogenesis.
- GWAS utilizing dense SNP mapping will likely locate many of the remaining genes in AS susceptibility.

TABLE 57.2 The Epidemiology of Spondyloarthritis

Ethnic Group	HLA-B27 Frequency (%)	Prevalence of AS %(Criterion)	Prevalence of PsA % (Criterion)	Prevalence of SpA % (Criterion)
Europe				
Azores		n.a.	n.a.	1.6 (ESSG)
Czech Republic		0.09 (mNY)	0.05 (Vasey)	n.a.
France		0.08 (questionnaire)	0.19 (questionnaire)	0.3 (questionnaire)
Germany		0.86 (mNY)	0.29 (ESSG)	1.9 (ESSG
Greece		0.24 (mNY)	0.17 (ESSG)	0.49 (ESSG)
Iceland		0.13 (mNY)	0.14 (other)	0.49 (ESSG)
Italy		0.37 (mNY)	0.42 (mNY)	n.a.
Netherlands		0.24 (questionnaire plus Xray)	n.a.	n.a.
N. Norway(Sami)	24	1.8 (NY)	n.a.	n.a.
Norway-general	24	1.1-1.4(mNY)	0.23 (ICD)	n.a.
South Sweden		0.25 (ICD10)	0.25 (ICD10)	0.45 (ICD10)
America				
United States	6.1	0.52 (questionnaire plus X-ray)	0.16	1.4 (ESSG)
Mexico	4.6	0.02 (mNY)	0.02	n.a.
Alaska (Eskimo)		0.4 (NY)	<0.1	2.5 (ESSG)
Argentina		n.a.	0.07 (CASPAR)	n.a.
Asia				
China		0.25 (mNY)	0.02	0.78 (ESSG)
Iran		0.12 (unclear)	n.a.	0.23 (COPCORD)
Japan		0.0065 (NY)	0.001	0.0095
Pakistan		n.a.	n.a.	0.1 (COPCORD)
Siberia	37	1.1 (NY)	0.3	2.5 (ESSG)
Taiwan		0.19–0.54 (mNY)	n.a.	n.a.
Turkey		0.49 (mNY)	n.a.	1.05 (ESSG)
Vietnam		n.a.	n.a.	0.28 (COPCORD)

AS, ankylosing spondylitis; HLA, human leukocyte antigen; n.a., not available; PsA, psoriatic arthritis; SpA, spondyloarthritis; NY (New York) and mNY (modified New York criteria for AS); ESSG (European Spondyloarthropathy Study Group Criteria)

Approximately 70% of patients with ReA have HLA-B27, except in Africa, where no association of HLA-B27 is seen in those with HIV-associated SpA.[13]

HLA-B27 is found in 60–70% of patients with psoriatic spondylitis and in 25% of those with peripheral PsA.[12,13] Up to 70% of those with IBD-associated spondylitis have HLA-B27, although no HLA-B27 association is seen with asymptomatic sacroiliitis. Approximately 50% of patients with AAU alone are HLA-B27 positive.[12]

Over 146 molecular subtypes of HLA-B27 have been described this far (http://www.ebi.ac.uk/ipd/imgt/hla/nomenclature/index.html). The most common subtypes (HLA-B*27:05, B*27:02, B*27:04, B*27:07, B*27:14) are clearly associated with SpA. Two subtypes of HLA-B27—HLA-B*27:06 and B*27:09—found in Southeast Asia and Sardinia, respectively, appear not to be associated with AS,[13] possibly in part as a result of amino acid differences in the "B" pocket of the HLA antigen-binding cleft at positions 114 and 116 that could alter the composition and anchoring of peptides presented by these HLA-B27 subtypes. The other subtypes of HLA-B27 are too rare to have had disease associations established. These subtypes evolved from the parent allele HLA-B*27:05 along three lines (Fig. 57.1) in three distinct geographical regions.

The exact mechanism underlying the effect of HLA-B27 on disease susceptibility has still not been determined. One theory suggests that SpA results from a unique set of antigenic peptides, either bacterial or self, that are bound and presented by all disease-associated HLA-B27 subtypes (but not by other HLA class I molecules) to CD8 T cells, resulting in an HLA-B27–restricted cytotoxic T-cell response found only in joints and other affected tissues (the so-called arthritogenic peptide hypothesis). However, such peptides have not been reproducibly identified to date. In fact, gene association data showing interaction between HLA-B27 and a gain-of-function (GOF) variant of endoplasmic reticulum (ER)–associated aminopeptidase I (ERAP1) suggest that aberrant antigen processing may play a central role in AS susceptibility.[15] The same has been described for HLA-C*06:02 and psoriasis.[16] An alternative concept focuses on self-association as a unique property of the HLA-B27 molecule. HLA-B27 heavy chains can form homodimers *in vitro* that are dependent on disulfide binding through their cysteine-67 residues in the extracellular α_1 domain (as well as other cysteine residues in other domains) (Fig. 57.2).[13,16] This occurs as a result of B27 misfolding within the ER. The accumulation of misfolded protein results in a proinflammatory intracellular stress response through the stimulation of secretion of type 1 interferon (IFN). Also, HLA-B27 homodimers are detectable at the cell surface in patients with SpA, are capable of peptide binding, and are more abundantly expressed when the cell's antigen-presenting function is impaired. They are ligands for a number of natural killer (NK) and related cell surface receptors. Populations of synovial and peripheral blood monocytes, NK cells, and B and T lymphocytes from patients with SpA and controls carry receptors for HLA-B27 homodimers, including killer immunoglobulin-like receptor (KIR)3DL1 and KIR3DL2 and immunoglobulin (Ig)–like transcript 4 (ILT4).[13,15] These homodimers may act as a

TABLE 57.3 Genetic Factors Implicated in Spondyloarthritis

HLA-B27	AS = 90%, reactive arthritis = 70%; psoriatic spondylitis = 70%; peripheral psoriatic arthritis = 24%; IBD-associated arthritis = 60%. No association with psoriasis, Crohn disease, ulcerative colitis, or peripheral enteropathic arthritis
Other MHC genes	_AS=B*40, B*51, A*02, MICA*007, *019_ IBD—None for Crohn disease per se; _HLA-DRB1*01:03_ for peripheral arthritis with IBD;_HLA-DRA_ locus with UC; _HLA-C*06:02_ with psoriasis, _HLA-B*38, B*39_ with psoriatic arthritis
AS, CD, UC, and PS	*IL23R, IL10, ADCY3, IL12B, CDKAL1, CCL21, ZMIZ1, ETS1, SH2B3, FOS, TNP2, RPS6KB1, RAVER1, TYK2, DNMT3B, UBE2L3*
AS, PS, and CD	*SP140, FUT2*
AS, PS, and UC	*CD28, IRF5*
AS, CD, and UC	*RORC, HMGN2P18, ITLN1, FCGR2A, DNMT3A, IL1R2, CXCR2, GPR35, MST1, ANTXR2, PTGER4, ERAP2, C5orf56, IRGM, BACH2, C7orf72, JAK2, TLR4, ZNF365, NKX2-3, OR5B21, SMAD3, NOD2, IL27, ZPBP2, PTPN2*
AS and CD alone	*IL6R, ANKRD55, FGFR1OP, TNFSF8, NOTCH1, TFRSF1A/LTBR, RIC8B, LSM14*
AS and UC alone	*ASAP2, ITGAL, FAM118A*
AS and PS alone	*ERAP1, TNFAIP3, ZC3H12C, PPP2R3C, NOS2*
CD, UC, and PS	*PARK7, OSMR, RN7SKP226, IL2RA, PLAU, TSPAN14, NFKBIA, STAT3*
CD and PS	*FIBP, TAGAP*
UC and PS	*GNA-12, HDAC7A*
CD and UC	*SDF4, PTPN22, DENND1B, GCKR, FOSL2, THADA, PUS10, IL18R1, BANK1, IL21, DAP, NDFIP1, PRDM1, SKAP2, TNFSF15, CREM, GLYAT, CD6, IFNG-AS1, LINC00284, FNDC3A, UBAC2, CLEC16A, PRKCB, ORMDL3, CD226, KEAP1, PKIG, CD40, ZNF831, TNFRSF6B, UQCR10*
AS alone	*IL1R1, CMC1, NPM1P17, ACTA2, NPEPPS, ERN1*
CD alone	*HIPK1, GBAP1, TEX41, TG16L1, GAL3ST2, CPEB4, CCR6, JAZF1, MYRF, LRRK2, SR1, C17orf67, NFATC1, SBNO2, STK11, PPP5C, IFNGR2*
UC alone	*TNFRSF14, RNF186, HSPA6, NR5A2, IL8RA; LSP1, NXPE1, SLC39A11, HNF4A, UBASH3A*
PS/PsA alone	*IFNLR1, LCE3B, IL36RN, STAT4, IL17RD, FBXL1, IL4/IL13, TNIP1, TRAF3IP2, DEFB4, IL23A, IL22, GJB2, SETD1A, CARD14, SOCS3, STARD6, ILF3, CARM1, SLC9A8, RNF114*

HLA, human leukocyte antigen; MHC, major histocompatibility complex; *IL23R*, interleukin-23 receptor; *ADCY3*, adenylate cyclase 3; *CDKAL1*, CDK5 regulatory subunit-associated protein 1-like 1; *CCL21*, chemokine; CC motif, ligand 21; *IL12B*, The Interleukin-12 p40 subunit; *ZMIZ1*, zinc finger MIZ-domain containing 1; *ETS1*, V-ETS avian erythroblastosis virus E26 oncogene homologue 1; *SH2B3*, SH2B adaptor protein 3; *FOS*, V-FOS FBJ murine osteosarcoma viral oncogene homologue; *TNP2*, transition protein 2; *RPS6KB1*, ribosomal protein S6 kinase, 70-kDa, 1; *TYK2*, tyrosine kinase 2; *DNMT3B,* DNA methyltransferase 3b; *UBE2L3,* ubiquitin-conjugating enzyme E2L; *SP140*, nuclear body protein 140; *FUT2*, fucosyl transferase 2; *CD28*, CD28 antigen; *IRF5*, interferon regulatory factor 5; *RORC,* RAR-related orphan receptor C-3; *HMGN2P18,* high mobility group nucleosomal binding domain 2 pseudogene 18; *ITLN,* intelectin 1; *FCGR2A*, Fcγreceptor 2A; *DNMT3A,* DNA methyltransferase 3b; *IL1R2,* interleukin-1 receptor, type II; *CXCR2,* chemokine, CXC motif, receptor 2; *GPR35,* G protein–coupled receptor 35; *MST1*, macrophage stimulating 1; *ANTXR2*, anthraxin receptor receptor 2; *PTGER4*, Prostaglandin E receptor 4, EP4 subtype; *ERAP2*, endoplasmic reticulum–associated aminopeptidase 2; *C5orf56,* chromosome 5 open reading frame 56; *IRGM,* immunity-related GTPase family, M; *BACH2*, basic leucine zipper transcription factor 2; *C7orf72*, chromosome 7 open reading frame 72; *JAK2*, Janus kinase 2; *TLR4*, Toll-like receptor 4; *ZNF365,* Zinc finger protein 365; *NKX2-3*, NK2 homeobox 3; *OR5B21,* olfactory receptor family 5 subfamily B member 21; *SMAD3*, Mothers against decapentaplegic, Drosophila, homologue of, 3; *NOD2*, Nucleotide-binding oligomerization domain protein 2; *IL27*, interleukin-27; *ZPBP2*, zona pellucida–binding protein 2; *PTPN2*, protein tyrosine phosphatase, non–receptor type, 2; *IL6R*, interleukin-6 receptor; *ANKRD55,* ankyrin repeat domain–containing protein 55; *FGFR1OP*, fibroblast growth factor receptor 1 oncogene partner; *TNFSF8*, tumor necrosis factor ligand superfamily, member 8; *NOTCH1,* NOTCH, Drosophila, homologue of, 1; *TFRSF1A/LTBR*, tumor necrosis factor receptor superfamily, member 1A/ lymphotoxin B receptor; *RIC8B,* RIC8, *C. elegans*, homologue of, b; *LSM14*, LSM Protein A; *ASAP2*, development-and differentiation-enhancing factor 2; *ITGAL*, integrin, alpha-L; *FAM118A*, family with sequence similarity 118 member A; *ERAP1*, endoplasmic reticulum aminopeptidase I; *TNFAIP3,* tumor necrosis factor-α-induced protein 3; *ZC3H12C*, zinc finger CCCH domain-containing protein 12C; *PPP2R3C*, protein phosphatase 2, regulatory subunit B-double prime, γ; *NOS2,* nitric oxide synthase 2A; *PARK7*, Parkinson disease 7, autosomal recessive early-onset; *OSMR,* oncostatin M receptor; *RN7SKP226,* RNA, 7SK small nuclear pseudogene 226; *IL2RA*, interleukin-2, receptor alpha; *PLAU*, plasminogen activator, urinary; *TSPAN14*, tetraspanin 14; *NFKBIA,* nuclear factor of κ light chain gene enhancer in B-cell inhibitor, α; *STAT3,* signal transducer and activator of transcription 3; *FIBP,* fibroblast growth factor, acidic, intracellular binding protein; *TAGAP*, T-cell activation GTPase-activating protein; *GNA-12*, guanine nucleotide-binding protein, α-12; *HDAC7A*, histone deacetylase 7A; *SDF4*, stromal cell-derived factor 4; *PTPN22*, protein tyrosine phosphatase, non–receptor type, 22; *DENND1B*, DENN/MADD domain-containing protein 1B; *GCKR*, glucokinase regulatory protein; *FOSL2*, FOS-like antigen 2; *THADA*, thyroid adenoma-associated gene; *PUS10*, pseudouridylate synthase 10; *IL18R1*, Interleukin-18 receptor; BANK1, B-cell scaffold protein with ankyrin repeats 1; *IL21*, interleukin-21; *DAP*, death-associated protein; *NDFIP1,* NEDD4 family-interacting protein 1; *PRDM1*, PR domain-containing protein 1;*SKAP2*, SRC kinase associated phosphoprotein 2; *TNFSF15*, tumor necrosis factor ligand superfamily, member 15; *CREM*, cAMP response element modulator; *GLYAT,* glycine N-acyltransferase; *CD6*, CD6 molecule; *IFNG-AS1*, interferon gamma antisense RNA 1; *LINC00284*, long intergenic non–protein coding RNA 284; *FNDC3A*, fibronectin type III domain containing 3A; *UBAC2,* UBA domain containing 2; *CLEC16A*, C-type lectin domain family 16, member A; *PRKCB*, protein kinase C, subunit B; *ORMDL3,* ORM1-like protein 3;*CD226*, CD226 antigen; *KEAP,* KELCH-like ECH-associated protein 1; *PKIG*, protein kinase, cAMP-dependent catalytic, inhibitor gamma; *CD40*, CD40 antigen; *ZNF831*, zinc finger protein 831; *TNFRSF6B,* tumor necrosis factor receptor superfamily, member 6B; *UQCR10,* ubiquinol-cytochrome c reductase complex, 7.2-kDa subunit; *IL1R1*, interleukin-1 receptor, type I; CMC1, COX assembly mitochondrial protein 1, *S. cerevisiae*, homologue of; NPM1P17, nucleophosmin 1 (nucleolar phosphoprotein B23, numatrin) pseudogene 17; *ACTA2*, actin, α_2, smooth muscle, aorta; *NPEPPS*, aminopeptidase, puromycin-sensitive; *ERN1*, endoplasmic reticulum-to-nucleus signaling 1; *HIPK1,* homeodomain-interacting protein kinase 1; *GBAP1*, glucosylceramidase β pseudogene 1; *TEX41,* testis expressed 41 (non–protein coding); *ATG16L1*, autophagy 16-like1; *GAL3ST2*, galactose-3-O-sulfotransferase 2; *CPEB4,* cytoplasmic polyadenylation element-binding protein 4; *CCR6*, chemokine, cc motif, receptor 6; *JAZF1*, juxtaposed with another zinc finger gene 1; *MYRF*, chromosome 11 open reading frame 9; *LRRK2*, leucine-rich repeat kinase 2; *KSR1*, kinase suppressor of RAS 1; *C17orf67*, chromosome 17 open reading frame 67; *NFATC1*, nuclear factor of activated T-cells 1; *SBNO2*, strawberry notch, Drosophila, homologue of, 2; *STK11*, serine/threonine protein kinase 11; *PPP5C*, protein phosphatase 5, catalytic subunit; *IFNGR*, interferon-γreceptor 1; *TNFRSF14,* tumor necrosis factor receptor superfamily, member 14; *RNF186*, Ring finger protein186; *HSPA6*, heat-shock 70-kDa protein 6; *NR5A2*, nuclear receptor subfamily 5, group A, member 2; *IL8RA*, interleukin-8 receptor A; *LSP1*, lymphocyte specific protein 1; *NXPE1*, neurexophilin and PC-esterase domain family member 1; *SLC39A11*, solute carrier family 39 (zinc transporter), member 11; *HNF4A*, hepatocyte nuclear factor 4-α; *UBASH3A,* ubiquitin-associated and SH3 domain-containing protein A; *IFNLR1*, interferon-λ receptor 1; *LCE3B*, late cornified envelope protein 3b; *IL36RN*, Interleukin-36 receptor antagonist; *STAT4*, signal transducer and activator of transcription 4; *IL17RD*, interleukin-17 receptor D *FBXL1*, S-phase kinase-associated protein 2; *IL4/IL13*, Interleukin-4/interleukin-13; *TNIP1*, TNFAIP3-interacting protein 1; *TRAF3IP2*, TRAF3-interacting protein 2; *DEFB4*, Defensin β, 4A; *IL23A,* interleukin-23; *IL22*, interleukin-22; *GJB2*, gap junction protein, β_2; *SETD1A*, SET domain-containing protein 1A; *CARD14,* caspase recruitment domain-containing protein 14; *SOCS3*, suppressor of cytokine signaling 3; *STARD6*, START domain-containing protein 6; *ILF3,* interleukin enhancer–binding factor 3; *CARM1*, coactivator-associated arginine methyltransferase 1; *SLC9A8,* Solute carrier family 9 (sodium/hydrogen exchanger), member 8; *RNF114*, ring finger 114 protein.

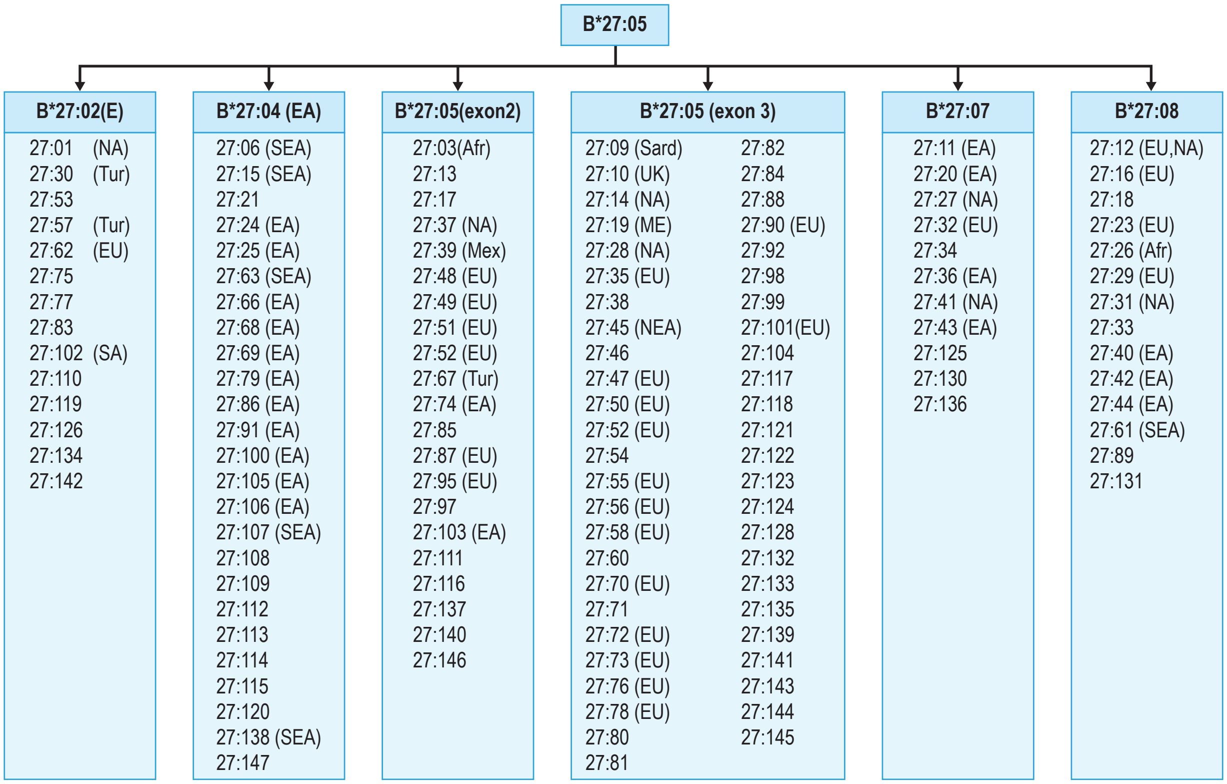

B*27:05
B*27:02(E)
27:01 (NA)
27:30 (Tur)
27:53
27:57 (Tur)
27:62 (EU)
27:75
27:77
27:83
27:102 (SA)
27:110
27:119
27:126
27:134
27:142
B*27:04 (EA)
27:06 (SEA)
27:15 (SEA)
27:21
27:24 (EA)
27:25 (EA)
27:63 (SEA)
27:66 (EA)
27:68 (EA)
27:69 (EA)
27:79 (EA)
27:86 (EA)
27:91 (EA)
27:100 (EA)
27:105 (EA)
27:106 (EA)
27:107 (SEA)
27:108
27:109
27:112
27:113
27:114
27:115
27:120
27:138 (SEA)
27:147
B*27:05(exon2)
27:03(Afr)
27:13
27:17
27:37 (NA)
27:39 (Mex)
27:48 (EU)
27:49 (EU)
27:51 (EU)
27:52 (EU)
27:67 (Tur)
27:74 (EA)
27:85
27:87 (EU)
27:95 (EU)
27:97
27:103 (EA)
27:111
27:116
27:137
27:140
27:146
B*27:05 (exon 3)
27:09 (Sard)
27:10 (UK)
27:14 (NA)
27:19 (ME)
27:28 (NA)
27:35 (EU)
27:38
27:45 (NEA)
27:46
27:47 (EU)
27:50 (EU)
27:52 (EU)
27:54
27:55 (EU)
27:56 (EU)
27:58 (EU)
27:60
27:70 (EU)
27:71
27:72 (EU)
27:73 (EU)
27:76 (EU)
27:78 (EU)
27:80
27:81
27:82
27:84
27:88
27:90 (EU)
27:92
27:98
27:99
27:101(EU)
27:104
27:117
27:118
27:121
27:122
27:123
27:124
27:128
27:132
27:133
27:135
27:139
27:141
27:143
27:144
27:145
B*27:07
27:11 (EA)
27:20 (EA)
27:27 (NA)
27:32 (EU)
27:34
27:36 (EA)
27:41 (NA)
27:43 (EA)
27:125
27:130
27:136
B*27:08
27:12 (EU,NA)
27:16 (EU)
27:18
27:23 (EU)
27:26 (Afr)
27:29 (EU)
27:31 (NA)
27:33
27:40 (EA)
27:42 (EA)
27:44 (EA)
27:61 (SEA)
27:89
27:131

FIG 57.1 The six major families of human leukocyte antigen (HLA)-B27 subtypes (HLA- B*27:59, B*27:64, B*27:65 and B*27:94 represent truncated genes with deletions of most of exons 2 and 3 and are not included above) are denoted in relationship to the "parent" subtype HLA-B*27:05. The ethnic or geographic origin of the subtypes is indicated in parentheses next to the allele: EU, European; NA, North American; SA, South American; EA, East Asian; SEA, Southeast Asian; NEA, Northeast Asian; ME, Middle East; Afr, African; UK, British Isles; Tur, Turkish; Mex, Mexican mestizo. Where no parentheses are given, the origin of the cell line from which the sequence came is unknown. Most B27 subtypes have evolved through five patterns of evolution along geographic lines. The first group, with HLA-B*27:02-B*27:23 (and B*27:30), appears to have evolved in Europe and the Middle East, entails anywhere from one to seven amino acid substitutions in the first (α_1) domain, and has the second (α_2) domain identical to B*27:05. The second group, including HLA-B*27:04 as the most common allele, evolved in Eastern Asia and includes a uniform amino acid substitution in the α_1 domain and anywhere from one to seven substitutions in the α_2 domain. The third and fourth groups seem to have evolved directly from HLA-B*27:05, from mutations in either exon 2 or 3, respectively. The fifth group, including HLA-B*27:07 as its leading member, evolved in southern and eastern Asia, the Middle East, and Sardinia with patterned amino acid substitutions in the α_2 domain. Finally, the sixth group, with HLA-B*27:08 as its common member evolved largely in East Europe, has different characteristic pattern of amino acid substitutions in the α_2 domain. Notable exceptions include HLA-B*27:13, B*27:109 and B*27:112-27:115, which have amino acid substitution outside α_1 and α_2 in the membrane proximal (α_3) domain. (Data are derived from http://www.ebi.ac.uk/cgi-bin/imgt/hla/allele.cgi.)

proinflammatory target or receptor for humoral or cell-mediated autoimmune responses. However, it is not yet known whether HLA-B27 homodimer formation is specific for, or even correlates with, the presence of SpA; in fact, disease does not develop in most HLA-B27–positive individuals.

Other MHC Genes and SpA Susceptibility

HLA-B27 is not the only genetic factor involved in AS and SpA susceptibility (see Table 57.3). Fewer than 5% of HLA-B27–positive individuals in the general population become affected,[11,13] whereas SpA will develop in up to 20% of HLA-B27–positive relatives of patients with AS. Family studies have suggested that HLA-B27 contributes less than 40% of the overall genetic risk for SpA,[12] whereas the entire effect of the MHC is about 50%.[11,13]

HLA-B60, a serologically defined HLA specificity that correlates with *HLA-B*40:01* on DNA analysis, has been described as augmenting the risk for AS in both HLA-B27–positive and HLA-B27–negative individuals from Europe and Taiwan.[11,13] Other MHC genes have also been implicated in AS in addition to B27, although their identification is complicated by the tight

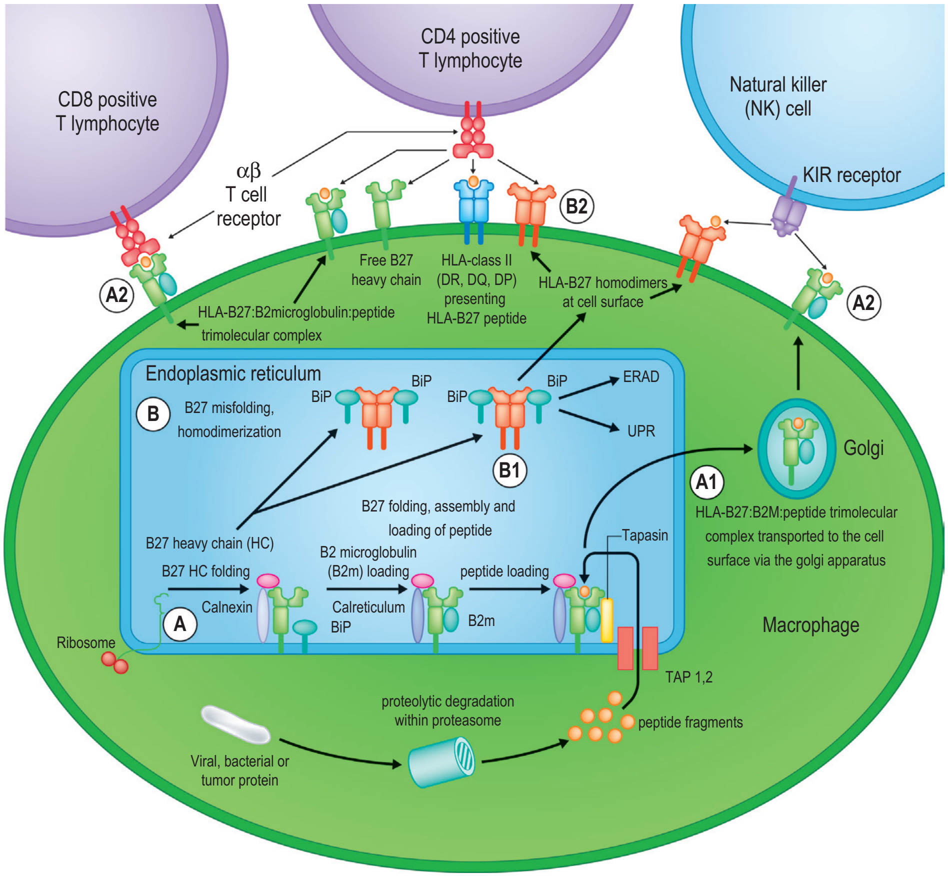

CD8 positive
T lymphocyte
CD4 positive
T lymphocyte
Natural killer
(NK) cell
αβ
T cell
receptor
KIR receptor
B2
A2
Free B27
heavy chain
HLA-class II
(DR, DQ, DP)
presenting
HLA-B27 peptide
HLA-B27 homodimers
at cell surface
HLA-B27:B2microglobulin:peptide
trimolecular complex
A2
Endoplasmic reticulum
B
B27 misfolding,
homodimerization
BiP
BiP
BiP
BiP
ERAD
UPR
B1
Golgi
A1
HLA-B27:B2M:peptide trimolecular
complex transported to the cell
surface via the golgi apparatus
B27 folding, assembly and
loading of peptide
Tapasin
B27 heavy chain (HC)
B27 HC folding
B2 microglobulin
(B2m) loading
peptide loading
Calnexin
A
Calreticulum
BiP
B2m
Macrophage
Ribosome
TAP 1,2
proteolytic degradation
within proteasome
peptide fragments
Viral, bacterial or
tumor protein

FIG 57.2 After transcription of the human leukocyte antigen (HLA)-B27 heavy chain on ribosomes, it is inserted into the endoplasmic reticulum (ER), glycosylated, and two pathways ensue. (A) The B27 heavy chain is retained through binding with calnexin and ERp57, folded into its tertiary structure, and bound to β_2 microglobulin. After that calnexin releases the complex and it is associated with calreticulum, which, in turn, chaperones the formation of the peptide loading on to the complex of heavy chain, β_2 microglobulin and antigenic peptide, via the TAP proteins and tapasin. The antigenic peptide is derived from intracellular proteins from viruses, bacteria, tumors, and so on, that have been degraded in proteasomes, and then the peptides are trimmed for optimal length for peptide loading by endoplasmic reticulum-associated aminopeptidases (ERAP1 and ERAP2). Then the trimolecular peptide complex (HLA-B27 heavy chain, β_2 microglobulin and peptide) travels through the Golgi apparatus (A1) to the cell surface, where the antigenic peptide is presented either to the α: β T-cell receptor on CD8 T lymphocytes or to the killer immunoglobulin (KIR) receptor on natural killer (NK) cells (A2); or (B) the HLA-B27 heavy chain misfolds in the ER, forming B27 homodimers and other misfoldings which are bound to the ER chaperone BiP. Then, they either (B1) accumulate there, causing either ER-associated degradation (ERAD) or a proinflammatory ER unfolded protein response (UPR); or (B2) the B27 homodimers migrate to the cell surface, where they either become antigenic themselves or present peptide to receptors on T cells and NK cells.

linkage disequilibrium found within the MHC, and many of the associations may reflect linkage to B27. These include the epithelial "stress" marker MICA, located adjacent to HLA-B27, which acts as a ligand for cells expressing a common activator NK receptor (NKG2D), as well as the tumor necrosis factor (*TNF*), heat shock protein (*HSP*)*-70*, *LMP-2* and *LMP-7*, *HLA-DRB1*01*, and *DRB1*04* alleles.[13] In addition, *HLA-DRB1*08* has been implicated both in susceptibility to uveitis in the setting of AS and to juvenile-onset AS.[1,13] In psoriasis, the primary MHC association is with HLA-Cw6 (*HLA-C*06:02*). Both *HLA-DRB1*04* and **07* alleles have been implicated in PsA. There is not a significant MHC association with IBD *per se*, although *HLA-DRB1*01:03* has been associated with enteropathic peripheral arthritis.[13]

Non-MHC Genes in Susceptibility to Spondyloarthritis

Recent genome-wide association studies (GWAS) in AS from the United Kingdom and North America[15,17] have implicated up to 70 genes in AS susceptibility (see Table 57.3). These genes fall into the following functional networks:

1. Genes involved in IL-17–mediated immunity (*IL23R*, *TYK2*, *IL6R*, *IL7R?*, *IL27*, *IL1R2/IL1R1*, *IL12B*, *JAK2*, *RORC*, *PTGER4*) (*Note:* Gene symbols are explained in the legend to Table 57.3.)

2. Genes involved in CD8 T-cell function (*RUNX3, EOMES, IL7R, ITGAL*)
3. Genes involved in peptide processing and presentation (*HLA-B, ERAP1, UBE2L3, NPEPPS, ERN1, ASAP2*)
4. Genes involved in microbial sensing (*CARD9, NOS2, NOD2, IRGM*)
5. Genes involved in nuclear factor (NF)-κB (NF-κB) activation (*TLR4, NOD2, NOTCH1, TNFAIP3*)
6. Others (*ZMIZ1, FCGR2A, KIF21B, SH2B3, TNFRSF1A, GPR65, SULT1A1, GPR35, BACH2, ICOSLG, NKX2-3*)

The most consistent (and best replicated) association is with ERAP1, which acts as a molecular ruler in the ER in trimming peptides processed in proteasomes to optimal length of nine amino acids for MHC class I binding and presentation. HLA-B27–negative patients with AS lack an association with ERAP1,[15] although an association with ERAP2 is seen.

A number of genes whose products are operative in the Th17 pathway have also been associated with AS susceptibility: the IL-23 receptor (*IL23R*), which pairs with the *IL12RB1* gene product to confer IL-23 (but not IL-12) responsiveness on cells expressing both subunits; prostaglandin E receptor 4 (*PTGER4*), which stimulates dendritic cell (DC) production of IL-23 and, in turn, Th17 expansion and is overexpressed in SpA synovium; signal transducer and activator of transcription 3 (*STAT3*), a key regulatory factor in Th17 responses; and IL-12β (*IL12B*), which encodes the IL12p40 protein, a component of both IL-12 and IL-23.

GWAS in psoriasis and PsA have identified over 60 genes implicated in disease susceptibility outside the MHC.[18] These include genes in common with those for AS (*IL23R, IL12B, CDKAL1,* and *ERAP1,* the last showing interaction with HLA-Cw6 similar to that seen with HLA-B27). Detailing all the non-MHC genes associated with psoriasis is outside the scope of this chapter. Non-MHC genes implicated in PsA susceptibility include *IL-1A,* as well as the *IL-2/IL-21* and the *IL-4/IL-13* gene complexes (both important in Th2-associated autoimmunity).

The first gene to be implicated in IBD susceptibility was *NOD2/CARD15,* which accounts for about 20% of CD susceptibility and whose protein product serves as a receptor for bacterial products in monocytes that transduces signals leading to NF-κB activation. A number of GWAS conducted since then have implicated over 130 genes in IBD susceptibility,[17] including 113 for CD and 100 for UC, and at least 83 shared genes between CD and UC. The autophagy 16-like 1 gene (*ATG16L1*), which is highly expressed in intestinal cell lines, is linked to CD susceptibility. DCs from patients with CD with susceptibility variants in *NOD2* or *ATG16L1* are deficient in autophagy induction, suggesting that these genes influence bacterial degradation and interact with the MHC class II antigen-presenting machinery. What is especially striking is the number of genes shared in common with the various types of SpA, suggesting a common pathogenesis.

Genes and Severity of SpA

Disease severity in AS also has a hereditary component. By defining severity on the basis of radiographic involvement, two genes have been identified in a large candidate gene study.[19] One was the gene for cyclooxygenase I (COX-I), the target of nonsteroidal antiinflammatory drugs (NSAIDs), the other for the receptor activator of NF-κB (RANK). No evidence was found for an MHC contribution to radiographic severity.

Infection

A role for triggering infections has been better documented in SpA than in most other rheumatic diseases. In developed countries, the most frequent type of ReA occurs following urogenital infections with *Chlamydia trachomatis* (endemic ReA). Postdysenteric ReA, more commonly encountered in "less technologically advanced countries," occurs after various *Shigella* and *Salmonella* (especially *S. typhimurium* and *S. enteriditis*), *Campylobacter jejuni* and *C. fetus,* and, in Europe, *Yersinia enterocolitica* species. Microorganisms implicated in ReA share common biological features: *(i)* They can invade mucosal surfaces and replicate intracellularly; and *(ii)* they contain lipopolysaccharide (LPS) in their outer membrane. Of particular note, antigens from *Salmonella, Yersinia,* and *Chlamydia* have been found in the synovial tissues and fluids of patients with ReA, often many years after the initial infection. Although only bacterial fragments of the enteric pathogens have been found, evidence for viable *C. trachomatis* and perhaps *C. pneumoniae* have been demonstrated in several studies. *Chlamydia* and other organisms have also been reported in the joints of healthy individuals, and this has led to questioning of the pathogenic significance of these findings. Other data, however, support the likelihood that bacterial persistence plays an important role in ReA, including the finding of specific IgA antibodies and synovial T-cell proliferation to the initiating infectious agent.

The Gut and Spondyloarthritis

In studies from Belgium and from Scandinavia, up to 50% of patients with AS have microscopic ileal inflammation seen on ileocolonoscopy. Moreover, two-thirds of patients with undifferentiated SpA have histological gut inflammation. Gut inflammation in AS appears to be immunologically related to that seen in CD. These observations have raised the speculation that the inciting event in the SpA may be a breakdown of the gut–blood barrier to intestinal bacteria, although such has yet to be proven. It has been established that patients with AS and their relatives have increased intestinal permeability compared with healthy controls.

The bacterial population inhabiting human intestines, referred to as the gut microbiota (Chapter 14), is a vast microbial community. The number of bacteria in the human gut outnumbers the cells in the body 10-fold, and they contain 100 times more genes than in the entire body.[20] In fact, there is a beneficial relationship between the gut microbiome and the human system. The human gut provides nutrients and the microbiome provides metabolic and physiological capacities, which possibly affect the education of the adaptive immune system. Many studies, from early antibody and fecal carriage studies suggesting a role for gut *Klebsiella pneumonia* in the pathogenesis of SpA (which has not been confirmed by modern sequencing studies)[20] to more recent findings showing an interplay between host genetics (*e.g.,* HLA-B27) and intestinal bacterial composition and others suggesting a link between intestinal dysbiosis and SpA, are suggesting that upsetting the homeostasis between the gut microbiome and the host immune systems.[20] This is clearly an area where more work needs to be done.

PATHOLOGY OF SPA

One of the biggest problems with studies of the synovium in SpA and PsA is that most lesions are examined late in the course of disease (*i.e.*, in the hips), and this only at joint replacement. Few data exist from early disease, and the difficulty with tissue access further complicates this.[21,22] For the most part, the synovium in SpA resembles that of rheumatoid arthritis (RA), with some notable differences. The synovium in SpA displays a tortuous vascular morphology compared with the rheumatoid synovium, which is linear and has diminished lymphoid aggregates. This may be caused by vascular endothelial growth factor (VEGF) and the angiogenic growth factor Ang2, the messenger RNA (mRNA) of which have been observed at higher levels in the synovium in PsA compared with RA. VEGF is particularly interesting because it can synergize with RANK ligand (RANKL) to induce bone resorption and also synergize with bone morphogenetic proteins to trigger bone formation, both processes typical of the altered bone remodeling seen in PsA and SpA.[21,22]

Increased production of the scavenger receptor CD163 by macrophages in both the lining and sublining layers is seen in SpA compared with RA.[20] Local production of soluble CD163 inhibits synovial T-cell activation, and levels of synovial CD163 fall with effective treatment. Increased expression of Toll-like receptors 2 and 4 (TLR2, TLR4) has been shown in SpA on CD163$^+$ peripheral blood mononuclear cells in patients with synovitis, which decreases with TNF-α blockade. This leads to the speculation that SpA represents an exaggerated inflammatory response of the innate immune system in genetically susceptible patients.[22]

Osteoclasts also appear to have a role and have also been observed at the bone–pannus junction in PsA. In addition, CD14$^+$ monocytes that are committed to becoming osteoclasts or osteoclast precursors are increased in the circulation of patients with PsA compared with healthy controls and decline rapidly following treatment with TNF antagonists. The clinical improvement is accompanied by an MRI-defined reduction of bone marrow edema.

Even fewer data exist on enthesitis (the enthesium being the insertion of tendons, ligaments, joint capsules, or fascia into bone). Pathological examination of enthesitis in AS demonstrates local inflammation, fibrosis, erosion, and ossification. Immunohistochemical staining for phosphorylated SMAD1/5 in entheseal biopsies of patients with SpA reveals active bone morphogenetic protein signaling.[21]

The pathology of psoriasis consists of an inflammatory cell infiltration in the dermis, with localized increased cytokine production and hyperproliferation of keratinocytes (Chapter 64). CD4 cells are prominent in the dermis, CD8 in the epidermis; Langerhans cells function as antigen-presenting cells (APCs). The synovium is infiltrated with CD8 T cells but demonstrates less pronounced intimal lining layer hyperplasia and fewer synovial T cells. It is more vascular than the synovium in RA, contains numerous B cells and macrophages, and has upregulation of adhesion molecules, such as intercellular adhesion molecule (ICAM)-1 and E-selectin, and overexpression of proinflammatory cytokines, such as TNF-α, IL-1β, IL-6, and IL-18.[22]

CLINICAL FEATURES

Ankylosing Spondylitis

Musculoskeletal Symptoms

KEY CONCEPTS

Clinical Features of Inflammatory Back Pain

- Low-back pain that is present every day for at least 3 months
- Age of onset <45 years
- Morning stiffness in the back lasting at least 30 minutes
- Pain that is relieved by exercise and worsened by rest
- Alternating buttock pain
- Relief with nonsteroidal antiinflammatory agents

The first symptoms of AS usually appear in adolescence or early adulthood and usually start before the age of 45 years. The hallmark of AS is the presence of inflammatory back pain,[5] a dull, persistent ache, usually in the buttocks or hips, that is worst in the early-morning hours (between 2 and 5 A.M.) and is associated with morning stiffness lasting >30 minutes (and sometimes several hours to all day). The pain is classically worsened by rest or recumbency and improves with activity. One important component of inflammatory back pain is the striking improvement that results from the use of NSAIDs (usually in high doses). Although the pain may be unilateral or intermittent at first (in fact, alternating buttock pain is a cardinal feature of the disease),[5] within a few months it usually becomes persistent and bilateral, and the lower lumbar area becomes stiff and painful.[4,5] Occasionally, the first symptom of AS comes from extraspinal sources, such as AAU, peripheral arthritis, or enthesitis, especially in patients with disease onset in childhood.

In patients with AS, the most commonly affected joints outside the spine are the hips (in up to 50% of patients),[22] with rapidly progressive destructive arthritis that necessitates joint arthroplasty at an early age. A characteristic radiographic finding is a fairly characteristic osteophytic collar that forms at the junction of the femoral head and the neck.[23] Peripheral arthritis other than in the hips and shoulders is uncommonly seen in patients with AS but, when present, is typical of that seen in other types of SpA, with an asymmetric oligoarthritis presenting predominantly in the lower extremities.

Chest pain, often pleuritic, can be seen in patients with AS as a result of involvement of the costovertebral and manubriosternal joints. This and progressive thoracic spinal involvement may result in fusion of the costovertebral joints, with loss of chest expansion and a mechanical restrictive ventilatory defect.

Enthesitis is a classic feature of AS and other SpA (Fig. 57.3). The most common (and most disabling) sites for enthesitis are in the foot, at the insertion of the Achilles tendon, and of the plantar fascia on to the calcaneus.[24]

Three physical measurements have been validated and recommended by an ASAS Working Group as useful for evaluating patients with AS specifically and with inflammatory back pain in general. The Schober test is measured as the increase with maximal forward spinal flexion with locked knees of a 10-cm segment marked on the patient's back with the inferior mark at the level of the posterosuperior iliac spines. The measured distance should increase from 10 cm to at least 13.5 cm in an adult. Chest wall expansion with inspiration is measured with a tape measure placed circumferentially around the chest wall

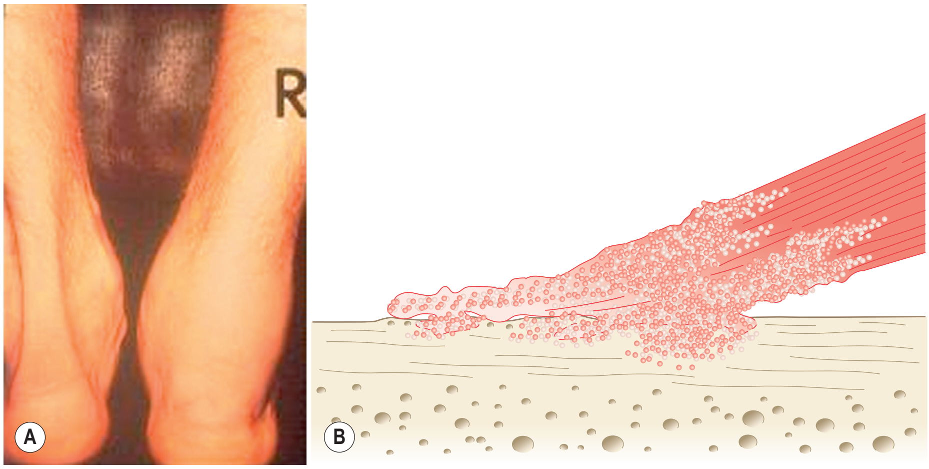


FIG 57.3 (A) Achilles tendinitis/enthesitis in a patient with reactive arthritis. (B) Schematic drawing of enthesitis, showing periosteal new bone formation, and subchondral bone inflammation and resorption.

at the fourth intercostal space. Normal chest expansion in an adult is >5 cm, although this may vary with age and gender. An occiput-to-wall distance of >2.5 cm is definitely abnormal. To measure the occiput-to-wall distance, the patient stands with the heels and buttocks touching the wall behind and with the knees straight. The patient is asked how far back he or she can move the head, still keeping the chin in the normal position. In the straight position, the distance between the posterior convexity of the occiput and the wall is measured to the nearest 0.1 cm. The better of two attempts is recorded. Anything >0 is regarded as abnormal.

Another commonly employed measurement is lateral bending. Here, the patient stands with the heels and back against the wall. There is no flexion of the knees or bending forward. The distance between the patient's middle fingertip and the floor is measured. The patient then bends sideways without bending the knees or lifting the heels. A second reading is taken, and the difference between the two is recorded. The best of two tries is recorded for the left and right sides. The mean of left and right gives the final result (in centimeters to the nearest 0.1 cm). Normal is >10 cm.

Extraarticular Manifestations

Uveitis. The anterior portion of the uvea consists of the *iris* and *ciliary body,* and the posterior portion is known as the *choroid.* Inflammation of the anterior uveal tract is known as *anterior uveitis* or *iritis* (Chapter 74). When the adjacent ciliary body is also inflamed, the process is known as *iridocyclitis.* AAU represents the typical uveitis found in SpA, occurring in up to 40% of patients with AS,[1] of whom approximately 90% are HLA-B27 positive (see Table 57.2).[1] Typically, AAU presents unilaterally with sudden onset, is self-limiting, and tends to be recurrent. Symptoms may include redness, pain, blurred vision, increased lacrimation, photophobia, and miosis. The diagnosis is characteristically confirmed by slit-lamp examination, which is also useful in monitoring response to treatment.

Prognosis is favorable in AAU, with resolution of symptoms within a few weeks. However, if treatment is delayed or inadequate, complications may occur; these include anterior synechiae (adherence of the iris to the cornea), posterior synechiae (adherence of the iris to the lens), which can lead to cataracts, and cystoid macular edema.[1] Rarely, increased ocular pressure is seen.[1] Macular edema has been shown to be the main factor that determines visual outcome in cases of uveitis.[1] Although AAU is the most common uveitis associated with AS, posterior uveitis has been reported and tends to be more severe, especially in those with coexistent IBD.[1]

Cardiac manifestations. The characteristic cardiac abnormalities in AS are aortitis, aortic regurgitation, and conduction abnormalities. Less commonly associated cardiac conditions include pericarditis, cardiomyopathy, and mitral valve disease. HLA-B27 is an important genetic risk factor for these cardiac conditions. Aortic regurgitation is well characterized and distinguished from aortic valvular dysfunction in other disorders. Three factors contribute to the development of incompetent aortic valves: dilatation of the aortic root, fibrotic thickening and downward retraction of the bases of the cusps, and inward rolling of the edges or margins of the cusps. Aortic regurgitation is present in 2–10% of patients with AS and increases in likelihood with greater disease duration.

Cardiac conduction abnormalities, including atrioventricular and intraventricular blocks, have been regarded as the most common cardiac complication in patients with AS. Complete heart block has been found in 1–9% of patients with AS. Electrophysiological studies show that the preferential level of block is in the atrioventricular node itself, which is in contrast to most cases of acquired complete heart block, where 80% are within or below the bundle of His. Rare complications include myocardial involvement, mitral regurgitation, and pericarditis.

Pulmonary manifestations. The incidence of pleuropulmonary involvement in AS is estimated to be 1%. The most frequently recognized manifestations are upper-lobe fibrosis, mycetoma formation, and pleural thickening. Fusion of the costovertebral joints caused by inflammation and ankylosis of the thoracic spine may lead to restrictive ventilatory impairment on pulmonary function testing. The upper-lobe fibrosis tends to be progressive. Another common finding is the presence of bilateral symmetric apical pleural thickening. Several recent studies have demonstrated that high-resolution computed tomography (HRCT) is more sensitive than chest radiography in detecting the presence of pulmonary abnormalities in AS, suggesting that pulmonary involvement in AS is more common than once thought. The clinical implications of these observations remain unclear as lung involvement in AS is usually asymptomatic.

Renal manifestations. Renal involvement in AS, although uncommon, can include secondary renal amyloidosis (AA type), NSAID nephropathy, and glomerulonephritis.

Osteoporosis. Measuring bone mineral density in patients with spondylitis is complicated by false increases in spinal density from dense syndesmophyte formation, leading some to recommend quantitative CT over standard dual-energy X-ray absorptiometry (DEXA) for bone mineral density measurements. Nevertheless, up to half the patients with long-standing AS have been reported as having osteopenia or osteoporosis,[25] which has been attributed to the impact of inflammation on bone remodeling as a result of aberrant activation of bone morphogenic protein and Wnt signaling. This may be further worsened by treatment factors and decreased mobility or physical activity, in addition to osteoclast/osteoblast imbalance.[25]

Spondylodiscitis and spinal fractures. An uncommon but well-recognized complication of AS is spondylodiscitis, a destructive discovertebral lesion also called *Andersson lesion.*[26] Typically, these lesions are confined to the thoracolumbar spine, sometimes with multiple-level involvement. Pain and tenderness localized to the affected disk are the most common presenting features of spondylodiscitis, although it can be asymptomatic and only detected on routine radiographic examination many years later. Spondylodiscitis usually occurs at an advanced stage of AS under the form of an erosive condition related to both mechanical factors and osteoporosis. However, early spondylodiscitis may occur as a result of the inflammatory process. Patients may or may not have a history of preceding trauma.

Even trivial falls can be catastrophic for AS patients, who are at risk for spinal fractures because of their spinal rigidity and osteoporosis. The estimated prevalence of vertebral fractures in AS varies from 4% to 18%. Fractures through the disk space, the weakest point in the ankylosed spine, are most common, with the cervical spine being the most frequently affected region, followed by the thoracolumbar junction, and may or may not be complicated by injury to the spinal cord, ranging from mild sensory loss to quadriplegia. Spontaneous atlantoaxial subluxation is also rarely seen.

Neurological manifestations. Neurological involvement in AS is most often related to spinal fracture, atlantoaxial subluxation, or cauda equina syndrome. The cauda equina syndrome in AS is characterized by a slow insidious progression and a high incidence of dural ectasia, although a rapid onset secondary to a traumatic event has been reported. It tends to be a late manifestation of AS, often when the disease is no longer active. The prevalence of neurological findings in cauda equina syndrome in AS is very high, presenting with a prodrome of sensory, motor, or reflex loss before the progression to sphincter disturbance. About half the patients have pain in the rectum or lower limbs that is presumably neurogenic in origin. Case reports have also been published about the occurrence of AS with a multiple sclerosis (MS)–like syndrome and transverse myelitis, although the association is not conclusive.

Fatigue and psychosocial manifestations. Fatigue is a common problem in AS and seems to be associated with more severe disease. Sleep disturbance has been reported to be as high as nearly 81% of female patients with AS and 50% of male patients. The disturbance is closely related to pain during the night characteristic of active disease. A high level of depressive symptoms has been reported in approximately one-third of patients with AS, with women reporting more depression than men. Pain was found to be a major determinant of depression for women but was of lesser importance for men.

AS in women. AS in women may not be as severe as it is in men and may present with isolated neck pain in the absence of typical back pain.[26] There tends to be a greater delay in the diagnosis of AS in women compared with that in men. Women tend to have less severe involvement of the spine, with peripheral joint involvement. A large review of the impact of AS on reproductive events on women concluded that AS did not adversely affect the ability to conceive, pregnancy outcome, or neonatal health.[27]

Reactive Arthritis

The classic triad of arthritis, urethritis, and conjunctivitis, representing what was formerly known as *Reiter syndrome,* is a presenting feature of only a minority of patients with ReA (comprising only a third of the cases in some series). In ReA, the clinical features are nowadays viewed more as a spectrum ranging from the classic triad to undifferentiated SpA. In fact, the manifestations vary among patients, depending on the genetic makeup, the triggering event, and the sequential immunological reaction.

Typically, the features start 1–4 weeks after a triggering event, frequently identified as an enteric or urogenital infection, but often the event passes unnoticed without any specific symptoms. The syndrome starts with constitutional symptoms, such as fatigue, malaise, and fever, and then is typically manifested by asymmetrical, additive lower-extremity oligoarticular inflammatory arthritis along with an array of different extraarticular features, including a sterile oligoarticular or monoarticular and asymmetrical arthritis of the lower extremities, especially knees, ankles, and, occasionally, hips. Upper-extremity involvement is encountered less commonly. Dactylitis occurs in the toes or fingers, resulting in the "sausage digits," which represent inflammation not only of the interphalangeal joints of the hands and feet but also of the surrounding soft tissue structures, including tendons and subcutaneous tissue.

Sacroiliitis and spondylitis are less common than peripheral arthritis, although inflammatory back pain does occur. Unilateral and bilateral sacroiliac involvement and even spondylitis occur, especially in those with chronic or long-standing disease. The most common sites for enthesitis are the Achilles tendon and plantar fascia insertions, although tenderness over the symphysis pubis, iliac crest, ischial tuberosity, greater trochanters, and thoracic cage ribs may also occur.

Mucocutaneous lesions may be difficult to distinguish from PsA, especially circinate balanitis and keratoderma blennorrhagica. Circinate balanitis is an ulcerative mucosal lesion over the glans or shaft of the penis that is demarcated by a serpiginous erythematous border. The lesion is usually painless and sterile unless a superimposed infection occurs. Keratoderma blennorrhagica is a painless desquamative psoriatic-like papulosquamous eruption and is sometimes referred to as *pustulosis palmoplantaris* and occurs on the palms and soles of the feet. Oral lesions have been described as shallow, painless ulcers or patches on the palate and tongue, or mucositis of the soft palate and uvula. Conjunctivitis and AAU also occur, as described in AS. Conjunctivitis may be unilateral or bilateral and is usually an early feature manifesting with irritation, erythema, and lacrimation. It is usually associated with a sterile discharge unless a superimposed infection occurs because of eye rubbing. It can be severe and occasionally progresses to episcleritis, scleritis, or keratitis. Other findings seen in AS, such as cardiac involvement, rarely occur. Although renal involvement is mainly described in the context of the urogenital triggering infectious process, sterile pyuria in conjunction with proteinuria and microscopic hematuria are sometimes encountered. Documented glomerulonephritis is rarely described.

Juvenile Spondyloarthritis

There are two clinical subsets of JSpA: *(i)* undifferentiated JSpA, which includes peripheral arthritis and enthesitis, primarily affects the lower limbs, may also present with sacroiliac tenderness and/or inflammatory spinal pain, and also includes isolated episodes of arthritis, enthesitis, tendinitis, dactylitis, and seronegative enthesopathy and arthropathy (SEA) syndrome[28]; *(ii)* differentiated JSpA (juvenile AS, PsA, IBD-related arthropathy) includes peripheral arthritis and enthesitis plus evidence of structural changes in juvenile AS (radiographic sacroiliitis, spinal disease, or tarsal ankylosis) and/or specific extraarticular symptoms (*e.g.,* psoriasis or IBD). SEA syndrome was originally referred to the combination of enthesitis and arthritis or arthralgia as an idiopathic disease or as part of a well-defined SpA.[27]

Psoriatic Arthritis

Skin involvement exhibits four clinical patterns.[29] The most common type is psoriasis vulgaris. Nearly as common is guttate psoriasis. The most severe type is the erythrodermic variety. Finally, pustular psoriasis is the type most closely associated with HLA-B27. Usually, the disease appears coincident with or after the onset of skin manifestations, although approximately 15–20% of patients will have preexisting arthritis. The joint disease likewise occurs in different subtypes, as defined by the Moll and Wright classification (Fig. 57.4), including oligoarticular, asymmetrical, polyarticular, symmetrical, distal interphalangeal (DIP)-predominant, spondylitis (sacroiliitis), arthritis mutilans, inflammation of DIP joints (often with nail involvement (≈80%)), dactylitis (“sausage digits”), and enthesitis.[26] Extraarticular features include nail pitting (which correlates best with DIP involvement) and uveitis (which occurs in some series as high as 33% but in most far less). Radiographically, large eccentric erosions are encountered.

Enteropathic Arthritis

The arthritis associated with IBD (enteropathic arthritis) is most commonly nondestructive and reversible. Two patterns have been recognized (Table 57.4). Type 1 is oligoarticular, involving the knees and ankles more than the upper extremities. It tends to resolve in <6 weeks. The second type has a polyarticular presentation, is more likely to involve the metacarpophalangeal (MCP) and proximal interphalangeal (PIP) joints than the lower extremities, and is more likely to have a chronic course. The symptoms of peripheral arthritis tend to coincide with activity of the bowel disease in UC but not in CD. Total colectomy is associated with remission of arthritis in half the patients. In contrast, axial

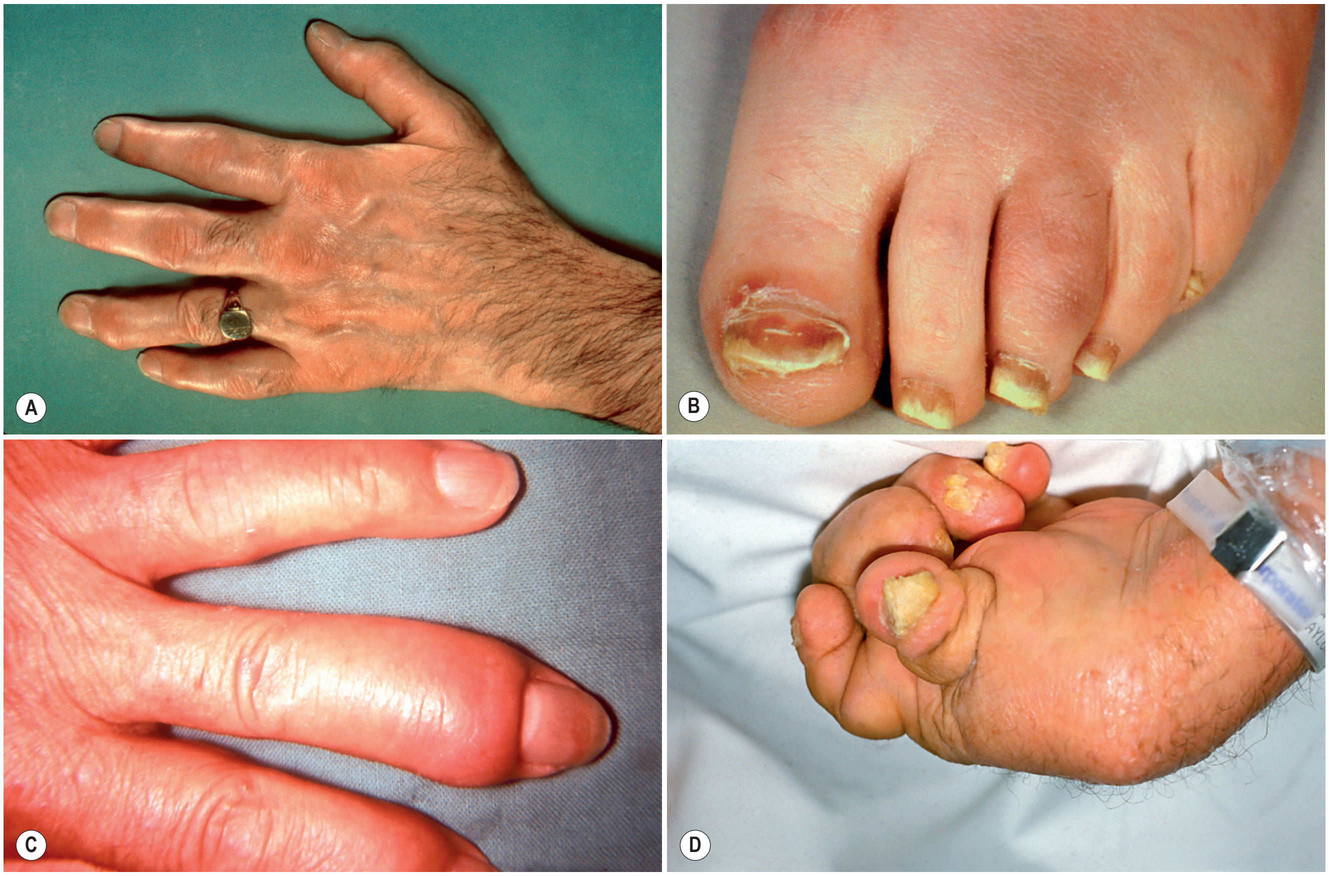


FIG 57.4 Patterns of psoriatic arthritis, showing (A) rheumatoid-like distribution; (B) sausage digits; (C) distal interphalangeal involvement; and (D) psoriatic arthritis mutilans.

TABLE 57.4 Extraintestinal Manifestations of Inflammatory Bowel Disease

Feature	Crohn Disease (%)	Ulcerative Colitis (%)
Peripheral arthritis	15	10
Axial arthritis	15–20	10–15
Septic arthritis	Rare	No association
Skin	<9	<1
Aphthous ulcers	Rare	8
Nephrolithiasis	<15	No association
Liver disease	3–5	7
Uveitis	13	4
Vasculitis	Takayasu arteritis	<5
Clubbing of fingers	4–13	1–5

TABLE 57.5 Frequencies of Different Symptoms and Signs in Patients With Undifferentiated Spondyloarthritis

Feature	Percent (%)
Demographic	
Males	62–88
Mean age at onset (years)	16–23
Clinical	
Low-back pain	52–80
Peripheral arthritis	60–100
Polyarthritis	40
Enthesopathy	56
Heel pain	20–28
Mucocutaneous involvement	16
Conjunctivitis	33
Genitourinary disease	28
Inflammatory bowel disease	4
Cardiac abnormalities	8
Laboratory	
Elevated erythrocyte sedimentation rate	19–30
Human leukocyte antigen (HLA)-B27 positive	80–84
Radiographic	
Sacroiliitis	16–30
Spinal radiographic changes	20

Adapted from Chen CH, Lin KC, Yu DT, et al. Serum matrix metalloproteinases and tissue inhibitors of metalloproteinases in ankylosing spondylitis: MMP-3 is a reproducibly sensitive and specific biomarker of disease activity. Rheumatol (Oxf) 2006; 45: 414–420.Table 56.6 Measurements of disease outcome in spondyloarthritis.

involvement may precede the development of IBD, has no gender predilection, and resembles the development of AS. The axial symptoms do not parallel activity of IBD. In addition to spondylitis, an isolated sacroiliitis occurs that is often asymmetric and not associated with HLA-B27.

Mucocutaneous complications of IBD include erythema nodosum, which occurs in fewer than 10% of those with CD and is rare in UC; pyoderma gangrenosum, seen in slightly over 1% of those with CD and rarely occurring in those with UC; and, rarely, erythema multiforme. Painful aphthous ulcers occur in about 8% of those with UC and are rare in CD.

The uveitis with IBD that is bilateral, posterior, insidious in onset, and/or chronic in duration contrasts with the uveitis associated with other types of SpA, which is predominantly anterior, unilateral, sudden in onset, and limited in duration. Only 46% of patients with uveitis associated with IBD are HLA-B27 positive, as opposed to 89% of the patients with SpA. Episcleritis, scleritis, and glaucoma are more common among patients with IBD than in those with SpA.

Undifferentiated Spondyloarthritis

Patients who do not meet criteria or clinical features of the "classic" spondyloarthritides are regarded as having undifferentiated SpA. Generally, at presentation, about 40% of patients will be classified as having undifferentiated SpA,[30] and the frequency of HLA-B27 reaches around 80% in Caucasians (Table 57.5). Follow-up studies suggest that over time about one-third will go into remission and more than half will develop a "classic" SpA, usually AS.[30]

LABORATORY INVESTIGATIONS

The most useful investigations in SpA come from musculoskeletal imaging, but some laboratory tests are informative.

Data on the correlation of erythrocyte sedimentation rate (ESR) and cross-reactive protein (CRP) in the assessment of disease activity in SpA show ambiguous results, although most studies suggest that CRP performs better. A recent literature review on the validity of ESR and CRP in AS concluded that ESR and CRP do not comprehensively represent the disease process and thus do not have the same validity as in RA.[13,30] Generally, it is felt that patients with peripheral joint involvement or with IBD more often have elevated ESR and CRP than patients with axial disease.[13] However, the presence of clinically active AS is not excluded by normal ESR and/or CRP levels. Synovial fluid does not differ in appearance or cytology from that of any inflammatory joint disease.

DIAGNOSIS

In most cases, SpA is largely diagnosed, or at least initially suspected, on clinical grounds. Current criteria for AS demand that the patient have radiographic sacroiliitis (at least grade II bilaterally or grade III unilaterally) in conjunction with clinical signs of inflammatory back pain and limitation of spinal mobility. However, given that up to 10 years can pass from the onset of inflammatory back pain and the development of radiographic sacroiliitis, many of those with inflammatory back pain might not have radiographic evidence of sacroiliitis. With the development of effective treatments (*i.e.,* anti-TNF blockers), criteria have been developed for earlier diagnosis of axial SpA that take into account recent advances in MRI scanning as well as the added benefit provided in HLA-B27 testing.[5,6]

Measures of SpA Activity and Severity

In the past few years, outcome measures have been developed and validated to quantitate disease activity and severity; these are summarized in Table 57.6. These instruments are extensively validated and easy to administer in clinical practice and have been shown to perform well in clinical trials. The Ankylosing Spondylitis Disease Activity Scale (ASDAS) has recently been developed to gauge disease activity.

TABLE 57.6 Measurements of Disease Outcome in Spondyloarthritis

Ankylosing Spondylitis

1) Disease activity
    (a) Bath Ankylosing Spondylitis Disease Activity Index (BASDAI)
    (b) Ankylosing Spondylitis Disease Activity Score (ASDAS)
    (c) Patient and Physician Global Assessments
2) Function
    (a) Bath Ankylosing Spondylitis Functional Index (BASFI)
    (b) Dougados Functional Index
3) Quality of Life
    (a) SF-36
    (b) Ankylosing Spondylitis Quality of Life Index (ASQOL)
    (c) ASAS Health Index
4) Metrometry
    (a) Schober test (lumbar flexion)
    (b) Chest expansion
    (c) Occiput-to-wall distance
    (d) Lateral bending
    (e) Bath Ankylosing Spondylitis Metrology Index (BASMI) lateral ending
5) Imaging
    (a) Standard radiography
    (b) Computed tomography
    (c) Magnetic resonance imaging
6) Assessment in Ankylosing Spondylitis (ASAS) 20
    (a) An improvement of 20% and absolute improvement of 10 units on a 0–100 scale in three of the following four domains:
        i. Patient global assessment (by visual analogue scale [VAS] global assessment)
        ii. Pain assessment (the average of VAS total and nocturnal pain scores)
        iii. Function (represented by BASFI)
        iv. Inflammation (the average of the BASDAI's last two VAS concerning morning stiffness, intensity, and duration)
        v. Absence of deterioration in the potential remaining domain (deterioration is defined as 20% worsening)

Psoriatic Arthritis

1) Arthritis
    (a) ACR response criteria
    (b) Psoriatic Arthritis Response Criteria (PsARC)
    (c) Ritchie Articular Index
2) Skin response
    (a) Psoriasis Area and Severity Index (PASI)
    (b) Target lesion score
    (c) Static global assessment
3) Quality of life (HAQ, SF-36, DLQI)
4) Radiographic

CLINICAL RELEVANCE

Utility of Human Leukocyte Antigen (HLA)-B27 Testing in the Evaluation of Inflammatory Back Pain and Spondyloarthritis

- Not indicated where the diagnosis is unquestionable, as it has little value in prognosis.
- Although patients with spondyloarthritis of African and Middle Eastern ancestry are more likely to be HLA-B27 negative, the finding of HLA-B27 in these patients has higher predictive value.
- Most useful in patients with either inflammatory back pain without radiographic changes or with other features of spondyloarthritis (unexplained lower extremity arthritis in a young adult, uveitis, etc.).
- If serological testing is used to ascertain HLA-B27, it must be ensured that the blood sample arrives in the laboratory within 24 hours of being drawn (false negatives will be caused by cell death).

Radiographic Imaging of Spondyloarthritis

Axial Spondyloarthritis

The bottom line in the diagnosis of AS is the demonstration of radiographic sacroiliitis (Fig. 57.5).[4] Two outcome instruments have been introduced in the assessment of disease damage and progression in AS: the Bath Ankylosing Spondylitis Radiographic Index (BASRI) and the modified Stokes Ankylosing Spondylitis Scoring System (mSASSS).[31] As a rule, these instruments have a low sensitivity to change (7.5% over 2 years), have been validated in long-duration disease only, and their predictive effect for disease activity is not yet ascertained.

One problem with radiographic imaging is the average decade-long interval from the onset of inflammatory back pain to the appearance of radiographic sacroiliitis.[5] The introduction of MRI imaging of the spine and entheses has allowed not only correct anatomical description of spinal structures but also differentiation of AS-related and unrelated inflammatory spinal lesions earlier than is possible with standard radiography.[32] MRI of the sacroiliac joints and spine is currently the only imaging tool to localize and quantify spinal inflammation accurately (Fig. 57.6) and is being developed as a measure of disease activity and treatment response.

Psoriatic Arthritis

PsA has some characteristic radiographic manifestations, including asymmetrical involvement, involvement of the DIP joints, and the classic "pencil-in-cup" deformities. Also seen are periostitis, bony ankylosis, and bony erosions with new bone formation. Radiographic severity is quantitated by the modified Sharp scoring method used in RA.

The measures used in PsA to assess disease severity include the ACR response criteria, the Psoriatic Arthritis Response Criteria (PsARC), which entail improvement in at least two of the following four criteria: *(i)* physician; and *(ii)* patient global assessments (on 0–5 visual analogue scales); *(iii)* tender and swollen joint scores (>30% improvement), with improvement in at least one of these two joint scales; and *(iv)* no worsening in any criteria. The Ritchie Articular Index is also used. The Psoriasis Area and Severity Index (PASI) is used to assess the extent of skin involvement, as well as general measures, such as the target lesion score, and the static global assessment. The PASI is a composite index of skin disease severity, including an overall evaluation and quantitation of the extent of scaling, erythema, and induration weighted: *(i)* by severity; and *(ii)* by body surface area. A target lesion is a single lesion >2 cm in diameter that is evaluated over time by a dermatologist and is graded for size, elevation, erythema, and scaling.

DISEASE COURSE AND PROGNOSIS

Ankylosing Spondylitis

AS significantly impacts the lives of those affected. Recent data suggest that patients with AS are more likely to be work disabled or even not participate in the labor force compared with population controls, especially in older patients and in those with longer disease duration. Moreover, in the same study, patients with AS were more likely to have never married or to be divorced. Women with AS were less likely than expected to have had children.[32]

Although AS is a chronic condition that can frequently have an unpredictable course, some studies suggest that those with higher levels of disease activity early in the course of the disease are more likely to have active disease in the future (Fig 57.7).[33]

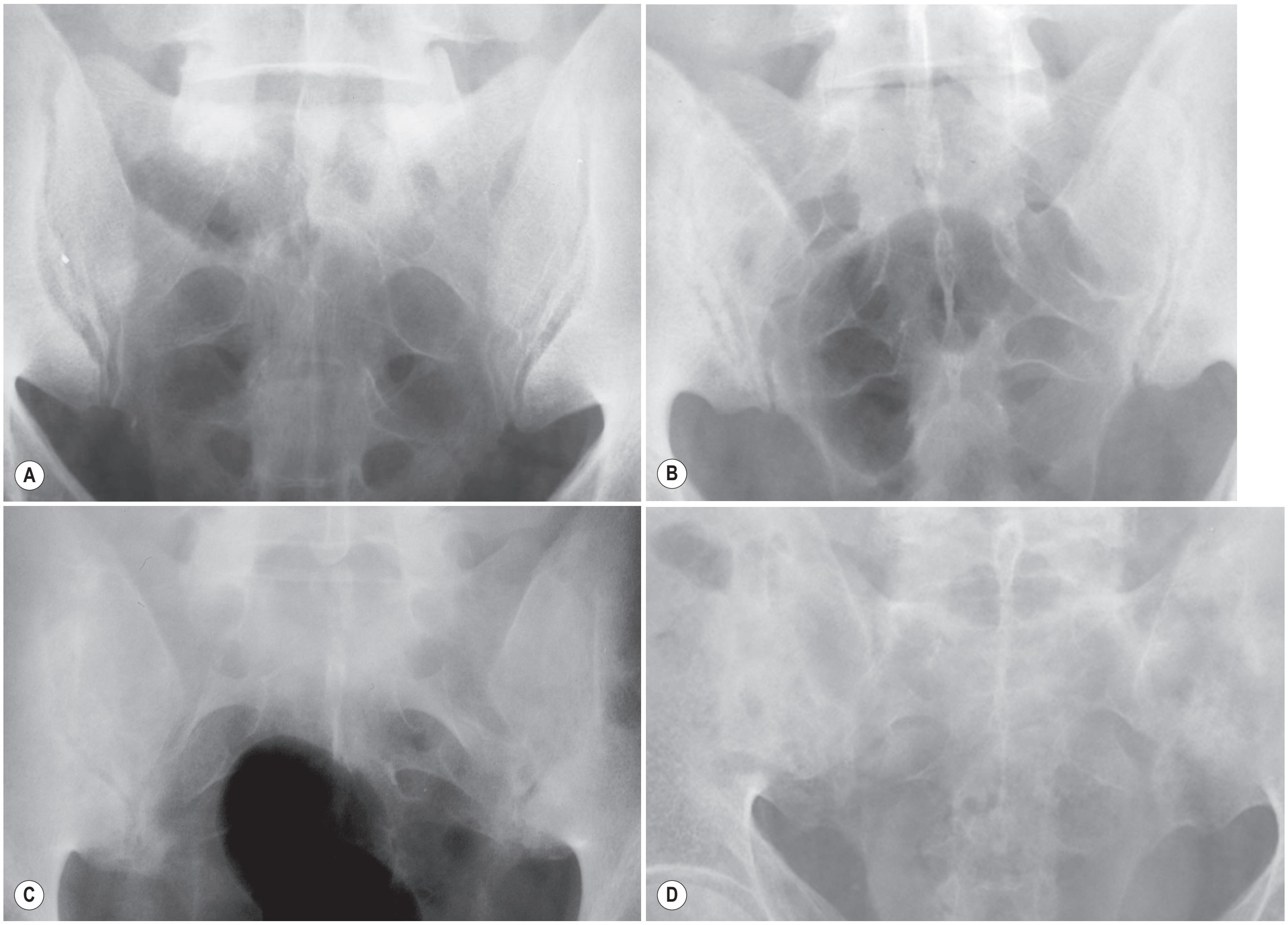


FIG 57.5 Grading of radiographic sacroiliitis, including (A) grade 0–1 (normal); (B) grade 2–3, with sclerosis and small erosions; (C) advanced grade 3, with joint space narrowing and large erosions; and (D) grade 4 (total sacroiliac fusion).

Hip involvement is a predictive factor for severe disease.[32] Other factors that may suggest severe disease and severe outcome include ESR >30 mm/hr; unresponsiveness to NSAIDs; limitation of lumbar spine; "sausage" digits; oligoarthritis; or onset at <16 years.[33] Longitudinal studies in patients with AS revealed that deformities and disability occurred during the first 10 years of disease and correlated significantly with the occurrence of peripheral arthritis, radiographic changes of AS in the spine, and the development of "bamboo" spine.[33]

Significant risk factors for work disability from several studies include older age, longer disease duration, lower level of education, reduced physical functioning, pain, and more physically demanding jobs. Patients with AS have an overall frequency of disability and economic costs similar to that of RA, with a threefold greater rate of disability than the general population, a higher rate of sick leave episodes (up to 50% more), and an overall 8% loss of productivity.[34] Moreover, a growing body of data has shown that patients with AS are at risk for early mortality as a result of cardiovascular disease.[35] However, the impact of newer agents, such as anti-TNF and anti–IL-17 drugs, on the natural history of this disease remains to be seen.

Reactive Arthritis

Early studies of outcome in ReA suggested a relatively poor prognosis. More recent studies, however, have found that, in general, the prognosis of ReA is fairly good.[36] Most cases appear to remit within 6 months of onset. Given the introduction of highly effective biological agents, such as anti-TNF blockers, it is likely that the long-term prognosis in ReA will improve even further.

Psoriatic Arthritis

Recent studies—although before the introduction of biological treatments—showed that the prognosis of PsA is worse than previously suggested.[29,37] In one large cohort, 40–57% of patients had deforming arthritis, 17% had[3] five deformed joints, and up to 19% of patients had disability. Overall, more rapidly progressive disease was associated with a greater number of actively inflamed joints, the early use of disease-modifying agents, and the presence of HLA-B27 and -B39. Patients with PsA have also been shown to have increased mortality, which is associated with a high ESR, high "advanced" medication use, and early radiographic damage.

Juvenile Spondyloarthritis

Although not extensively studied, the prognosis in JSpA is guarded.[28] Available data suggest that children with disease activity for >5 years are more likely to be disabled. In fact, the probability of remission was only 17% after 5 years of disease. Nearly 60% of children with JSpA have moderate to severe limitation after 10 years of disease. What is not clear is the extent to which the

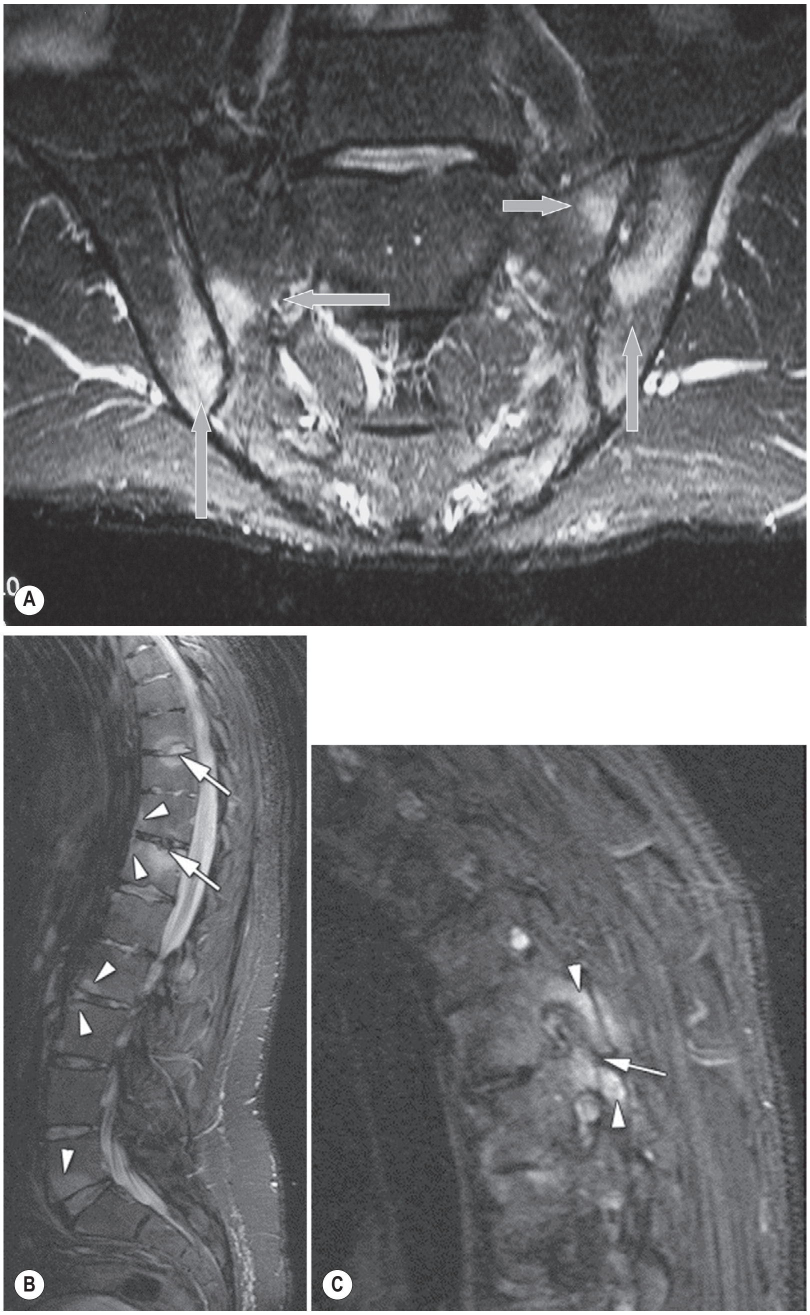


FIG 57.6 (A) Magnetic resonance imaging (MRI) of the sacroiliac joints, showing areas of marrow edema (indicated by arrows) on STIR sequences. (B) Lateral spine, showing enhancement of the insertion of the annulus fibrosis on the disk (arrowheads) and subchondral bone (arrows). (C) Involvement of the subchondral bone of the apophyseal joints.

outcome in juvenile-onset AS is different from that in adult-onset disease.

TREATMENT

Recent guidelines for treatment of SpA have been put forth by the American College of Rheumatology and the European League against Rheumatism based on current literature and clinical trials (Table 57.7).[38,39]

Patient Education and Physiotherapy

A great deal of educational information is available for patients (http://www.spondylitis.org and http://www.arthritis.org). Unsupervised recreational exercise improves pain and stiffness, and back exercise improves pain and function in patients with AS and other types of SpA, but these effects differ with disease duration. Health status is improved when patients perform recreational exercise at least 30 minutes per day and back exercises at least 5 days per week.

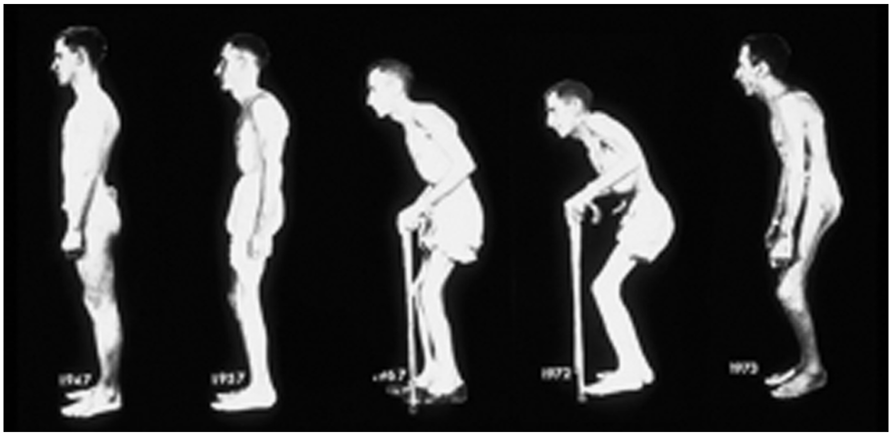


FIG 57.7 The "classic" course of ankylosing spondylitis, showing disease progression from shortly after disease onset in 1947 until just before the patient's death in 1973. The slight improvement between 1972 and 1973 was as a result of his having undergone total hip arthroplasties.

TABLE 57.7 Treatment of Spondyloarthritis

- Patient education
- Physiotherapy
- Medications
- Nonsteroidal antiinflammatory drugs
- Disease-modifying antirheumatic drugs
    - Sulfasalazine (especially for peripheral arthritis)
    - Methotrexate (especially for psoriatic arthritis, psoriasis)
    - Leflunomide
- Corticosteroids
    - Systemic
    - Intraarticular, intralesional
- Biological agents
    - Tumor necrosis factor blockers
    - Interleukin (IL)-17/-IL-23 blockers
- Treatment of osteoporosis
- Surgery
- Hip replacement
- Corrective spinal surgery

Medical Treatment

Nonsteroidal Antiinflammatory Drugs

NSAIDs remain the starting point of treatment, and many patients will attain satisfactory symptom control with these agents alone. There are no strong data to suggest the superiority of any specific NSAID in patients with SpA. NSAIDs when taken regularly (not on an as-needed basis) and at full antiinflammatory doses retard the radiographic progression of AS,[40] an observation that has recently been replicated.[41] COX-2 antagonists are recommended mainly for patients with proven peptic ulcer disease. Of concern is the association of the use of NSAIDs with flares of colitis, suggesting they should be used with care in this setting.

Disease-Modifying Antiinflammatory Drugs

Sulfasalazine. The efficacy of sulfasalazine in the treatment of peripheral joint involvement in AS and other SpA has been shown in several controlled trials, including two large multicenter studies in the United States and France.[42] It is not efficacious in axial disease. Coincident with improvement in peripheral arthritis is a fall in acute-phase reactants, such as the ESR and CRP.

THERAPEUTIC PRINCIPLES

Treatment Principles for Medical Management of Spondyloarthritis

- Patient education, regular exercise, smoking cessation, and physiotherapy should be initiated early in the disease course.
- Nonsteroidal antiinflammatory drugs (NSAIDs) remain the "first-line" treatment.
- Disease-modifying antirheumatic drugs (DMARDs: sulfasalazine, methotrexate) are used for peripheral arthritis.
- Intraarticular/intralesional corticosteroid injections are administered.
- Biological (anti–tumor necrosis factor [TNF], anti–interleukin-17 [IL-17] agents) for axial disease refractory to NSAIDs, peripheral arthritis refractory to DMARDs, and entheseal lesions refractory to NSAIDs.
- It is important to remember to treat coexistent/complicating conditions (inflammatory bowel disease [IBD], psoriasis, osteoporosis, premature atherosclerosis).

Other DMARDs. Although less well studied than sulfasalazine, methotrexate has been shown to be effective in some but not all studies of peripheral arthritis and psoriasis in patients with AS and other SpA. Its efficacy in treating axial SpA has not been established.

The use of leflunomide in patients with SpA has not been well examined. Limited data suggest that it may be useful in the treatment of peripheral joint involvement in SpA as well as PsA, although not for axial involvement. Apremilast, an oral phosphodiesterase 4 inhibitor, has been shown to be effective in psoriatic arthritis but not in AS.[43]

Corticosteroids. Although not well studied in patients with AS, many clinicians add low-dose glucocorticoids to the management of active SpA where NSAIDs or DMARDs fail to achieve a satisfactory response. On occasion, pulse steroids have also been utilized. Given the lack of controlled data as to their effectiveness, the side effects of long-term glucocorticoid therapy (including osteoporosis, a major cause of morbidity in AS patients, and possible worsening of psoriasis), and the emergence of more effective treatments, their use is not recommended unless more effective treatments are not available.

Intraarticular/Intralesional Corticosteroids

Intraarticular and peritendinous injections of depot steroid preparations are frequently employed by clinicians for symptomatic relief of local flare-ups, although they have not been extensively studied in controlled trials. Injecting around the Achilles tendon is generally not recommended because of the risk of tendon rupture.

Antibiotics

Early data suggested that a 3-month course of antibiotics in the acute phase after disease onset may have a beneficial effect on the course of ReA, specifically in those with *Chlamydia trachomatis*–triggered ReA, but not in other patients. Recent long-term follow-up data, however, suggest that tetracycline treatment did not change the natural history of the disease. In another recent study, however, a 6-month course of azithromycin and rifampin in the acute phase was found to have a beneficial effect on the long-term prognosis, although this has not been reproduced in all studies.[44] It is clear that early antibiotic therapy for chlamydial genital infections can prevent ReA, although the same is not true for enteric pathogens. For patients at risk for enteric infections and ReA, prophylactic antibiotics should be considered. There is no evidence that antibiotics have any place in the management of other SpA.

TNF-α Blockers

This category of medications has been shown to be effective in controlling inflammation and improving function in patients with AS (Table 57.8). The use of TNF blockers in the treatment of AAU is less clear, however. Recent data suggest they are useful, at least those comprising monoclonal antibodies (mAbs).[1] Currently five agents that have been approved for use by the Food and Drug Administration (FDA) are in clinical use for the treatment of AS in the United States, including infliximab, an infusion of a chimeric mAb to TNF-α (infliximab) at 5 mg/kg of infliximab every 6–8 weeks; etanercept, given 50 mg subcutaneously weekly; adalimumab, which is used at a dose of 40 mg administered subcutaneously every other week; golimumab, at a dose of 50 mg administered subcutaneously monthly; and certolizumab, at either 200 mg administered subcutaneously every other week[45] or 400 mg once a month. The onset of action is quite rapid, usually following the first infusion or injection. These agents are effective also in PsA[46] and axial SpA.[45,47] Improvement was seen not only clinically but also radiographically, with clearing of lesions suggestive of disease activity on MRI, and, in studies extending >4 years, radiographically. Although earlier studies following up patients for 2 years did not demonstrate that anti-TNF agents affect radiographic severity in AS, data from observational cohorts from North America and Europe spanning longer periods have shown they do, indeed, slow radiographic progression, especially in those with shorter disease duration treated for an extended period.[48]

TABLE 57.8 International Assessment in Ankylosing Spondylitis (ASAS) Consensus Statement for the Use of Anti–Tumor Necrosis Factor (TNF) Agents in Patients With Ankylosing Spondylitis (AS) and Axial Spondyloarthritis (ASpA)*

1. For the initiation of anti–TNF-α therapy:
    (a) A diagnosis of definitive AS by modified NY criteria or axial SpA by ASAS criteria
    (b) Presence of active disease for at least 4 weeks as defined by both a sustained Bath AS Disease Activity Index (BASDAI) of at least 4 *and* an expert opinion based on clinical features, acute-phase reactants, and imaging modalities
    (c) Presence of refractory disease defined by failure of at least two nonsteroidal antiinflammatory drugs during a single 4-week period, failure of intraarticular steroids if indicated, and failure of a disease-modifying antirheumatic drug, preferably sulfasalazine, in patients with peripheral arthritis
    (d) Application and implementation of the usual precautions and contraindications for biological therapy
2. For the monitoring of anti–TNF-α therapy: both the BASDAI and the ASAS core set for clinical practice should be followed regularly
3. For the discontinuation of anti-TNF-α therapy: in nonresponders, consideration should be made after 12 weeks' treatment. Response is defined as improvement of:
    (a) At least 50% or 2 units (on a 0–10 scale) of the BASDAI
    (b) Expert opinion that treatment should be continued

*van der Heijde D, Sieper J, Maksymowych WP, et al. 2010 Update of the international ASAS recommendations for the use of anti-TNF agents in patients with axial spondyloarthritis. Ann Rheum Dis 2011; 70 :905–8.

Interleukin-17 Blockers (Secukinumab)

Secukinumab is an anti–IL-17A mAb that has been shown to control the symptoms of active AS in one phase II trial and two phase III clinical trials[49] and has been approved by the FDA. The impact of this medication on other disease features, such as radiographic progression, remains to be determined.

Surgical Treatment of AS Complications

Because the hip is the joint most commonly involved in patients with AS, total hip arthroplasty is the most common surgical procedure.[24] Heterotopic new bone formation may be a potential problem.

Limited prevalence data suggest that patients with AS, even those with mild disease, are at increased risk for vertebral fracture, often resulting in neurological compromise.[25] In general, halo vest immobilization is recommended. Surgical intervention may be necessary when neurological impairment is seen. The fixed kyphotic deformities seen in patients with advanced AS are of considerable distress to patients and may result in substantial functional impairment. A small minority of patients with AS will seek surgical correction of their spinal deformities. In general, open, polysegmental, and closing wedge osteotomies are employed. Loss of correction is seen least commonly in closing wedge osteotomy. In a meta-analysis of the literature between 1945 and 1998, an average correction of 37°–40° was seen, with perioperative mortality of 4% as a result of pulmonary, cardiac, and intestinal problems.[50]

CONCLUSIONS AND RESEARCH OPPORTUNITIES

ON THE HORIZON

Research Opportunities in Spondyloarthritis (SpA)

- Improved understanding of pathogenetic mechanisms of SpA
- Elucidation of the roles of non–major histocompatibility complex (MHC) genes in SpA
- Definition of the link between gut inflammation and ankylosing spondylitis (AS) triggering
- Improved measures of treatment outcomes
- Advances in biological therapies of SpA

Progress has been made in the classification and epidemiology of SpA, particularly in the elucidation of the factors involved in SpA pathogenesis in recent years. It has become clear that HLA-B27, the primary genetic factor, functions in a variety of roles, including "classic" antigen presentation. GWAS have identified a vast number of susceptibility genes for SpA, including those in the antigen processing and Th17 pathways, some of which are shared by some or all SpA and others that are specific for a given disease. Their specific roles and how they interact are currently being elucidated, and with larger-scale studies and whole-genome sequencing, new candidates are being identified.

However, less progress has been made with regard to nongenetic factors. Despite what has been learned in ReA, no infectious trigger has been identified in AS (and, in fact, there might be no specific trigger). The link of gut inflammation to the triggering of AS is strongly suggested by data thus far, especially its possible link to the gut microbiome. This is clearly an area of further investigation.

Novel outcome measures have been developed that will help provide better care to patients, especially in the area of imaging. The advances in treatment have been most exciting, particularly the development of biological treatments, which hold promise for a better future for patients with these diseases.

REFERENCES

1. Monnet D, Breban M, Hudry C, et al. Ophthalmic findings and frequency of extraocular manifestations in patients with HLA-B27 uveitis: a study of 175 cases. Ophthalmology 2004;111:802–9.
2. Nurmohamed MT, van der Horst-Bruinsma I, Maksymowych WP. Cardiovascular and cerebrovascular diseases in ankylosing spondylitis: current insights. Curr Rheumatol Rep 2012;14:415–21.
3. Dougados M, van der Linden S, Juhlin R, et al. The European Spondylarthropathy Study Group preliminary criteria for the classification of spondylarthropathy. Arthritis Rheum 1991;34:1218–27.
4. van der Linden S, Valkenburg HA, Cats A. Evaluation of diagnostic criteria for ankylosing spondylitis. A proposal for modification of the New York criteria. Arthritis Rheum 1984;27:361–8.
5. Taylor W, Gladman D, Helliwell P, et al. Classification criteria for psoriatic arthritis: development of new criteria from a large international study. Arthritis Rheum 2006;54:2665–73.
6. Rudwaleit M, van der Heijde D, Landewe R, et al. The development of Assessment of SpondyloArthritis International Society classification criteria for axial spondyloarthritis (part II): validation and final selection. Ann Rheum Dis 2009;68:777–83.
7. Rudwaleit M, van der Heijde D, Landewe R, et al. The Assessment of SpondyloArthritis International Society classification criteria for peripheral spondyloarthritis and for spondyloarthritis in general. Ann Rheum Dis 2011;70:25–31.
8. Stolwijk C, van Onna M, Boonen A, et al. The global prevalence of spondyloarthritis: a systematic review and meta-regression analysis. Arthritis Care Res (Hoboken) 2016;68:1320–31.
9. Thom N, Ritchlin CT, Zhang X, et al. Prevalence of chronic axial pain, inflammatory back pain and spondyloarthritis in diagnosed psoriasis. Arthritis Care Res (Hoboken) 2015;67:829–35.
10. Loftus EV. Clinical epidemiology of inflammatory bowel disease: incidence, prevalence, and environmental influences. Gastroenterology 2004;126:1504–17.
11. Brown MA, Kennedy LG, MacGregor AJ, et al. Susceptibility to ankylosing spondylitis in twins: the role of genes, HLA, and the environment. Arthritis Rheum 1997;40:1823–8.
12. Robinson PC, Claushuis TA, Cortes A, et al. Genetic dissection of acute anterior uveitis reveals similarities and differences in associations observed with ankylosing spondylitis. Arthritis Rheumatol 2015;67:140–51.
13. Reveille JD. The MHC and Ankylosing Spondylitis. Clin Rheum 2014;33:749–57.
14. Reveille JD, Hirsch R, Dillon CF, et al. The prevalence of HLA-B27 in the United States: data from the U.S. National Health and Nutrition Examination Survey, 2009. Arthritis Rheum 2012;64:1407–11.
15. International Genetics of Ankylosing Spondylitis Consortium (IGAS), Cortes A, Hadler J, et al. Identification of multiple risk variants for ankylosing spondylitis through high-density genotyping of immune-related loci. Nat Genet 2013;45:730–8.
16. Delay ML, Turner MJ, Klenk EI, et al. HLA-B27 misfolding and the unfolded protein response augment interleukin-23 production and are associated with Th17 activation in transgenic rats. Arthritis Rheum 2009;60:2633–43.
17. Ellinghaus D, Jostins L, Spain SL, et al. Analysis of five chronic inflammatory diseases identifies 27 new associations and highlights disease-specific patterns at shared loci. Nat Genet 2016;48:510–18.
18. Genetic Analysis of Psoriasis Consortium and The Wellcome Trust Case Control Consortium 2. A genomewide association study identifies new psoriasis susceptibility loci and an interaction between HLA-C and ERAP1. Nat Genet 2010;42:985–90.
19. Cortes A, Maksymowych WP, Wordsworth BP, et al. Association study of genes related to bone formation and resorption and the extent of radiographic change in ankylosing spondylitis. Ann Rheum Dis 2015;74:1387–93.
20. Costello ME, Robinson PC, Benham H, et al. The intestinal microbiome in human disease and how it relates to arthritis and spondyloarthritis. Best Pract Res Clin Rheumatol 2015;29:202–12.
21. Melis L, Elewaut D. Progress in spondylarthritis. Immunopathogenesis of spondyloarthritis: which cells drive disease? Arthritis Res Ther 2009;11:233.
22. van Kuijk AWR, Tak PP. Synovitis in psoriatic arthritis: immunohistochemistry, comparisons with rheumatoid arthritis, and effects of therapy. Curr Rheumatol Rep 2011;13:353–9.
23. Sweeney S, Gupta R, Taylor G, et al. Total hip arthroplasty in ankylosing spondylitis: outcome in 340 patients. J Rheumatol 2001;28:1862–6.
24. McGonagle D, Gibbon W, Emery P. Classification of inflammatory arthritis by enthesitis. Lancet 1998;352:1137–40.
25. Carter S, Lories RJ. Osteoporosis: a paradox in ankylosing spondylitis. Curr Osteoporos Rep 2011;9:112–15.
26. Mitra D, Elvins DM, Speden DJ, et al. The prevalence of vertebral fractures in mild ankylosing spondylitis and their relationship to bone mineral density. Rheumatology (Oxford) 2000;39:85–9.
27. Lee W, Reveille JD, Davis JC, et al. Are there gender differences in the severity of ankylosing spondylitis: results from the PSOAS cohort. Ann Rheum Dis 2007;66:633–8.
28. Colbert R. Classification of juvenile spondyloarthritis: Enthesitis-related arthritis and beyond. Nat Rev Rheumatol 2010;6:477–85.
29. Ritchlin C. Psoriatic disease—from skin to bone. Nat Clin Pract Rheumatol 2007;3:698–706.
30. Zeidler H, Brandt J, Schnarr S. Undifferentiated spondyloarthritis. In: Weisman MH, Reveille JD, van der Heijde D, editors. Ankylosing spondylitis and the spondyloarthropathies. Philadelphia: Elsevier; 2006. p. 75–93.
31. van der Heijde D, Landewe S, van der Linden S. How should treatment effect on spinal radiographic progression be measured? Arthritis Rheum 2005;52:1979–85.
32. Maksymowych WP, Inman RD, Salonen D, et al. Spondyloarthritis Research Consortium of Canada magnetic resonance imaging index for assessment of sacroiliac joint inflammation in ankylosing spondylitis. Arthritis Rheum 2005;53:703–9.
33. Ward MM, Reveille JD, Learch TJ, et al. Impact of ankylosing spondylitis on work and family life: comparisons with the US population. Arthritis Rheum 2008;59:497–503.
34. Boonen A, Brinkhuizen T, Landewe R, et al. Impact of ankylosing spondylitis on sick leave, presenteeism and unpaid productivity, and estimation of the societal cost. Ann Rheum Dis 2010;69:1123–8.

35. Szabo SM, Levy AR, Rao SR, et al. Increased risk of cardiovascular and cerebrovascular diseases in individuals with ankylosing spondylitis: a population-based study. Arthritis Rheum 2011;63:3294–304.
36. Sairanen E, Paronen I, Mahonen H. Reiter's syndrome: a follow-up study. Acta Med Scand 1969;185:57–63.
37. Gladman DD, Chandran V. Review of clinical registries of psoriatic arthritis: lessons learned?: value for the future? Curr Rheumatol Rep 2011;13:346–52.
38. Ward MM, Deodhar A, Akl EA, et al. American College of Rheumatology/Spondylitis Association of America/Spondyloarthritis Research and Treatment Network 2015 Recommendations for the Treatment of Ankylosing Spondylitis and Nonradiographic Axial Spondyloarthritis. Arthritis Rheumatol 2016;68:282–98.
39. van den Berg R, Baraliakos X, Braun J, et al. First update of the current evidence for the management of ankylosing spondylitis with non-pharmacological treatment and non-biologic drugs: a systematic literature review for the ASAS/EULAR management recommendations in ankylosing spondylitis. Rheumatology (Oxford) 2012;51:1388–96.
40. Wanders A, van der Heijde D, Landewe R, et al. Nonsteroidal anti-inflammatory drugs reduce radiographic progression in patients with ankylosing spondylitis: a randomized clinical trial. Arthritis Rheum 2005;52:1756–61.
41. Poddubnyy D, Rudwaleit M, Haibel H, et al. Effect of non-steroidal anti-inflammatory drugs on radiographic spinal progression in patients with axial spondyloarthritis: results from the German Spondyloarthritis Inception Cohort. Ann Rheum Dis 2012;71:1616–22.
42. Clegg DO, Reda DJ, Abdellatif M. Comparison of sulfasalazine and placebo for the treatment of axial and peripheral articular manifestations of the seronegative spondylarthropathies: a Department of Veterans Affairs cooperative study. Arthritis Rheum 1999;42:2325–9.
43. Edwards CJ, Blanco FJ, Crowley J, et al. Apremilast, an oral phosphodiesterase 4 inhibitor, in patients with psoriatic arthritis and current skin involvement: a phase III, randomised, controlled trial (PALACE 3). Ann Rheum Dis 2016;75:1065–73.
44. Barber CE, Kim J, Inman RD, et al. Antibiotics for treatment of reactive arthritis: a systematic review and meta-analysis. J Rheumatol 2013;40:916–28.
45. Coates LC, Helliwell PS. Treating to target in psoriatic arthritis: how to implement in clinical practice. Ann Rheum Dis 2016;75:640–3.
46. Haibel H, Rudwaleit M, Listing J, et al. Efficacy of adalimumab in the treatment of axial spondylarthritis without radiographically defined sacroiliitis: results of a twelve-week randomized, double-blind, placebo-controlled trial followed by an open-label extension up to week fifty two. Arthritis Rheum 2008;58:1981–91.
47. Landewé R, Braun J, Deodhar A, et al. Efficacy of certolizumab pegol on signs and symptoms of axial spondyloarthritis including ankylosing spondylitis: 24-week results of a double-blind randomised placebo-controlled Phase 3 study. Ann Rheum Dis 2014;73:39–47.
48. Haroon N, Inman RD, Learch TJ, et al. The Impact of TNF-inhibitors on radiographic progression in Ankylosing Spondylitis. Arthritis Rheum 2013;65:2645–54.
49. Baeten D, Sieper J, Braun J, et al. Secukinumab, an Interleukin-17A Inhibitor, in Ankylosing Spondylitis. N Engl J Med 2015;373:2534–48.
50. Van Royen BJ, De Gast A. Lumbar osteotomy for correction of thoracolumbar kyphotic deformity in ankylosing spondylitis. A structured review of three methods of treatment. Ann Rheum Dis 1999;58:399–406.

Small- and Medium-Vessel Primary Vasculitis

Raashid Ahmed Luqmani, Ana Águeda, Lorraine O'Neill

Small- and medium-vessel vasculitides are characterized by inflammation of the blood vessel wall resulting in end organ failure or irreversible tissue damage and necrosis. In some cases, this is relatively trivial and may lead to minor inconvenience for the patient. However, in many forms of vasculitis, the consequences of rapid onset of ischemia and occlusion of blood vessels are devastating, leading to organ failure and death.

The distinction between small- and medium-vessel vasculitis entities is arbitrary, with three main patterns of vasculitis: small, medium, and large. Although there is merit in this classification, there are different patterns of involvement for patients with predominantly small-vessel involvement (characterized by capillaritis) compared with, but overlapping with, patients with medium-vessel involvement typified by small arteriolar inflammation. For example, in the kidney, small-vessel involvement leads to inflammation of glomeruli (glomerulonephritis); by contrast, medium-vessel inflammation of renal arterioles results in infarction of the kidney with tissue loss. Therefore patterns of disease are well recognized, with typical dominance of kidney, lung, and upper airway involvement in patients with granulomatosis with polyangiitis (GPA), previously termed Wegener's granulomatosis. In microscopic polyangiitis (MPA), there is lack of upper airway involvement but significant renal and lung involvement. Finally, in eosinophilic glomerulonephritis with polyangiitis (EGPA), previously termed Churg-Strauss syndrome, the pattern of clinical involvement is dominated by upper and lower airway disease combined with neurological (peripheral nerve) features. These three entities are grouped together by their association with the presence of antineutrophil cytoplasmic antibody (ANCA) in most, but not all, cases.

In patients with polyarteritis nodosa (PAN), one of the main forms of medium-vessel vasculitis, the most characteristic findings are bowel ischemia or infarction and peripheral neuropathy. In a childhood onset form of medium-vessel vasculitis (Kawasaki disease [KD]), the clinical features are diverse and include mucocutaneous inflammation and systemic upset with fever, and in 2–4% of cases, there is coronary artery dilatation and/or aneurysm development, which can potentially rupture, leading to fatal consequences.[1]

The diversity of different forms of vasculitis with overlapping features suggests that the underlying mechanisms are varied, but some pathways are likely to be shared. This is reflected in treatment approaches, which, with some exceptions, are often very similar across diseases.

EPIDEMIOLOGY

Despite improvement in our understanding of the epidemiology of the systemic vasculitides, patients are often not diagnosed as having vasculitis for extended periods. The discovery of antibodies to neutrophil cytoplasm (ANCA) and their association with small-vessel vasculitis has improved recognition.[12] Greater awareness of vasculitis may be a factor that explains an apparent increase in incidence from 1.5/million/year to 6.1/million per year.[2] The incidence of ANCA-associated vasculitides (AAV) is 10–20 new cases per million per annum.[3] In southern Europe, the number of cases of MPA is greater than those of GPA.

MPA is also more common than GPA in Japan. ANCA directed against myeloperoxidase is the predominant antibody detected in patients with AAV in Japan, whereas autoantibodies directed against proteinase 3 (PR3) are rarely seen in Japanese patients but are the most frequent ANCA antibody in patients in northern Europe.

The epidemiology of EPGA is less well understood compared with other forms of AAV. EGPA is characterized by an elevated eosinophil count in patients with late-onset asthma. Around 50% of cases have ANCA, usually directed against myeloperoxidase. EGPA is less common than either GPA or MPA, with an incidence of around 0.6/million per annum. There is a potential overlap condition called *hypereosinophilic syndrome* (HES), and it is not clear whether some cases of HES are really cases of EGPA, or vice versa. Indeed, if patients are ANCA negative, distinguishing the two conditions can be challenging.[4] Furthermore, because bronchospasm is a key feature of EGPA, it is possible that some cases of asthma, an extremely common condition, may, in fact, represent mild forms of EGPA. This has been brought more to light in cases of drug-induced EGPA, specifically in the setting of the use of montelukast, a leukotriene inhibitor, as a treatment for moderate to severe asthma. It has been suggested that these patients probably had underlying EGPA, which had previously been suppressed with systemic glucocorticoids, but when these were withdrawn, features of EGPA became more apparent.

AAV typically affects older individuals in their 60s or 70s but can occur at any age. Most patients survive their initial illness as a result of effective immunotherapy, and therefore the prevalence of these diseases is growing. In southern Sweden, estimates of prevalence are 160 per million (95% confidence interval [CI] 114–206) for GPA, 94 (58–129) for MPA, 31 (11–52) for PAN,

TABLE 58.1 **Diagnosis, Incidence and Prevalence in KD, PAN, AAV and Leukocytoclastic Vasculitis**

Diagnosis	Incidence	Prevalence	References/ Reviews
KD	For children <5 years old: 2431/million (Japan) 1131 per million (Korea) 690 per million (Taiwan) 36–185/million (Italy)	Not applicable	Makino N, et al., 2015[96]
PAN	3.6/million adults	2.6–14/million adults	Nesher G, et al., 2016[5] Mohammad AJ, et al., 2007[4]
ANCA-associated vasculitides	9.5–16/million/year (Germany) 22.6/million (Japan) 21.8/million (UK)	149/million	Reinhold-Keller E, et al., 2005[10] Herlyn K, et al., 2014[11] Fujimoto S, et al., 2011[94]
Leukocytoclastic vasculitis	45/million (equal male and female; increased incidence with age)	No data	Arora A, et al., 2014[7]

AAV, ANCA-associated vasculitis; *ANCA,* antineutrophil cytoplasmic antibody; *KD,* Kawasaki disease; *PAN,* polyarteritis nodosa.

and 14 (0.3-27) for EGPA.[6] However, these are based on a relatively small population size and are higher than those reported from a Spanish series, where the prevalence of all forms of AAV was under 45/million.

The two main forms of medium-vessel vasculitis are PAN and KD. PAN is extremely rare. There is very wide misconception among many physicians that PAN is the most common form of vasculitis, and this is partly encouraged by the older literature, which refers to all forms of vasculitis as PAN (initially called *periarteritis nodosa* and subsequently *polyarteritis nodosa*). In fact, the majority of patients with so-called PAN probably did not have this disease but were more likely suffering from one of the forms of small-vessel vasculitis, particularly MPA and GPA. PAN has been associated with infection, especially hepatitis B and hepatitis C. The epidemiology of hepatitis B has been transformed by effective immunization; as a result of this, hepatitis B associated with PAN is now an extremely rare disease. Recent figures suggest an incidence of PAN of 0.6–3.6/million adults.[5]

KD is most common in children under the age of 5 years but can occur in older children and young adults, in which case it is more difficult to diagnose because it is not suspected. In a recent Italian study of children under the age of 14 years, the incidence rate was 17.6/100 000 children under the age of 5 years, with a slight increase in reported cases during spring and winter compared with other times of the year. This is a slightly higher incidence rate than previously reported by other studies in Europe with a range of 3.6–15.2/100 000 children under the age of 5 years. Earlier figures from a large-scale study from the United States of over 6000 children admitted to hospital because of KD suggested that the peak age of onset was 1 year, and children under the age of 2 years accounted for over a third of all cases. The under-5-year incidence was reported as 8.1 per 100 000 children in 1988, rising to 18.5 per 100 000 children in 1997, similar to the recent Italian experience. Males were more commonly affected than females (approximately 60% males); there was no obvious seasonal variation. In a recent nationwide hospital survey in Japan, however, the incidence rates for KD during 2011 and 2012 were over 2400 cases per million children under the age of 5 years. The higher incidence of KD in patients of Japanese ethnicity appears to be independent of their geographical location. In fact, during 1996–2006, the average annual incidence of KD in Japanese American children in Hawaii was 210.5/100 000 children, in contrast to the rate for Caucasian children in Hawaii of 13.7/100 000, similar to the rate of 12.0/100 000 among Caucasian children in continental United States.[6]

Leukocytoclastic skin vasculitis is one of the more common forms of small vessel vasculitis occurring with an annual incidence of about 45/million (Table 58.1).[7] Less than a third is found in association with immunoglobulin A (Henoch-Schönlein purpura [HSP], or IgA vasculitis). IgA vasculitis is very common in children and is usually self-limiting; annual incidence is 100–200/million children ≤17 years of age.[8] By contrast, this is a much less common disease in adults (around 13/million/year).

Cryoglobulinemic vasculitis is strongly associated with hepatitis C, and its epidemiology is likely to mirror the prevalence of the hepatitis C virus (HCV). However, there are no published incidence and prevalence figures for cryoglobulinemic vasculitis itself.

PATHOGENESIS OF AAV

The Pathogenic Role of ANCA in GPA and MPA

There is some evidence that ANCA play a significant role in the pathogenesis of GPA, MPA, and EGPA. Based on the immunofluorescence pattern, different forms of ANCA can be distinguished, but only two are of direct clinical relevance: cytoplasmic c-ANCA (corresponding to antibodies directed against PR3) and perinuclear p-ANCA (predominantly corresponding to antibodies directed against myeloperoxidase [MPO]). p-ANCA can also be directed against other antigens, including bactericidal/permeability-increasing (BPI) protein, lactoferrin, human neutrophil elastase (HNE), cathepsin G, and azurocidin, but their clinical significance is not well characterized. Although ANCAs are associated with vasculitis, titers are not reliable for monitoring disease status because there is no clear relationship with remission or relapse.

Transfer of MPO-ANCA in humans (maternal–fetal route) and animal models (necrotizing pauciimmune glomerulonephritis after passive transfer of purified antibody or splenocytes from MPO-deficient mice immunized with purified murine MPO) has resulted in features typical of MPA.[13] By contrast, the pathogenicity of antiPR3 antibodies is less well established. In one animal model of autoimmune-prone nonobese diabetic (NOD) mice, immunization with recombinant mouse PR3 (rmPR3) in complete Freund's adjuvant (CFA) had no clinical

KEY CONCEPTS

Pathogenic Mechanisms in ANCA-Associated Vasculitides

- The recognition of an association between some forms of vasculitis and the presence of antibodies to neutrophil cytoplasm (ANCA) has transformed understanding of the group of diseases considered to be associated with ANCA and in which ANCAs are suspected to have a pathogenic role.
  - Granulomatosis with polyangiitis (GPA)
  - Microscopic polyangiitis (MPA)
  - Eosinophilic granulomatosis with polyangiitis (EGPA)
- A genome-wide association study (GWAS) has demonstrated strong associations with specific alleles.
  - Anti-proteinase 3 ANCA with specific HLA-DP alleles and alleles encoded by genes for α_1-antitrypsin and proteinase 3
  - Anti-myeloperoxidase ANCA with specific HLA-DQ alleles
- Environmental factors
  - Silica exposure
  - Specific strains of *Staphylococcus aureus*
- Inflammatory processes
  - Ineffective T-cell regulation
  - Neutrophil generation of extracellular traps (NETs) containing proteinase 3 and myeloperoxidase
  - Activation of alternative complement pathway by ANCA-activated neutrophils

effect but resulted in high levels of circulating c-ANCA; transfer of splenocytes from these immunized animals into mice with severe combined immunodeficiency (SCID) resulted in vasculitis and severe segmental and necrotizing glomerulonephritis.[14] Transfer of splenocytes from the CFA-alone-immunized mice (controls) resulted in no disease, further suggesting that disease development depends on PR3-specific immune responses. In a second model of PR3ANCA vasculitis, based on animals with a human–mouse chimeric immune system,[15] >70% of mice treated with IgG from patients with antiPR3 AAV (as compared with IgG from patients with nonvasculitic renal disease; or healthy controls) developed mild kidney disease with glomerular hypercellularity and focal pulmonary hemorrhage. Fifteen (83%) mice treated with antiPR3 IgG later showed mild kidney disease with glomerular hypercellularity, and three (17%) had severe glomerular injury. In the lungs, 13 (72%) showed areas of focal pulmonary hemorrhage, whereas lungs of the control group ($n = 8$) appeared normal ($p < 0.01$). There were no granulomatous lesions, but because granulomatous lesions are dependent on a robust T cell–mediated response, the authors argued that a refinement of the model to include greater levels of chimerism and administration of interleukin-7 (IL-7)–Fc protein to augment T-cell development would be required to study this.[15]

Even though these studies strongly support a pathogenic role for ANCA, conventional serological assays fail to detect ANCA in some patients with classic clinical and pathological features of AAV, and titers do not correlate well with disease activity. Roth et al. examined MPO epitopes specificities,[16] reporting 25 different epitopes bound by anti-MPO antibodies; although some epitopes were associated with active disease, others were either not specific to active disease or not associated with disease at all. Igs purified from patients with ANCA-negative vasculitis could bind to a specific MPO epitope. Furthermore, the absence of ANCA in some patients could be explained by competitive binding to a fragment of ceruloplasmin (CP), the natural inhibitor of MPO. This CP fragment decreased anti-$\text{MPO}_{447-459}$ autoantibody reactivity by 30–50%, whereas full-length CP did not have any effect.[16] ANCAs are reported in small numbers of healthy individuals. How the pathogenic transformation of ANCA occurs is still unclear, but it is probably a multifactorial process requiring a complex interaction among genetics, the environment, and the immune system facilitating a break in immune tolerance.

GENETICS

There has been significant progress in the understanding of AAV genetics following the publication of two genome-wide association studies (GWAS).[17] The strongest human leukocyte antigen (HLA) association is with the HLA-DPB1 haplotype, initially described in German cohorts and confirmed in both GWAS, especially for the PR3-ANCA–positive subgroup, regardless of the clinical diagnosis.[14] Further analysis has revealed an association between MPO-ANCA and a single nucleotide polymorphism (SNP) in the HLA-DQ region, which had probably been masked previously as a result of the small number of MPO-ANCA–positive patients included in initial analyses. Other HLA associations are reported, such as HLA-DRB1*09:01 and HLA-DQB1*03:03 in Japanese patients with MPA. Less robust findings, not replicated in other studies, include a protective effect of HLA-DR13(6) and HLA-DR1, but an increased proportion of HLA-DR4 in Dutch patients with GPA compared with controls[18]; HLA-DRB1 in PR3-ANCA-positive (but not MPO-ANCA) patients[17]; HLA-B50, HLA-DR1, HLA-DR9, HLA-DQw7, and HLA-DR3 in GPA.[19-21] In EGPA, the most robust association is with HLA-DRB4. Overall, there is evidence for genetic susceptibility to AAV, related to specific SNPs in the HLA region. Other genetic associations with GPA include *PRTN3, SERPINA 1, PTPN22,* and *CTLA4.*

PR3 is either stored in neutrophil azurophilic granules or exposed on the cell membrane (where it can interact directly with ANCA). Although the proportion of neutrophils displaying membrane PR3 (mPR3^+) is stable over time, surface expression of PR3 may be enhanced. The percentage of mPR3^+ neutrophils is genetically determined. Schreiber et al. showed that among 125 healthy controls, 35 patients with GPA, 15 patients with other inflammatory diseases, and 27 pairs of monozygotic (MZ) and dizygotic (DZ) twins, the percentage of mPR3^+ neutrophils correlated significantly in MZ twins (but not in DZ twins) and the heritability percentage was estimated as 99%. Furthermore, the absolute number of PR3 molecules expressed on the cell membrane was correlated among MZ (but not DZ) twins, with a heritability estimate of 96.7%.[22] Patients with GPA have a higher percentage of mPR3^+ neutrophils compared with healthy controls or patients with other inflammatory diseases (mPR3^+ percentage of 76.8% vs 56.1% vs 59.4%, respectively).[22] Following neutrophil activation and enzyme release, PR3 can mediate direct tissue damage. α_1-antitrypsin, encoded by the gene *SERPINA1* (found on chromosome 14), is the major inhibitor. Two α_1-antitrypsin alleles, Z and S, are associated with low enzymatic activity. A significant correlation with the Z allele (odds ratio [OR] 0.3; p = 1.25 × 10–5), but not the S allele, is reported in PR3-ANCA–positive patients with GPA.[17] *PRTN3* (the gene encoding PR3) is associated with GPA, especially among patients who are PR3-ANCA-positive.[17]

The 620WPTPN22 variant of *PTPN22* (encoding lymphoid tyrosine phosphatase) correlates with abnormal regulatory CD4 T regulatory cell (Treg) function, increased humoral activity, and enhanced neutrophil function in patients with GPA.[23] CTLA-4 (encoding cytotoxic T lymphocyte antigen-4) polymorphisms are associated with GPA,[14] especially the G allele of CT60 (p = 0.001) and +49 (p = 0.03). By contrast, only borderline significance was noted between CT60 SNP and MPA (p = 0.05).[24] CTLA-4 is a negative regulator of T-cell activation, which competes with the costimulatory molecule CD28 for binding of CD80 or CD86 on antigen-presenting cells (APCs). Abatacept, a monoclonal antibody (mAb), containing the binding domain of CTLA-4 reduces CD28–CD80/CD86 interaction and therefore T-cell stimulation, which could explain its potential benefit in a small trial of non–life-threatening GPA.[25] Haplotypes of IL-10 (a pleiotropic cytokine with complex and multiple effects in immune modulation) are associated with several immunological disorders, including systemic lupus erythematosus (SLE), rheumatoid arthritis (RA), and giant cell arteritis. IL-10 production is largely (50–70%) determined by genetic factors, and elevated plasma levels are reported in EGPA, but not in GPA.[26] Wieczorek et al. evaluated the impact of functionally relevant IL-10 polymorphisms in 403 patients with GPA, 103 with EGPA (compared with 507 controls).[26] There was a significant association with the 3575/1082/592 TAC haplotype, part of the extended ancient haplotype of IL-10.2, in ANCA-negative EGPA, but not in GPA.[26]

EPIGENETICS

Genetic factors alone are not enough to explain the range of phenotypic presentations in AAV. Epigenetic dysregulation is increasingly recognized as a contributor to immune-mediated diseases. Epigenetics relates to heritable changes in the function of genes, without altering the DNA sequence. It includes DNA methylation, histone, and microRNAs (miRNAs). Epigenetic modifications can be stable over time or can respond to developmental and environmental triggers, leading to phenotypic aberrances. Central to the pathogenesis of AAV is a dysregulated immune response resulting in ANCA production and aberrant expression of their target autoantigens, MPO and PR3. The expression of MPO and PR3 occurs primarily during early neutrophil development to produce intragranular constituents and is silenced in mature cells. However, in AAV, expression remains active. One epigenetic change that helps explain this is the reduced levels of the H3K27me3 histone modification at both the MPO and PR3 gene loci. H3K27me3 histone is associated with transcriptional silencing in patients with AAV compared with healthy individuals.[27]

ENVIRONMENTAL AND INFECTIOUS TRIGGERS

Genetic and epigenetic modifications may render a patient susceptible to developing disease in the presence of an appropriate trigger. Several triggers have been associated with AAV, including toxins, viral and bacterial infections, and some licit and illicit drugs.

Silica dust exposure has been associated with the development of AAV and other autoimmune diseases, such as SLE, RA, and scleroderma.[28] Silica is an abundant earth material found in sand, grain, grass, and wool. Processing these materials may expose workers to respirable crystalline silica. Silica-induced ANCA-positive disease is often associated with p-ANCA, targeting MPO, and the clinical presentation is usually MPA rather than GPA.[29] Case-control studies show that a high percentage (22–46%) of patients with AAV were previously exposed to silica.[25,29] A systematic review and meta-analysis of 332 cases and 516 controls found that prior silica exposure was associated with development of AAV, with a summary odds ratio (OR) of 2.56 (95% CI 1.51–4.36). Pooled ORs were similar for GPA and MPA, 3.56 (95% CI 1.85–6.82) and 3.95 (95% CI 1.89–8.24), respectively. Other confounders, such as other occupational exposures or tobacco use, were not analyzed, which could have resulted in an overestimation of the risk attributed to silica.[25] The mechanism by which silica exposure triggers the development of AAV is not fully understood. *In vitro,* silica can activate monocytes and macrophages, releasing cytokines, such as IL-1 and tumor necrosis factor-α (TNF-α), as well as releasing oxygen radicals and lysosomal enzymes, including proteinase 3 and myeloperoxidase.[30] Furthermore, silica can inactivate α_1-antitrypsin. Asbestos, another silicon-containing mineral, has been suggested as a trigger for AAV in a small case-control study of 31 patients (22 GPA, 8 MPA, 1 EGPA) and 30 healthy controls; three patients had prior exposure to asbestos versus none of the controls.[31] There are few published studies investigating the potential pathogenic role of asbestos in the development of AAV.

It is estimated that 63% of patients with GPA are chronic nasal carriers of *Staphylococcus aureus,* resulting in an increased risk of relapse. Maintenance treatment with trimethoprim-sulfamethoxazole (cotrimoxazole) has been shown to reduce the incidence of relapse in a double-blind, placebo-controlled study in patients with GPA, where a maintenance dose of 960 mg two times daily of trimethoprim-sulfamethoxazole resulted in a 60% reduction in relapse. However, this has not been replicated in other studies. The mechanism for the pathogenic role of *S. aureus* is still unclear. Potential pathways include stimulation of B and/or T cells by *S. aureus* superantigens; polyclonal activation of B cells by cell wall components of the bacterium, resulting in persistence of ANCA; and neutrophil priming leading to surface expression of PR3.[28]

Although parvovirus B19 has been proposed as a trigger for AAV in a few case reports, a case-control study failed to demonstrate any association because IgG antibodies to parvovirus B19 were detected equally in the sera of patients with AAV and control subjects, and all 13 patients with AAV and 39 controls were negative for IgM antibodies and viral DNA.[32] Hepatitis B virus (HBV) is implicated in the pathogenesis of PAN,[28] but there is no evidence to support a role for HBV or HCV in AAV. The presence of circulating c-ANCA and PR3-ANCA has been reported to be higher in patients with chronic hepatitis B and C compared with controls. In a study of patients with liver disease (11 patients with primary biliary cirrhosis, 216 with chronic hepatitis B, 516 with chronic hepatitis C, and 44 healthy controls), ANCAs were detected in 55.6% of patients with hepatitis C (all had a c-ANCA pattern with PR3 specificity and 4.8% had concurrent MPO specificity).[33] ANCA-positive patients had a higher prevalence of skin involvement, anemia, abnormal liver functions, and elevated α-fetoprotein (FP) levels[33]; however, the findings were not replicated in other studies.[34]

Lysosomal-associated membrane protein 2 (LAMP-2) is a heavily glycosylated type 1 membrane protein, abundant on neutrophil and endothelial cell surfaces, which shuttles between lysosomes and the cell membrane. LAMP-2 cross-reacts with FimH, a gram-negative adhesin, which facilitates bacterial entry to host tissues. Preliminary studies have suggested a role for LAMP-2 in small-vessel vasculitis. The LAMP-2 epitope P41–49 has been reported to have a 100% homology with amino acids 72–80 of mature FimH.[35] Antibodies to human LAMP-2 have been shown to injure human microvascular endothelium *in vitro* and induce focal necrotizing glomerulonephritis (FNGN) in rats; immunization with FimH-induced pauciimmune FNGN associated with antibodies that bound human and rat LAMP-2; furthermore, patients with pauciimmune FNGN were found to have an increased likelihood of infections with FimH-expressing bacteria shortly before presentation of their FNGN.[35] LAMP-2 antibodies were found in patients with active disease or relapse but not in those in remission.[35] Although this a promising new explanation for some cases of small-vessel vasculitis, the lack of replication of these findings makes them hard to substantiate.

CpG-oligodeoxynucleotides (ODN) is a short synthetic DNA containing unmethylated CpG motifs, highly prevalent in bacterial DNA, and recognized by Toll-like receptor 9 (TLR9), which is expressed by a variety of cells in the immune system. TLR9 triggering results in proinflammatory IL production. It has been reported that CpG motifs and IL-2 exposure can induce B lymphocytes to proliferate, increase antigen presentation, produce a range of cytokines, and differentiate into Ig-producing cells, ultimately leading to ANCA production.[36]

DRUG INDUCED AAV

The most commonly implicated drugs that can induce an AAV-like syndrome are propylthiouracil (PTU), hydralazine, levamisole-adulterated cocaine, TNF-α inhibitors, sulfasalazine, D-penicillamine, and minocycline (Fig. 58.1).

Propylthiouracil

The most commonly described drug associated with the presence of ANCA is propylthiouracil (PTU). PTU is used to treat hyperthyroidism and was the first drug described to induce a condition mimicking AAV in the 1990s. Even though its use has declined over time, reports of PTU-induced AAV still emerge. Clinically, PTU-induced AAV may manifest as acute kidney injury caused by pauciimmune necrotizing and crescentic glomerulonephritis in addition to respiratory tract, joint, and skin diseases.

Reports of the proportion of patients who develop ANCA as a result of exposure to antithyroid medications varies widely from 4% to 46%; by contrast, the prevalence of antithyroid drug-induced AAV is much lower (0–1.4%). Although PTU is the most commonly reported antithyroid drug, methimazole and carbimazole have also been implicated. In a study by Noh et al.,[37] the estimated annual incidence of antithyroid drug-induced vasculitis was 0.53–0.79 patients per 10 000 patients with Graves disease, especially for patients receiving propylthiouracil; the ratio of the estimated incidence for methimazole and propylthiouracil was 1 : 39. Most patients with PTU-induced ANCA will not develop clinical features of vasculitis. Patients with PTU-induced AAV (usually MPO-positive) are younger, more commonly female compared with those with primary AAV, and usually have significantly higher anti-MPO ANCA titers compared with patients without vasculitis. Other ANCA subtypes may be found, including HNE-ANCA and lactoferrin (LF)-ANCA; however, when PTU-induced AAV develops, MPO-ANCA is usually present, regardless of the existence of other forms of ANCA. The pathogenesis of AAV related to antithyroid drugs is unclear. PTU is predominantly metabolized in the liver, but a proportion is modified by MPO in neutrophils.[38] Neutrophils are responsible for neutrophil extracellular trap (NET) formation, and PTU can induce abnormal configured NETs *in vitro;* these abnormal NETs cannot easily be released into liquid phase and are not effectively digested by DNase I, therefore remaining in the tissue.[39] Immunization of rats with abnormal NETs results in production of MPO-ANCA.[39]

Hydralazine

Hydralazine is widely used to treat hypertension, acting as a smooth muscle relaxant and causing arterial and arteriolar vasodilation. It has been implicated in the development of drug-induced SLE and AAV. High titers of MPO-ANCA, ANA, anti–double-stranded DNA (dsDNA), and antihistone antibodies are found; hypocomplementemia is also frequent. The underlying mechanism of action for hydralazine-induced AAV is not clear. One potential pathway by which hydralazine acts is through reverse epigenetic silencing of tumor suppressors, but also potentially of MPO and PR3.

Levamisole-Contaminated Cocaine

It is commonplace for illicit drugs, such as cocaine, to be combined with adulterants, such as levamisole, in an effort to increase profits. Levamisole is similar in physical appearance and has possible potentiating effects on levels of dopamine in the central nervous system. It is estimated that >75% of cocaine users in the United States are exposed to levamisole.

Clinical manifestations of AAV induced by levamisole-contaminated cocaine include constitutional features; arthralgia; retiform purpura involving the ears, face, and extremities; and, less commonly, renal and lung diseases. Laboratory abnormalities include leukopenia, neutropenia, and high-titer p-ANCA, directed against MPO-ANCA or against atypical p-ANCA–associated antigens, such as HNE, lactoferrin, and cathepsin G. PR3-ANCA, ANA, and antiphospholipid autoantibodies have also been described in these patients. This multiplicity of antibodies helps distinguish AAV induced by levamisole-contaminated cocaine from primary AAV, which usually targets just one antigen.

Levamisole has an estimated mean half-life of 5.6 hours; therefore serum testing is likely to be negative if the most recent exposure occurred >24 hours prior to sample collection. Urinary detection of levamisole is highly suggestive of drug-induced disease and is useful if exposure occurred <48 hours prior to testing.

The mechanism by which levamisole-contaminated cocaine induces AAV is unclear. Levamisole, like PTU, could serve as a substrate for MPO to form reactive metabolites that might induce autoimmunity. Levamisole can enhance NETosis through engagement of muscarinic (subtype 3) receptors; furthermore, neutrophil extracellular traps (NETs) generated by levamisole can induce endothelial cell death and vascular dysfunction by disrupting normal endothelium-dependent vasorelaxation. Abusing cocaine without levamisole can also trigger a syndrome

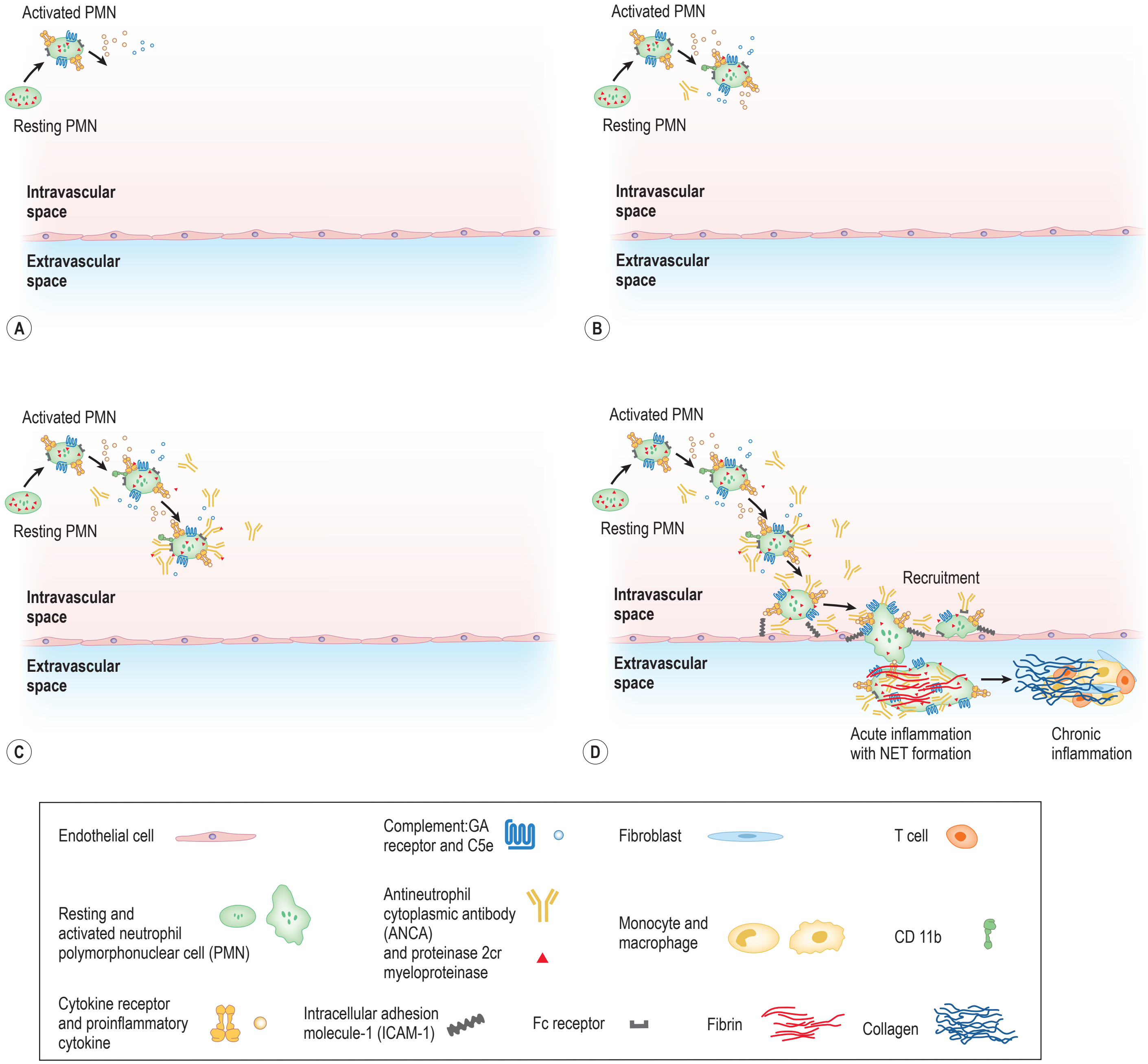

Activated PMN
Resting PMN
Intravascular space
Extravascular space
A
Activated PMN
Resting PMN
Intravascular space
Extravascular space
B
Activated PMN
Resting PMN
Intravascular space
Extravascular space
C
Activated PMN
Resting PMN
Recruitment
Intravascular space
Extravascular space
Acute inflammation with NET formation
Chronic inflammation
D
Endothelial cell
Complement:GA receptor and C5e
Fibroblast
T cell
Resting and activated neutrophil polymorphonuclear cell (PMN)
Antineutrophil cytoplasmic antibody (ANCA) and proteinase 2cr myeloproteinase
Monocyte and macrophage
CD 11b
Cytokine receptor and proinflammatory cytokine
Intracellular adhesion molecule-1 (ICAM-1)
Fc receptor
Fibrin
Collagen

FIG 58.1 Proposed pathogenic mechanisms in antineutrophil cytoplasmic antibody (ANCA) associated vasculitis. The interaction between activated neutrophils and ANCA is depicted in the figure, with resting neutrophils becoming activated (A), interacting with ANCA in the intravascular space (B), triggering a series of cascading events leading to vascular invasion (C) by inflammatory cells and the formation of neutrophil extracellular traps (NETs) and eventually chronic inflammation in the extravascular space (D).

mimicking AAV. The clinical presentation can be identical to that of patients with GPA, although cerebral angiitis, urticarial vasculitis, and EGPA-like syndromes have also been described. In GPA-like vasculitis, patients typically present with cutaneous vasculitis, nasal septal destruction, and pauciimmune crescentic glomerulonephritis. The autoantibody profile usually includes c-ANCA with PR3 specificity, although there are reports of cases with negative ANCA or p-ANCA with PR3 specificity. In these cases, the target of p-ANCA may be atypical, such as HNE.

Cocaine-induced midline destructive lesions (CIMDLs) in the upper respiratory tract mimic limited forms of GPA.[40] Patients with CIMDLs may present with or without ANCA, and when ANCA-positive, the pattern often varies. Some patients are positive for c-ANCA with PR3 specificity, but more often, patients with CIMDLs present p-ANCA with specificity for atypical ANCA, such as HNE. Wiesner et al. reported that 76% of the patients with CIMDLs were ANCA-positive (mostly p-ANCA): 57% had PR3-ANCA and 86% had HNE-ANCA; this compares with the absence of HNE-ANCA in GPA and MPA, suggesting that the presence of HNE-ANCAs may be helpful in distinguishing CIMDL from GPA.[41]

Table 58.2 provides a summary of the most significant associations between drugs and ANCA or AAV. The management of all forms of drug-induced AAV is withdrawal of the offending agent, supportive measures, and, in severe cases, immunosuppression, dialysis, and plasma exchange.

LOSS OF B- AND T-CELL TOLERANCE IN ANCA-ASSOCIATED VASCULITIS

The imbalance of effector and Tregs underpins the autoimmune dysregulation seen in AAV with multiple alterations in the circulating T-cell population. Patients with AVV have a reduced number of circulating Tregs,[42] which are functionally impaired, and an expanded population of CD4 effector memory T cells with an increased number of activated T cells.[43] Persistence of CD4 T-cell activation in peripheral blood correlates with disease severity.[44] Aberrant T helper (Th) polarization, with an increase in proinflammatory Th17 responses, further contributes to vascular injury.[43] It has been suggested that these alterations to the peripheral T-cell compartment could be influenced by environmental factors, such as infection.

The discovery of B cells in the inflammatory lesions in AAV together with the success of B-cell depletion therapies suggests that B cells play a significant role in the pathogenesis of AAV; however, the exact mechanism of this involvement is still not known. Regulatory B cells are reduced in AAV, whereas B-lymphocyte stimulator (BLyS) levels are significantly increased, thereby promoting B-cell differentiation, proliferation, and survival.[45] B cells may play a number of roles in AAV: as precursors to antibody-producing plasma cells; as APCs; producing proinflammatory mediators; or in costimulation of T cells.

ROLE OF NEUTROPHILS

In addition to containing the antigens for ANCA, activated neutrophils release many mediators that modulate the inflammatory response and can directly contribute to tissue inflammation, vascular injury, and damage in AAV via phagocytosis, degranulation, and cytokine production. Neutrophils release B cell–activating factors (BAFFs) that enhance B-cell proliferation and retard apoptosis. Neutrophils from patients with AAV are more prone to spontaneous release of NETs.[46] In the normal immune system, NETs consist of chromatin fibers released from dying neutrophils and are designed to trap and kill extracellular pathogens. NETs not only contain proinflammatory proteins that directly cause endothelial cell damage and complement activation but also form a link between innate and adaptive immunity via providing access to MPO and PR3.

ROLE OF COMPLEMENT

Despite the apparent paucity of immune complexes in AAV, complement (and in particular activation of the alternative complement pathway) plays a crucial role in the pathogenesis of AAV. When primed neutrophils are activated by ANCA, they produce C5a, which, in addition to recruitment, primes additional neutrophils for further activation by ANCA.[47] C3a, C5a, and soluble C5b-9 levels are elevated in active disease; plasma levels of complement factor H, a regulator of the alternative complement pathway, are significantly lower in patients with active AAV.[48] Low serum levels of C3 at diagnosis are associated with a worse prognosis.[49] There is also a suggestion that C5a may play a role in the hypercoagulability associated with active AAV.

TABLE 58.2 **Drug-Induced Vasculitis Associated With ANCA Positivity—Implicated Drugs, Proposed Mechanisms of Action and Laboratory Findings**

Drug/Class	Proposed Mechanism of Action	IF Pattern	ANCA Serotype	Other ANCA Autoantigens	Other Antibodies	Ref.
Allopurinol	Limited data	p-ANCA	MPO-ANCA	-	ANA	50
Anti–TNF-α (ADA, ETN, IFX)	TNF-α may induce the formation of immune complexes, activation of complement and mediate inflammation by switching from a cytokine response of T-helper type 1 to type 2, upregulating antibody production	p-ANCA p-ANCA c-ANCA	MPO-ANCA MPO-ANCA PR3-ANCA	-	ANA	51
Benzylthiouracil	Limited data	p-ANCA	MPO-ANCA	HNE Lactoferrin	-	52
Carbimazole	Limited data	p-ANCA c-ANCA	MPO-ANCA PR3-ANCA	-	-	
Cocaine	Enhanced formation of NETs enriched in neutrophil elastase and inflammatory mitochondrial DNA with enhanced release of B cell–activating factor belonging to the TNF family (B cell–activating factor [BAFF])	c-ANCA p-ANCA	PR3-ANCA	HNE		40, 41

Continued

TABLE 58.2 Drug-Induced Vasculitis Associated With ANCA Positivity—Implicated Drugs, Proposed Mechanisms of Action and Laboratory Findings—cont'd

Drug/Class	Proposed Mechanism of Action	IF Pattern	ANCA Serotype	Other ANCA Autoantigens	Other Antibodies	Ref.
D-penicillamine	Limited data	p-ANCA	MPO-ANCA	HNE Lactoferrin	ANA Anti-dsDNA abs	[53]
Hydralazine	Hydralazine mediated reversal of epigenetic silencing of MPO and PR3, with subsequent increased expression of both autoantigens. Induction of cytotoxic products and neutrophil apoptosis mediated by hydralazine binding MPO	P-ANCA	MPO-ANCA PR3-ANCA	HNE Lactoferrin	ANA Anti-dsDNA abs Antihistone abs Anticardiolipin abs	
Levamisole contaminated cocaine	Enhanced NETosis through engagement of muscarinic subtype 3 receptor Metabolism of levamisole by MPO with formation of reactive metabolites Genetic susceptibility to agranulocytosis in HLA B27+ patients	p-ANCA c-ANCA	MPO-ANCA PR3-ANCA (double + is very common)	HNE Cathepsin G Lactoferrin	ANA Anticardiolipin abs	[54, 40]
Methimazole	Limited data	p-ANCA c-ANCA Atypical ANCA	MPO-ANCA PR3-ANCA	HNE	ANA	[55, 56]
Minocycline	Minocycline oxidation by MPO, with abnormal production of reactive metabolites and modification of enzymatic activity, triggering the induction of ANCA Cytotoxicity leading to premature apoptosis of neutrophils, with abnormal release of nucleosomes and drug-modified proteins (including myeloperoxidase, elastase, LL37, and HMGB1), which can be bound to neutrophil extracellular traps (NETs) triggering lupus/vasculitis via type I IFN production	p-ANCA Atypical ANCA	MPO-ANCA PR3-ANCA	HNE Cathepsin G BPI	ANA Anti-dsDNA abs	[57]
Propylthiouracil (PTU)	Cross-reactivity between anti-thyroperoxidase (anti-TPO) antibody and MPO Alteration of NET configuration PTU-induced structural change in MPO, which serves as a neoantigen Metabolism of PTU by MPO (in the presence of H_2O_2 and Cl^-) into strong toxic metabolites. Subsequent binding to neutrophils proteins and recognition by T cells with B cells activation resulting in autoantibody production Competitive inhibition of MPO oxidation activity by PTU, in a dose dependent manner	p-ANCA atypical ANCA	MPO-ANCA PR3- ANCA	HNE Lactoferrin BPI Azurocidin Cathepsin G	ANA Antihistone abs Anticardiolipin abs	[27, 55, 95, 58]
Sulfasalazine	Sulfasalazine-induced neutrophil apoptosis, with subsequent membrane expression of PR3 and MPO	p-ANCA c-ANCA	MPO-ANCA PR3-ANCA	Lactoferrin	ANA Anti-dsDNA abs	[50, 59]

abs, antibodies; *ADA,* adalimumab; *ANA,* antinuclear antibodies; *ANCA,* antineutrophil cytoplasmic antibody; *BPI,* bactericidal permeability increasing protein; *dsDNA,* double-stranded DNA; *ETN,* etanercept; *HLA,* human leukocyte antigen; *HNE,* human neutrophil elastase; *IF,* immunofluorescence; *IFN,* interferon; *IFX,* infliximab; *MPO,* myeloperoxidase; *NET,* neutrophil extracellular trap; *PR3,* proteinase 3; Ref, references; *TNF,* tumor necrosis factor.

PATHOGENESIS OF KAWASAKI DISEASE

The incidence of KD is highest in East Asian countries, such as Japan, Korea, and Taiwan. The incidence in Japanese children is 10-fold higher compared with that in the United States. Japanese data show a seasonal variation in incidence, with an increase seen in the winter; furthermore, siblings of patients with KD are at increased risk of disease compared with the general population,[60] and the risk of sibling patients is increased among patients whose parents previously had the disease.[60] KD shows geographical clustering, with shared clinical features between clusters of disease. These findings suggest a role for an infectious environmental trigger, together with genetic susceptibility. Several disease susceptibility genes have been identified, including genes that affect the function of molecules involved in the calcineurin-nuclear factor of activated T cells (NFATs) pathway; transforming growth factor-β (TGF-β) lipopolysaccharide-induced endothelial cell inflammation pathways; and genes encoding Fcγ receptors, such as *ITPKC, FAM167-BLK, CD40, FCGR2A, HLA, CASP3, TGFB2, TGFBR2, SMAD3,* and *PLCB4/PLCB1.*[61] *CD40* and *FCGR2A* have been associated with disease susceptibility in Europeans.

SNPs identified within the *ITPKC* and *PLCB4/PLCB1* genes are associated with the risk of coronary artery aneurysms.[61] *ITPKC* and *CASP3* SNPs are associated with lack of response to intravenous immunoglobulin (IVIG) treatment.[62] ITPKC is a negative regulator of T-cell activation via downregulation of the Ca^{2+}/NFAT signaling pathway. The SNP located in the intron of *ITPKC* regulates splicing, decreasing negative regulation of T cells and leading to increased T-cell activation and IL-2 production.[63]

TNF-α, nuclear factor-κB (NF-κB), IL-17, TGF-β, interferon-γ (IFN-γ), granulocyte–colony-stimulating factor (G-CSF), IL-1β, IL-6, follistatin-like protein 1, and TLR2 are reported to be elevated in patients with KD.[64,65] In the acute phase of disease, circulating IL-17 levels and IL-17–induced cytokinesis are increased, suggesting an imbalance between Th17 and Tregs. Th17 levels are higher in children with coronary lesions.[66] There is evidence for marked activation of cytotoxic and Th lymphocytes and dendritic cells, with upregulation of type I IFN responses, supporting a possible viral etiology of KD, based on a transcriptomic study of patients who died or who were undergoing heart transplantation.[67]

KD has been suspected to be triggered by bacterial or viral infections, but no single causative agent has been identified. It has been suggested that superantigens (SAgs) are implicated. Several microorganisms can produce SAgs, including bacteria (*e.g.,* staphylococci, streptococci, mycobacteria, *Mycoplasma, Yersinia, Lactobacillus)* and viruses (*e.g.,* Epstein-Barr virus). SAg-producing bacteria have been isolated from children with acute KD, including TSST-1 producing *S. aureus* and pyrogenic exotoxin–producing *Streptococcus.*[68] In a mouse model involving injections of intraperitoneal *Lactobacillus casei*, disease could not be induced in recombinase activating gene-1–deficient mice, implicating T cells as critical drivers of the disease process.[69]

PATHOGENESIS OF POLYARTERITIS NODOSA

The pathogenesis of idiopathic PAN is unknown; clinical responses to immunosuppressive therapy strongly support an underlying immunological mechanism. Elevated levels of IL-2, IL-8, IFN-γ, TNF-α, and IL-1β have been reported in PAN, as well as elevated levels of circulating soluble adhesion molecules, such as intercellular adhesion molecule-1, vascular cell adhesion molecule-1, and E-selectin, suggesting that endothelial cell activation could perpetuate and potentiate the inflammatory milieu. IFN-γ and TNF-α increase the expression of class I major histocompatibility complex (MHC) antigens and induce MHC class II expression, resulting in antigen presentation to T cells. In one study of 24 patients with idiopathic PAN, macrophages and CD4 T lymphocytes were the predominant cell types found in lesions. T lymphocytes were more abundant in nerve biopsy specimens compared with muscle biopsy specimens (52% vs 35%; $p < 0.001$), whilst macrophages were predominant in muscle specimens (45% vs 33%; $p = 0.005$). Histologically, T lymphocytes were distributed throughout the vessel wall, whereas macrophages predominated in the periphery.[70]

In HBV-induced PAN, there is direct injury to the vessel wall as a result of viral replication[59] and deposition or *in situ* formation of immune complexes, which activate complement, thereby recruiting and activating neutrophils. The immunological process usually occurs within 6 months of HBV infection. During the active phase of disease, complement levels are low, consistent with complement consumption resulting from immune complex deposition.

Alterations in the *MEFV* gene (encoding pyrin) may represent an important susceptibility factor for PAN. Pyrin is critically important in regulating IL-1β production; mutations of pyrin may be associated with loss of regulation of inflammatory pathways via apoptosis and IL-1β release, resulting in augmented inflammation, similar to patients with familial Mediterranean fever, who can be carriers of *MEFV* mutations.[71]

PATHOGENESIS OF CRYOGLOBULINEMIC VASCULITIS

The pathological hallmark of cryoglobulinemic vasculitis is deposition of cold-precipitating Igs in small vessels, resulting in vascular inflammation and damage. In most patients, chronic HCV infection is the trigger, but cryoglobulinemic vasculitis can also occur in the context of other chronic infections and in some autoimmune rheumatic diseases and lymphoproliferative disorders.

HCV plays several important roles in the pathogenesis of cryoglobulinemic vasculitis. HCV glycoproteins interact directly with B-cell surface receptors, lowering the activation threshold and resulting in polyclonal activation and expansion of B cells.[72] Circulating, clonally expanded B cells produce monoclonal IgM rheumatoid factor (RF). Mixed cryoglobulins contain a monoclonal IgM RF directed against the Fc segment of IgG and polyclonal IgG. When IgM RF is exposed to cold conditions, it undergoes conformational changes resulting in cryoprecipitation. Cold-insoluble immune complexes are formed predominantly by IgM RF linked to IgG (which, in turn, is bound to HCV core protein). Cryoglobulinemic vasculitis should be considered an antigen-driven disease, because HCV persistence ensures sustained lymphoid proliferation and continued cryoglobulin production.

C1q protein levels and C1q binding are significantly increased in cryoglobulinemic vasculitis. C1q is required for binding of immune complexes to endothelial cells. Although mixed cryoglobulins can be detected in up to 60% of patients with chronic hepatitis C, cryoglobulinemic vasculitis occurs only in a minority, suggesting that host factors must be equally important in the pathogenesis of the disease.

PATHOGENESIS OF EGPA

Despite the established role of ANCA in GPA and MPA, the pathogenic role of ANCA in EGPA is less clear. In EGPA, only 30–40% of patients are ANCA positive; the presence or absence of ANCA in EGPA defines distinct clinical phenotypes. The pathological hallmark of EGPA is prominent tissue and peripheral blood eosinophilia, suggesting a central pathogenic role for the eosinophil. EGPA is considered to be Th2 mediated, given the striking early allergic manifestations, together with the marked increase in circulating Th2 cytokines, including IL-5, which plays a crucial role in eosinophil differentiation in bone marrow as well as recruitment and activation of eosinophils at sites of inflammation. IL-5 could delay apoptosis of eosinophils and modulate the function of mast cells and basophils.[73] Elevated levels of IFN-γ (a potent Th1 cytokine involved in granuloma formation) provide evidence for the involvement of other Th responses. Th17 responses are upregulated, and B lymphocytes and humoral responses are likely to be important, given that B-cell depletion can induce remission and reduce circulating IL-5 and eosinophils.

The precise mechanisms of eosinophil-mediated inflammation in EPGA remains unclear. Eosinophils are granulocytic cells capable of releasing multiple proinflammatory cytokines, chemokines, and reactive oxygen species, which have direct effects on the vessel and perivascular tissues, including tissue fibrosis, thrombosis, and allergic inflammation; indirect effects include recruitment and activation of other inflammatory cells to perpetuate the inflammatory response.

CLASSIFICATION

Several sets of classification criteria have been published, the most recent being one by the American College of Rheumatology (ACR) in 1990. These are primarily for the purposes of epidemiological studies or for defining patients for inclusion in clinical studies, rather than being used in clinical practice. However, as is often the case with classification criteria, they are misapplied as diagnostic criteria in daily practice and may fail to provide a clear-cut distinction between vasculitis and other conditions.[74] An international effort is currently underway to develop improved classification criteria as well as diagnostic criteria for patients with vasculitis or suspected vasculitis, and it is likely to lead to a new set of evidence-based criteria to improve diagnosis and classification of vasculitis in the future.[74] There is considerable overlap between different forms of vasculitis. The group of patients with diseases clustered around the presence of ANCA share many features in common, particularly the presence of glomerulonephritis and pulmonary infiltrates. In the case of EGPA, patients also experience rhinitis, nasal polyps, asthma, and rashes and may develop mononeuritis multiplex or sensory peripheral neuropathy. In addition to renal and lung inflammation, patients with GPA typically suffer from upper respiratory problems, such as nasal crusting or discharge; subglottic stenosis; or hearing loss. Patients with MPA do not usually develop significant upper airway disease; their disease tends to be more limited to the kidneys and/or lungs. All of these conditions can also affect any other organs, such as skin. Skin is the most commonly affected site in most forms of vasculitis. This can vary from small infarcts around nail edges to purpura, ulcers, nodules, and even gangrene.

The development of consensus criteria for diagnosis by the Chapel Hill group has provided a more rational approach to define disease entities. They are not true classification or diagnostic criteria because they are not evidence based (they are consensus of opinion from several experts); but they have advantages over the 1990 ACR criteria. They were developed after recognition of ANCAs, which are an important part of the 2013 definitions. One of the major flaws of the ACR criteria is the failure to recognize MPA as an entity separate from PAN.[74] This failure to separate PAN and MPA has probably prevented proper diagnosis, and it is a historical legacy from an assumption that almost all forms of vasculitis are, in fact, variations of PAN.

DIAGNOSIS

Individual Diseases

AAV are a group of three main entities: GPA, MPA, and EGPA. These three conditions overlap with many shared clinical features. Figs. 58.2 and 58.3 describe clinical patterns in patients with MPA and GPA and show that there are many similarities, with a dominance of kidney and lung involvement in both conditions. EGPA is characterized by eosinophilia, late-onset asthma, and neuropathic involvement, but cardiac involvement may develop in a subset of patients with EGPA, particularly in association with hypereosinophilia. Systemic features, such as malaise, fever, weight loss, or myalgia, could be mistaken for a variety of other conditions, and it may delay recognition of disease. Patients with MPA can present just with isolated microscopic hematuria and hypertension, with nonspecific systemic features. In contrast, patients with GPA typically suffer from upper airway problems, with nasal crusting, blood-stained discharge, sinusitis, or hearing loss. In anti–glomerular basement membrane (anti-GBM) antibody disease, clinical features sometimes overlap with those of AAV, but the dominant features are pulmonary hemorrhage leading to hemoptysis and rapid deterioration in renal function. It is not entirely clear whether the anti-GBM disease should be regarded vasculitis or vasculopathy, but it has been included in the current Chapel Hill Consensus definitions of a small-vessel vasculitis. Confusingly, some patients with anti-GBM antibodies also have ANCA.

The clinical course in MPA is typically acute, whereas GPA may remain undiagnosed for several months or even years before the upper airway and or lower airway symptoms are recognized as being related to vasculitis. For almost all forms of acute AAV, the differential diagnosis is infection, cancer, or drug toxicity. One of the drugs recently recognized to mimic AAV is cocaine,[75a] which can cause local tissue necrosis, especially in the palate, leading to palatal and nasal septal perforation. Interestingly cocaine can induce ANCA production and therefore can be a true mimic of AAV. Typically the ANCA is directed against elastase or is nonspecific. Careful history taking, sometimes with sensitive, confidential, repeated questioning, may be required to eventually identify this as a cause of a patient's condition. Contamination of cocaine with levamisole is becoming a growing problem. Presumably, levamisole (a veterinary anthelmintic agent) is a cheap white powder that can be mixed readily with cocaine but can cause an acute necrotizing vasculitis of skin and the extremities. Discontinuation of levamisole and cocaine usually stops further progression of disease, but once palatal or nasal perforation has developed, patients may require local surgical repair.

Cryoglobulinemic vasculitis typically presents with purpuric lesions on the legs, and the lesions may ulcerate. Patients often complain of joint pains and malaise. Neuropathies can occur with either mononeuritis multiplex or sensory neuropathy;

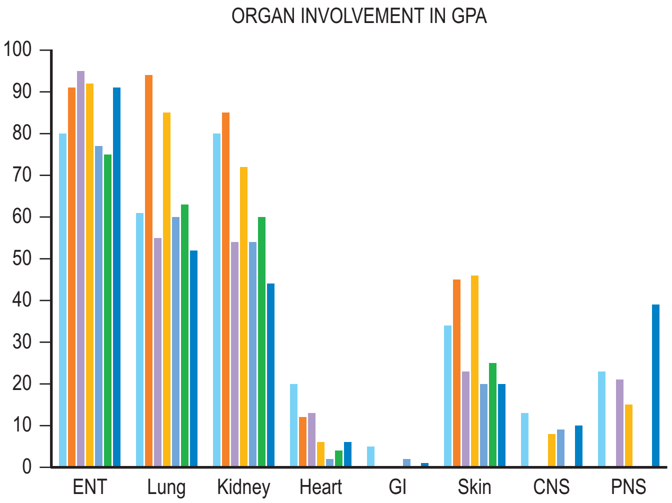


FIG 58.2 Organ involvement in patients with granulomatosis with polyangiitis (GPA). (Data compared across seven cohorts—reviewed by Luqmani R, Ponte C. ANCA associated vasculitides and polyarteritis nodosa. In: Bijlsma JWJ, Hachulla E, eds. *EULAR Textbook on Rheumatic Diseases*. London, UK: BMJ Publishing Group Ltd; 2015. Chapter 27, p.717-753)[75]

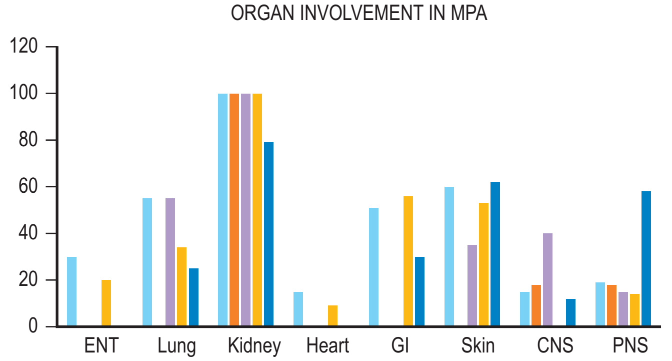


FIG 58.3 Organ involvement in patients with microscopic polyangiitis (MPA). (Data compared across seven cohorts—reviewed by Luqmani R, Ponte C. ANCA Associated Vasculitides and Polyarteritis Nodosa. In: Bijlsma JWJ, Hachulla E, eds. EULAR TEXTBOOK ON RHEUMATIC DISEASES. London, UK: BMJ Publishing Group Ltd; 2015. Chapter 27, p.717-753)[75]

membranous glomerulonephritis is a recognized complication. Most cases appear to be associated with hepatitis C (>90% in some series), and eradication of hepatitis C appears to be effective in reducing the manifestations of disease.

True PAN typically presents with neuropathies, systemic inflammation, and ischemic abdominal pain (difficulty eating because of medium-vessel supply to the gut and nerves). If it is associated with hepatitis B, the patient may or may not have obvious signs of liver disease.

Isolated cutaneous vasculitis typically presents as purpuric lesions, skin ulceration, or broken livedo, and biopsy of the skin will show typical changes of a necrotizing vasculitis.

KD usually presents in childhood as an acute illness, with relapsing high fevers, significant lymphadenopathy, and a systemic inflammatory syndrome. The child typically has mucosal inflammation with strawberry tongue; approximately 10 days after the onset, skin desquamation is a very typical feature. The most worrying feature of KD is the development of coronary artery aneurysms, which occur in around 2–4% of children and can be fatal, if untreated. This can be detected by using echocardiography.

Laboratory Investigations

The investigation of patients with suspected small- or medium-vessel vasculitis should follow a careful history and examination to determine the likely diagnosis and underlying illness. The differential diagnosis is very wide. It is important to be vigilant in looking for positive signs of vasculitis, but it is equally important not to forget to look for more common causes of the clinical presentation. Many of the studies performed can result in nonspecific findings, such as an elevated white blood cell count, platelet count, or the erythrocyte sedimentation rate. The C-reactive protein (CRP) level is typically raised. Patients may be anemic. Liver or, more importantly, renal function may be abnormal. The presence of hypereosinophilia can be suggestive of a diagnosis of EGPA, but there are other causes of hypereosinophilia, particularly drug reactions. It is important to test renal function in patients with suspected vasculitis in case there is nephritis, and it is always important to perform urinalysis for microscopic hematuria or proteinuria. Although these might be explained by the presence of a kidney infection or other causes, they raise a strong suspicion of glomerular inflammation. An abnormal urinary sediment in combination with hypertension should alert the physician to the possibility of kidney involvement by small-vessel vasculitis. Since the 1980s, the discovery of ANCA[12] has transformed the recognition of renal vasculitis, so that these patients can be managed more effectively.

Histology remains a very important diagnostic test, not only to make a positive diagnosis but also to exclude other causes. Although histology from the airways can be nondiagnostic, this may still assist in ensuring that the patient does not have cancer, sarcoidosis, o tuberculosis, or IgG4-related disease, all of which could present in a similar way with inflammation of the upper or lower airway. Renal histology is still the gold standard to diagnose suspected glomerulonephritis and may be useful in predicting the prognosis.[76] Four categories of renal lesion have been proposed: focal, crescentic, mixed, and sclerotic. Follow-up of patients with different patterns has shown outcomes: progressively worse renal outcome from focal (the best) through to sclerotic (the worst) over the subsequent 5 years.

ASSESSMENT

The outcome of small- and medium-vessel vasculitis has been completely transformed by immunosuppressive therapy. With treatment almost all patients with KD recover from their initial illness. In AAV, >70% survive at least 5 years from starting treatment. The majority of patients with PAN respond to initial therapy. However, the risk of recurrence of disease is high in small-vessel vasculitides. Relapse rates are likely to exceed 50% over time.[46] Patients experience comorbidities as a consequence of the disease or its treatment and can suffer from infections or other complications that contribute to their overall poor health. In the long term, small-vessel vasculitides and their treatment can be associated with effects on the cardiovascular system with increased risk of hypertension, coronary heart disease, and stroke.

Immunosuppressants given to patients can significantly improve disease control but may increase the risk of malignancy. The continued use of high doses of glucocorticoid therapy contributes to hypertension, diabetes, heart diseases, osteoporosis, and infection.

It is therefore obvious that careful evaluation of a patient's status is required throughout the course of disease. Constant vigilance is required to detect and manage any flare-up of disease because the consequences of untreated inflammation of vessels to vital organs can be severe. Although for most patients, initial therapy is very successful and improvements in disease status is obvious to the patient, the subsequent disease course can be much more complicated because of comorbidities, the evolution of disease-related damage, and morbidity caused by treatment as well as disease flare-ups. There are no suitable biomarkers that can be universally applied in small- and medium-vessel vasculitides to determine the patient's disease state to provide an evidence-based rationale for adjustment of therapy. ANCA titers fluctuate during the course of disease, and although ANCA remains a very useful diagnostic test, its value in managing variations in disease activity is very limited. Up to 40% of patients have an elevation of ANCA without any new deterioration in clinical state, but a recent study has shown that patients with renal involvement as part of their presentation are likely to have relapses associated with ANCA rises.[77] Among 166 patients with AAV, all were positive for ANCA and 104 had renal involvement (a mixture of PR3-ANCA and MPO-ANCA). The hazard ratio for ANCA rises predicting subsequent relapse was 11.09 (CI 5.01–24.55), suggesting that this test may be of value in detection of future relapse in this subset of patients. However, for the majority of patients with systemic vasculitis, careful clinical evaluation remains a cornerstone of effective disease management as recommended by the EULAR guidelines on management of systemic vasculitis.[78]

In simple terms, at the onset of the condition, disease activity is the dominant problem, and treatment should be directed toward it; but over the course of time, with the development of consequences of vasculitis or its treatment, there is an increasing component of damage or scarring and also side effects of drug therapy. Similarly, different patients function at different levels with the same amount of disease, or damage, and therefore the ability to perform normal tasks is an important component of their overall condition to consider (Table 58.3).

The primary tool used is the Birmingham Vasculitis Activity Score (BVAS) in adults and the Paediatric Vasculitis Activity Score (PVAS) in children (reviewed in Ponte et al., 2014).[79] Training is recommended for assessment in the Birmingham Vasculitis Activity Score (BVAS). BVAS provides a quantitative score based on individual items, providing an effective means to define the patient's status with regards to response to therapy. Many recent studies of different immunosuppressive agents have made use of the BVAS either to define improvement in terms of a fall in BVAS score or to define a cut-off representing active disease, or inactive disease, or flare-up, depending on the number of items present. Although it is subject to observer variability, it provides an effective means by which groups of patients can be compared against each other and allows individual patients to be followed up over the course of their condition. The score is weighted according to the organ system and the individual manifestation, to reflect the severity of the disease. The range of scores for BVAS and PVAS are 0–63. For BVAS GPA, items are divided into major (scoring 3 points) and minor (scoring 1 point each); with 15–19 major items and 19–23 minor items, the range of sores is 0–76. The PVAS was developed as a pediatric version of the BVAS, but validated in children with vasculitis and demonstrated to be effective at discriminating different disease states and is increasingly used as a research tool in pediatric vasculitis.

KEY CONCEPTS

Assessment of Disease Activity in Vasculitis

- For small- and medium-vessel vasculitis in adults, version 3 and a specific version of the BVAS for granulomatosis with polyangiitis (GPA) are used.
- BVAS version 3 is the most generic and applicable across different forms of vasculitis; BVAS for GPA has been specifically designed for use in GPA and contains very specific items related to that condition.
- During sequential monitoring, a time frame of 3 months is recommended so that disease activity is considered to be of most relevance during this time. The time frame is based on pragmatic clinical experience that this is the usual time taken during which immunosuppressive therapy is likely to have a significant effect on active disease manifestations.

Damage Assessment in Vasculitis

The concept of damage in patients with vasculitis is about the permanent consequences of having vasculitis. It is an attempt to measure disease burden regardless of cause. The Vasculitis Damage Index (VDI) is the most widely used and validated measure for assessment of damage in vasculitis (reviewed in Ponte et al., 2014).[79] The VDI captures the long-term consequences of a diagnosis of vasculitis and its treatment and associated comorbidities. Damage is defined as lasting at least 3 months or occurring at least 3 months ago for single time point events (*e.g.,* stroke or myocardial infarction) and should be recorded as a permanent change to the patient's damage status. A VDI of >5 points recorded within 6 months of disease carries a significant increased risk of subsequent mortality compared with a lower damage index at 6 months (reviewed in Ponte et al., 2014).[79] VDI is a useful index of future harm. A further development of the damage index has been the combined damage assessment (CDA). In a comparative study with the VDI, CDA was shown to be inferior to it (reviewed in Ponte et al., 2014).[79]

TREATMENT

Once a diagnosis of small- and medium-vessel vasculitis has been established, treatment should be focused on patients and their problems rather than the specific diagnosis. Treatment for different forms of vasculitis may look very different, but many aspects of different forms of vasculitis require the same therapeutic approach. Without a clear understanding of the underlying pathogenesis of disease, we are inevitably led by the need to suppress inflammation and reduce damage to prevent mortality and improve survival. However, as well as modifying immune dysregulation, it is important to consider other aspects of the patient's condition, such as comorbidities and the prevention of future comorbidities.

TABLE 58.3 Assessing Disease Activity in Vasculitis

System Assessed	BVAS	BVAS WG	PVAS	Comments
General	Myalgia, arthralgia/ arthritis, fever, weight loss	Arthritis, fever	Myalgia, arthralgia/ arthritis, fever, weight loss	These are typical features in many forms of systemic vasculitis.
Skin	Infarct, purpura, skin ulcer, gangrene	Purpura, skin ulcer, gangrene	Polymorphous exanthema, livedo, panniculitis, purpura, skin nodules, infarct, ulcer, gangrene, other skin vasculitis	Skin is a common organ involved in most forms of small and medium vessel vasculitis. Skin manifestations in ANCA-associated vasculitides tend to be less serious compared with manifestations in kidneys, lungs, and upper airways. For skin vasculitis in children the manifestations are more diverse.
Mucous Membrane/ eyes	Mouth ulcers, genital ulcers, adnexal inflammation, proptosis, red eye, (epi)scleritis, conjunctivitis, blepharitis, keratitis, blurred vision, visual loss, uveitis, retinal vasculitis	Mouth ulcer, conjunctivitis/ episcleritis, retro-orbital mass, uveitis, scleritis, retinal exudates	Mouth ulcers, genital ulcers, adnexal inflammation, proptosis, red eye (epi)scleritis, conjunctivitis, blepharitis, keratitis, blurred vision, visual loss, uveitis, retinal vasculitis	Eye involvement is most common in GPA and less often seen in MPA or EGPA.
Ear, nose, and throat	Bloody nasal discharge, paranasal sinus involvement, subglottic stenosis, conductive hearing loss, sensorineural hearing loss	Blood nasal discharge, paranasal sinus involvement, swollen salivary glands, subglottic inflammation, conductive hearing loss, sensorineural hearing loss	Blood nasal discharge, paranasal sinus involvement, subglottic stenosis, conductive hearing loss, sensorineural hearing loss	ENT manifestations are most common in GPA and rarely seen in MPA or EGPA. However, EGPA is often characterized by the presence of nasal polyps, which are inflammatory in nature.
Chest	Wheezing, nodules or cavities, pleural effusion, infiltrates, endobronchial changes, massive hemoptysis, respiratory failure	Nodules or cavities, pleurisy, infiltrates, endobronchial changes, alveolar hemorrhage hemoptysis, respiratory failure	Wheezing, nodules or cavities, pleural effusion, infiltrates, endobronchial changes, massive hemoptysis, respiratory failure	The chest is commonly affected in all three forms of ANCA vasculitis. Wheezing is a common feature of EGPA as are infiltrates. In contrast, infiltrates, nodules, and endobronchial disease dominate in GPA. Massive hemoptysis can occur in patients with GPA or MPA and less frequently in patients with EGPA, but may also be a typical feature with GBM (glomerular basement membrane) disease.
Cardiovascular system	Loss of pulses, ischemic cardiac pain, cardiomyopathy, congestive cardiac failure, valvular heart disease, pericarditis	Pericarditis	Loss of pulses, bruits, blood pressure discrepancy, claudication, ischemic cardiac pain, cardiomyopathy, congestive cardiac failure, valvular heart disease, pericarditis	Cardiovascular manifestations are most widely recognized in KD, which is not particularly well covered by PVAS. Although CVD manifestations do occur in small- and medium-vessel vasculitis, they are more typically seen in large vessel diseases, such as Takayasu arteritis.
Abdominal system	Ischemic abdominal pain, peritonitis, bloody diarrhea	Mesenteric ischemia	Abdominal pain, peritonitis, bloody diarrhea, bowel ischemia	Gut involvement is more typical in medium-vessel vasculitis, especially polyarteritis nodosa, but is well recognized in patients with GPA especially with colitis giving rise to bloody diarrhea.
Renal system	Hypertension, proteinuria, hematuria, impaired renal function, deterioration in renal function	Hematuria, red cell casts, or glomerulonephritis, deterioration in renal function	Hypertension, proteinuria, hematuria, impaired renal function, deterioration in renal function	Renal involvement in small-vessel vasculitis is one of the major manifestations leading to organ failure and death and should be carefully assessed. Renal involvement in medium-vessel vasculitis is much less common and takes the form of infarction of segments of the kidney leading to hematuria and hypertension with consequent impairment of renal function.
Nervous system	Headache, meningitis, organic confusion, seizures, stroke, cord involvement, cranial nerve, lesion, sensory or motor neuropathies	Meningitis, stroke, cord involvement, cranial nerve lesion, sensory or motor neuropathies	Headache, meningitis, organic confusion, seizures, stroke, cord involvement, cranial nerve, lesion, sensory or motor neuropathies	Neurological involvement is a common feature in small- and medium-vessel vasculitis; often it does not lead to immediate loss of life. Strokes are less common, whereas peripheral neuropathies are more common and can cause long-term disability.

No Treatment/Symptom Relief

Small-vessel vasculitis, such as isolated cutaneous vasculitis resulting from infection or use of pharmaceutical agents, may respond to simple withdrawal of the offending agent or resolution of the infection without the need for specific treatment. However, for more recalcitrant disease, symptom relief may be required and occasionally systemic steroids. Symptom relief could be provided in the form of antipruritic agents or topical cream to reduce skin inflammation and/or topical steroids. NSAIDs can be helpful in relieving symptoms of joint pain or swelling. As monotherapy they are unlikely to resolve skin manifestations but could be tried in combination with other therapies. Colchicine has been used for skin vasculitis and occasionally can be effective, although the doses required should remain below 2 mg/day to avoid the predictable side effects of abdominal cramps and diarrhea.

Target-Directed Therapies

In diseases where there is a clearly defined provoking agent, as in hepatitis B–related PAN or hepatitis C–related cryoglobulinemic vasculitis, eradication of the virus is a key part of treatment of the disease. Effective antiviral agents play a vital role in the management of virus, associated with the need for immunosuppression. Hepatitis B–related PAN is treated with a combination of antiviral therapy plus plasma exchange to remove immune-complexes and other inflammatory mediators combined with a course of glucocorticoid therapy. For hepatitis C–related cryoglobulinemic vasculitis, recent reports of virus eradication may also be transforming the outcome of this disease. Unfortunately, the toxicity of these regimens can be considerable, with >40% requiring erythropoietin, red blood cell transfusions, and or G-CSF.

Specific Therapies

In KD, although the etiological factors have not yet been defined, it seems likely that this is related to some kind of infectious agent (see earlier section on pathogenesis). The most effective therapy is use of high doses of IVIG (0.4 g/kg/day for 5 days) combined with high doses of aspirin, which is usually curative. Whether or not this will prevent long-term harm to the cardiovascular system, particularly the coronary arteries, remains to be explored.

Glucocorticoids

Glucocorticoids (Chapter 86) remain a cornerstone in the management of most forms of multisystem vasculitis. They are relatively contraindicated in KD because they may potentially worsen the development of coronary artery aneurysm, but they have been used in combination with IVIG and aspirin with beneficial outcomes. However, they are an integral part of almost all therapeutic regimens for management of vasculitis. In some instances, such as isolated skin vasculitis, they are the only treatment required, but more often they are insufficient on their own without causing significant morbidity from side effects. Typical doses of glucocorticoid therapy are 1 mg/kg/day over a period of 2–4 weeks, reducing to around 10–15 mg/day within 6 months, and then slowly withdrawing steroids in the next 6–12 months. The use of high-dose intravenous methylprednisolone is popular but lacks evidence. The only randomized trial of intravenous methylprednisolone compared it against plasma exchange in patients with severe AAV.[80] This study demonstrated the superiority of plasmapheresis over IV prednisolone, both used as adjunct therapies (in combination with cyclophosphamide and high doses of oral prednisolone) for treatment of severe AAV (mainly MPA plus some cases of GPA) with significant renal impairment (creatinine levels above 500 µmol/ L [5.66 mg/dL]).

The role of glucocorticoid therapy is been increasingly challenged by more recent trials using smaller doses for shorter periods or even eliminating steroid use completely in some instances. Glucocorticoid side effects are well known: weight loss, increased appetite, mood swings, hypertension, risk of diabetes, risk of infection, risk of cataract, and skin striae. The risk of osteoporosis is largely preventable with concurrent use of bisphosphonate therapy (unless there is significant renal dysfunction) with supplementary calcium and vitamin D replacement.

Other Immunosuppressive Therapies

Cyclophosphamide (Chapter 87) has been available since the 1950s but was first used for the management of systemic vasculitis in the 1970s and remains the most effective agent we have for managing multiorgan systemic vasculitis. Initially it was used as a daily oral agent at 2–3 mg/day, and it transformed the outcome of patients with AAV from inevitable mortality to a high likelihood of survival. It is a cytotoxic agent and carries with it the risk associated with chemotherapy, including increased risk of malignancy, especially in the bladder because it is predominantly excreted through the kidneys and accumulates in the bladder. Initial protocols were associated with excessive risks of bladder cancer (approximately 33-fold), but despite this, the daily oral cyclophosphamide dosing regimen was not effective in maintaining control of disease. Therefore over the past 20 years or so, there have been a number of trials comparing reduced doses of cyclophosphamide, high-dose intermittent pulse therapy,[81] or with combination strategies of induction with short courses of cyclophosphamide followed by a switch to another drug for maintenance,[81] or by replacing cyclophosphamide with another agent, such as methotrexate. All these studies[81] have demonstrated the equivalence of using shorter courses of daily oral cyclophosphamide and, more recently, of using of high-dose intermittent pulses of cyclophosphamide to reduce the total dose even further. The total cumulative dose from six cycles of cyclophosphamide over a period of 3 months would be 6 g (based on 15 mg/kg/treatment on six occasions). This compares with 9–12 g of cyclophosphamide given as daily oral therapy over 4–6 months.[81]

Although the relapse rate for patients given high-dose intermittent cyclophosphamide was higher during the subsequent 5 years compared with that for patients who were not given daily oral cyclophosphamide, relapse was always effectively managed with reintroduction of therapy and never led to mortality.[78] The consequences of exposure to cyclophosphamide are likely to be risk of cancer (more recent studies suggest that this risk is relatively modest now that the regimens include much lower total doses), infertility, hair loss, nausea, vomiting, diarrhea, cytopenia, and increased risk of infection. Less common complications of cyclophosphamide include hyponatremia. The introduction of rituximab for AAV has had a significant impact on the use of cyclophosphamide, with increasing numbers of patients being managed with rituximab in place of cyclophosphamide, especially if patients are of child-bearing years, or if there is a potential contraindication to using cyclophosphamide, such as a previous history of bladder cancer.

Because of the nature of AAV, relapse is common. Therefore a single course of therapy rarely achieves long-lasting remission. Repeat cycles of treatment are likely to be required, which accounts for the accumulation of higher doses of cyclophosphamide, especially in the pre-rituximab era. Therefore although each individual course of therapy may only contain 6–9 g, during a patient's life time, they may require treatment for several relapses, which would then start building up the total exposure to cyclophosphamide. Nevertheless, even though it is being slowly replaced, it remains an important aspect in the therapy of vasculitis even in the present time.

Azathioprine is an immunomodulator with cytostatic properties. It inhibits cell division; it has been an effective immunosuppressant agent for decades. It was first used in combination with steroids and found to reduce mortality in systemic vasculitis in an open-label retrospective study of 64 patients.[82] The 5-year survival of patients given no therapy was 12%, and those given steroids alone had a 53% survival rate, whereas those treated with steroids plus another agent (mainly azathioprine but a few had cyclophosphamide) had a survival rate of 80%. It is an oral medication given at 2–2.5 mg/kg/day, and it has largely been superseded as an induction agent by cyclophosphamide. Azathioprine is usually now used as maintenance therapy once the disease has been controlled with another agent. It has been shown to be equivalent to methotrexate and superior to mycophenolate[76] as maintenance therapy. It is a relatively safe immunosuppressant and can be used safely throughout pregnancy.

Methotrexate (MTX) is popular among rheumatologists but less so among renal physicians who regard it with some suspicion because of its potential for nephrotoxicity. The latter is only the case for patients who have well-established renal disease (typically with a creatinine level >300 μmol/L). MTX is an effective immunosuppressant agent used very widely in the management of inflammatory arthritis and has found its place in the management of GPA, where studies have demonstrated its efficacy in comparison to oral cyclophosphamide.[78] However, it needs to be given continuously, rather than as induction regimen over the short period of time. Although MTX is as effective as cyclophosphamide in inducing remission in GPA, stopping the drug inevitably leads to relapse. MTX is recommended for non–life-threating AAV, usually in combination with steroid treatment. MTX is available either as oral, intramuscular, or subcutaneous administration. The dose is 20–25 mg/week in most studies of vasculitis. It is contraindicated in pregnancy.

Cotrimoxazole, an antibiotic containing sulfonamide and trimethoprim, was fortuitously discovered to have beneficial effects in patients with GPA who had been treated coincidentally for infections. There has been a significant advance in our understanding of GPA, partly as a result of this historical experiment and partly based on suggestions that *S. aureus* plays a role in initiating disease by its effect on the nasal mucosa. Eradication of the organism has been suggested as one mechanism by which it works, although the drug is not particularly effective against this organism. It is more likely that the drug has an immunosuppressive effect in itself; it has been demonstrated to be effective in combination with low-dose steroids in a randomized trial of localized forms of GPA.[78] It is also commonly used as a prophylactic agent against *Pneumocystis jiroveci* infection; it is given 3 times per week for patients receiving other more potent immunosuppressant therapy, such as cyclophosphamide or even MTX (despite the potential for drug interactions leading to anemia).

Mycophenolate mofetil is a widely used transplantation drug that has been tested in AAV but is less effective than azathioprine as maintenance agent for patients who have achieved remission. It is currently being tested against cyclophosphamide as an induction agent in AAV. Is typically given as 2–3 g/day as an oral dose, along with reducing courses steroids. It is contraindicated in pregnancy.

Cyclosporine is a well-established immunosuppressive drug that has been used for transplantation for many decades. Cyclosporine has been given to limited numbers of patients with systemic vasculitis, with one small well-conducted trial[82] in 32 patients with GPA. In combination with plasmapheresis, it was as effective as continuous oral cyclophosphamide as a maintenance agent. It is generally limited by its toxicity and is not routinely used.

Leflunomide, an antilymphocyte agent used extensively for the management of inflammatory arthritis, has been tested in patients with AAV in limited trials demonstrating its ability to maintain remission. Indeed, in a recent meta-analysis, it was reported to be superior to azathioprine, MTX, and mycophenolate as a maintenance agent for AAV, but more trial data are needed. It is an oral agent that is characterized by a very long half-life, and it is not suitable for use in pregnancy.

The role of hydroxychloroquine in mild forms of vasculitis is uncertain. There is anecdotal evidence that it is beneficial in patients with skin manifestations with small-vessel vasculitis. Its use probably stems from its known effects in the treatment of connective tissue diseases with skin and joint manifestations, such as SLE. Other similar agents, such as mepacrine and dapsone, have also been used for skin vasculitis with occasional reported positive outcomes. However, the potential toxicity for each of these agents should be considered alongside the relatively limited evidence for benefit.

Specific Immunotherapy

Better understanding of the pathogenesis of vasculitis (see section on pathogenesis) has led to the development of targeted immunotherapy for some of these diseases. Rituximab is an mAb against B cells (Chapter 89) and has been in widespread use for treatment of RA. Its role as an effective agent in AAV is well established with two randomized trials demonstrating efficacy comparable with cyclophosphamide.[78] Rituximab induction therapy is just as effective as cyclophosphamide for moderate and moderate-to-severe AAV. Maintenance therapy with rituximab is a real possibility with the potential to provide long-term control and reduce relapse risk. The long-term consequences of repeat cycles of rituximab, however, are unexplored. The risks are hypogammaglobulinemia, which occurs in most cases, and the potential increase in incidence of infections. The most feared complication is the risk of reactivation of the JC virus leading to the complication of progressive multifocal leukoencephalopathy (PML), which has a very high mortality rate.

Other Therapies

Belimumab is currently undergoing clinical trials as a maintenance agent for AAV. The drug is a BAFF inhibitor and may be an effective means to control the disease over the long term. Mepolizumab is an mAb directed against IL-5, which controls eosinophil production. Mepolizumab has been successfully used in the treatment of hypereosinophilic syndrome and is undergoing a trial in EGPA. IVIG (Chapter 84) has been available as a replacement therapy for patients with hypogammaglobulinemia for several years, and it is the standard of care for KD, but its use in AAV has been limited. An initial study suggested benefit, but the likelihood is relatively short-term control.[82] Plasmapheresis has been available for several decades. It is not entirely clear how it works; there are many theories suggesting the removal of circulating immune mediators is effective in reducing inflammation. It has been successful for rescue therapy in patients with very aggressive AAV or anti-GBM disease. The MEPEX trial demonstrated that it was able to reduce renal dysfunction in patients with severe AAV.[78] However, the long-term follow-up of the MEPEX trial demonstrated that the difference between patients treated with plasma exchange and those given methylprednisolone pulses (both as adjunct treatment alongside cyclophosphamide and steroid) did not last.[83] It has been suggested this is accounted for by the fact that many patients with severe renal disease had already developed irreversible changes to their kidneys and that renal dysfunction would have been secondary to damage rather than active disease. A new study is underway to test the role of plasma exchange in patients with less severe renal disease, but results are not yet available. Plasmapheresis is effective in conjunction with antiviral therapy and steroids in the management of hepatitis B–related PAN.

OUTCOMES

The majority of patients have a successful initial outcome. Either the condition is self-limiting in the case of more isolated forms of skin vasculitis or the initial immunosuppressive therapy is successful. Over 94% of patients with generalized AAV would expect to survive the first 18 months,[78] whereas patients with more severe disease, especially with significant renal impairment, have a mortality of around 25% after 2 years.[82] This contrasts with >80% likelihood of dying without adequate treatment. However, with current therapy, 5-year survival figures suggest around 25–30% mortality among AAV-affected patients.[84]

The bigger problem, however, is morbidity. The quality of survival for most patients with multiorgan disease is complicated by episodes of relapse in 50–70% of cases and low-grade grumbling disease, which never quite goes into full remission in up to a third of cases.[85] This is added to the comorbidity experienced by older patients, usually a combination of vasculitis-related damage, steroid-induced side effects, and the long-term consequences of immunosuppressive agents. In the first year of diagnosis, the most likely cause of mortality is active vasculitis or infection,[62] the latter being a surrogate measure of the severity in immunosuppressant therapy required to control the disease.

Long-term adverse outcomes in vasculitis can be measured using a structured VDI (see Assessment section above). One of the most important outcomes is the development of end-stage renal failure and the requirement for dialysis. It is likely that this is significantly reduced as a result of effective therapy given within the first 4 months of diagnosis. Transplantation is successful in patients with AAV, and these patients should be offered this treatment. Ten-year survival rates (32.5%) are similar to those reported for other patients without diabetes receiving a kidney transplant.[86] The immunosuppressive regimens used for maintaining the transplant (Chapter 81) are often sufficient to keep the vasculitis in remission, but there is a need for ongoing review.

Infection is a significant concern, especially in the early course of disease when potent treatment being commenced, especially high doses of steroids. The risk of serious infection requiring hospitalization is very high in the first year,[62] especially if the steroid doses remain high after 6 months. Interstitial lung disease (nonspecific interstitial pneumonia) is reported in around 20% of Japanese patients, especially those with MPA. This is a higher prevalence than that seen in other populations and may reflect genetic and environmental differences unique to Japan. Chronic neuropathy occurs in 15% of patients with AAV[87] and can be very distressing for patients. Upper airway disease generally continues to cause long-term problems in 65% of patients with GPA because of chronic mucosal damage causing symptoms of chronic nasal congestion, discharge, and discomfort.[88] Symptom relief is only partially successful in alleviating these problems. The risk of cardiovascular disease among patients with small-vessel vasculitis is probably around four times greater in patients with AAV compared with the general population. The risk of cardiovascular events is around 14% within 5 years of diagnosis, especially in older patients who have baseline hypertension and MPO antibodies.[89] Cancer is associated with the presence of small-vessel vasculitis. It may predate as well as occur at the same time as diagnosis or develop subsequently, but it is recognized as a risk among patients treated with immunosuppressive and cytotoxic agents. Cancer of the bladder particularly has been an established risk arising from treatment with cyclophosphamide for many years; the original data from the 1970s suggested up to 33-fold increased risk of bladder cancer among patients treated with cyclophosphamide for vasculitis compared with background controls. However, this risk has been reduced with the use of more limited courses of cyclophosphamide (typically 3–6 months duration) and particularly with use of intermittent cyclophosphamide delivery. In a recent large series from the European Vasculitis Study Group (EUVAS), the only increased risk of cancer was for nonmelanoma skin cancer, and this may also have reflected the use of azathioprine as well as use of cyclophosphamide.[55] Shang et al.[90] showed that in a meta-analysis of over 2500 patients, the standardized incidence rate of late-occurring malignancies, particularly nonmelanoma skin cancer, leukemia, and bladder cancer, was 1.74 (95% CI = 1.37–2.21). We advise all of our patients to wear sun protection.

The ability to work is significantly affected by AAV; among 410 patients interviewed, 26% of those of working age were classified as work disabled.[91] The strongest influences on this outcome were fatigue, depression, high levels of damage (measured using the VDI), and being overweight. Patients' functional outcome can be variably affected by vasculitis and its treatment. Patients report impairment of function as measured by using generic tools, such as EQ-5D or the short form 36.[92] The impairment is similar to that found in other chronic diseases. Physical functions tend to be more affected than mental functions, especially in older patients with evidence of neurological involvement, usually peripheral neuropathy. Functional outcome is not directly correlated to disease activity, although in a Japanese cohort, 18 months after initiation of therapy, many aspects of function had started to improve.[93] One of the problems of determining long-term outcomes in patients with vasculitis is the compounding effect of the very intensive immunosuppressive therapy required to control disease. Over the last 3 decades, we have seen dramatic shift away from long course cyclophosphamide toward short courses of intermittent dose therapy, but we are now witnessing an era when targeted biological therapies are able to take the place of cyclophosphamide. Therefore eliminating the use of cyclophosphamide altogether in some patients may reset potential future outcomes. If this is coupled with a reduced use of glucocorticoid therapies and maintaining better disease control with less frequent relapses, the outcome may well be improved considerably for patients in the future.

ON THE HORIZON

- Improved understanding of the underlying pathogenic mechanisms in different forms of vasculitis is starting to lead to more logical, rational therapies, some of which have become established as standard of care or should be available in the near future.
- Rituximab is available and is being increasingly used in place of cyclophosphamide in the management of ANCA-associated vasculitis (AAV). Rituximab therapy is also effective in patients with cryoglobulinemic vasculitis, even when associated with the presence of hepatitis C. There is a concern that this might lead to reactivation of the virus and increase the risk of hepatoma or induce the development of the B-cell clone. However, better control of hepatitis C with effective antiviral agent is now possible.
- Multiple proinflammatory cytokines are being explored as therapeutic targets, although this will only allow for the suppression of common pathways in the end stage of inflammation.
- Mepolizumab has recently been shown to be effective in the treatment of eosinophilic granulomatous polyangiitis (EGPA) in a large randomised controlled trial.
- Blocking interleukin (IL)-6 may be of benefit for some patients.
- The discovery of the involvement of the complement pathway (C5) in AAV has led to the testing of therapies directed against C5a in humans.
- The ultimate goal remains to identify upstream pathologies leading to vasculitis, which promises to fundamentally change the management of these devastating diseases, as it may permit the induction and maintenance of drug-free remission.

ACKNOWLEDGEMENTS

The authors thank Jana Vaskova for administrative and secretarial support.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Kawasaki T, et al. A new infantile acute febrile mucocutaneous lymph node syndrome (MLNS) prevailing in Japan. Pediatrics 1974;54(3):271–6.
2. Andrews M, Edmunds M, Campbell A, et al. Systemic vasculitis in the 1980s-is there an increasing incidence of Wegener's granulomatosis and microscopic polyarteritis? J R Coll Physicians 1990;24(4):284–8.
3. Watts RA, Mooney J, Skinner J, et al. The contrasting epidemiology of granulomatosis with polyangiitis (Wegener's) and microscopic polyangiitis. Rheumatology (Oxford) 2012;51(5):926–31.
4. Mohammad AJ, et al. Prevalence of Wegener's granulomatosis, microscopic polyangiitis, polyarteritis nodosa and Churg-Strauss syndrome within a defined population in southern Sweden. Rheumatology 2007;46(8):1329–37.
5. Nesher G, et al. The Incidence of Primary Systemic Vasculitis in Jerusalem: A 20-year Hospital-based Retrospective Study. J Rheumatol 2016;43(6):1072–7.
6. Holman RC, et al. Racial/ethnic differences in the incidence of Kawasaki syndrome among children in Hawaii. Hawaii Med J 2010;69(8):194–7.
7. Arora A, et al. Incidence of Leukocytoclastic Vasculitis, 1996 to 2010: A Population-Based Study in Olmsted County, Minnesota. Mayo Clinic Proceedings 2014;89(11):1515–24.
8. Watts RA, Scott DGI. Epidemiology of the vasculitides. Semin Respir Crit Care Med 2004;25(5):455–64.
9. Nakamura Y, et al. Epidemiologic Features of Kawasaki Disease in Japan: Results of the 2009-2010 Nationwide Survey. J Epidemiol 2012;22(3):216–21.
10. Reinhold-Keller E, et al. Stable incidence of primary systemic vasculitides over five years: Results from the German vasculitis register. Arthritis Rheum 2005;53(1):93–9.
11. Herlyn K, et al. Doubled prevalence rates of ANCA-associated vasculitides and giant cell arteritis between 1994 and 2006 in northern Germany. Rheumatology 2014;53(5):882–9.
12. van der Woude FJ, et al. Autoantibodies against neutrophils and monocytes: tool for diagnosis and marker of disease activity in Wegener's granulomatosis. Lancet 1985;1(8426):425–9.
13. Xiao H, et al. Antineutrophil cytoplasmic autoantibodies specific for myeloperoxidase cause glomerulonephritis and vasculitis in mice. J Clin Invest 2002;110(7):955–63.
14. Primo VC, et al. Anti-PR3 immune responses induce segmental and necrotizing glomerulonephritis. Clin Exp Immunol 2010;159(3):327–37.
15. Little MA, et al. Anti-proteinase 3 anti-neutrophil cytoplasm autoantibodies recapitulate systemic vasculitis in mice with a humanized immune system. PLoS ONE 2012;7(1):e28626.
16. Roth AJ, et al. Epitope specificity determines pathogenicity and detectability in ANCA-associated vasculitis. J Clin Invest 2013;123(4):1773–83.
17. Lyons PA, et al. Genetically distinct subsets within ANCA-associated vasculitis. N Engl J Med 2012;367(3):214–23.
18. Stassen PM, et al. HLA-DR4, DR13(6) and the ancestral haplotype A1B8DR3 are associated with ANCA-associated vasculitis and Wegener's granulomatosis. Rheumatology (Oxford) 2009;48(6):622–5.
19. Cao Y, et al. DRB1*15 allele is a risk factor for PR3-ANCA disease in African Americans. J Am Soc Nephrol 2011;22(6):1161–7.
20. Papiha SS, et al. Association of Wegener's granulomatosis with HLA antigens and other genetic markers. Ann Rheum Dis 1992;51(2):246–8.
21. Spencer SJ, et al. HLA class II specificities in vasculitis with antibodies to neutrophil cytoplasmic antigens. Kidney Int 1992;41(4):1059–63.
22. Schreiber A, et al. Membrane expression of proteinase 3 is genetically determined. J Am Soc Nephrol 2003;14(1):68–75.
23. Alberici F, Martorana D, Vaglio A. Genetic aspects of anti-neutrophil cytoplasmic antibody-associated vasculitis. Nephrol Dial Transplant 2015;30(Suppl. 1):i37–45.
24. Kamesh L, et al. CT60 and +49 polymorphisms of CTLA 4 are associated with ANCA-positive small vessel vasculitis. Rheumatology (Oxford) 2009;48(12):1502–5.
25. Langford CA, et al. An open-label trial of abatacept (CTLA4-IG) in non-severe relapsing granulomatosis with polyangiitis (Wegener's). Ann Rheum Dis 2014;73(7):1376–9.
26. Wieczorek S, et al. Functionally relevant variations of the interleukin-10 gene associated with antineutrophil cytoplasmic antibody-negative Churg-Strauss syndrome, but not with Wegener's granulomatosis. Arthritis Rheum 2008;58(6):1839–48.
27. Ciavatta DJ, et al. Epigenetic basis for aberrant upregulation of autoantigen genes in humans with ANCA vasculitis. J Clin Invest 2010;120(9):3209–19.
28. Chen M, Kallenberg CG. The environment, geoepidemiology and ANCA-associated vasculitides. Autoimmun Rev 2010;9(5):A293–8.
29. Gomez-Puerta JA, Gedmintas L, Costenbader KH. The association between silica exposure and development of ANCA-associated vasculitis: systematic review and meta-analysis. Autoimmun Rev 2013;12(12):1129–35.
30. Tervaert JW, Stegeman CA, Kallenberg CG. Silicon exposure and vasculitis. Curr Opin Rheumatol 1998;10(1):12–17.
31. Rihova Z, et al. Silica and asbestos exposure in ANCA-associated vasculitis with pulmonary involvement. Ren Fail 2005;27(5):605–8.
32. Eden A, et al. Lack of association between B19 or V9 erythrovirus infection and ANCA-positive vasculitides: a case-control study. Rheumatology (Oxford) 2003;42(5):660–4.
33. Wu YY, et al. Proteinase 3 and dihydrolipoamide dehydrogenase (E3) are major autoantigens in hepatitis C virus (HCV) infection. Clin Exp Immunol 2002;128(2):347–52.

34. Calhan T, et al. Antineutrophil cytoplasmic antibody frequency in chronic hepatitis B patients. Dis Markers 2014;2014:982150.
35. Kain R, et al. Molecular mimicry in pauci-immune focal necrotizing glomerulonephritis. Nat Med 2008;14(10):1088–96.
36. Hurtado PR, et al. CpG oligodeoxynucleotide stimulates production of anti-neutrophil cytoplasmic antibodies in ANCA associated vasculitis. BMC Immunol 2008;9:34.
37. Noh JY, Yasuda S, Sato S, et al. Clinical characteristics of myeloperoxidase antineutrophil cytoplasmic antibody-associated vasculitis caused by antithyroid drugs. J Clin Endocrinol Metab 2009;94(8):2806–11.
38. Waldhauser L, Uetrecht J. Oxidation of propylthiouracil to reactive metabolites by activated neutrophils. Implications for agranulocytosis. Drug Metab Dispos 1991;19(2):354–9.
39. Nakazawa D, et al. Abnormal conformation and impaired degradation of propylthiouracil-induced neutrophil extracellular traps: implications of disordered neutrophil extracellular traps in a rat model of myeloperoxidase antineutrophil cytoplasmic antibody-associated vasculitis. Arthritis Rheum 2012;64(11):3779–87.
40. Berman M, Paran D, Elkayam O. Cocaine-Induced Vasculitis. Rambam Maimonides Med J 2016;7(4).
41. Wiesner O, et al. Antineutrophil cytoplasmic antibodies reacting with human neutrophil elastase as a diagnostic marker for cocaine-induced midline destructive lesions but not autoimmune vasculitis. Arthritis Rheum 2004;50(9):2954–65.
42. Tsurikisawa N, et al. Decreases in the numbers of peripheral blood regulatory T cells, and increases in the levels of memory and activated B cells, in patients with active eosinophilic granulomatosis and polyangiitis. J Clin Immunol 2013;33(5):965–76.
43. Abdulahad WH, Boots AMH, Kallenberg CGM. FoxP3(+) CD4(+) T cells in systemic autoimmune diseases: the delicate balance between true regulatory T cells and effector Th-17 cells. Rheumatology 2011;50(4):646–56.
44. Marinaki S, et al. Persistent T-cell activation and clinical correlations in patients with ANCA-associated systemic vasculitis. Nephrol Dial Transplant 2006;21(7):1825–32.
45. Holden NJ, et al. ANCA-stimulated neutrophils release BLyS and promote B cell survival: a clinically relevant cellular process. Ann Rheum Dis 2011;70(12):2229–33.
46. Nakazawa D, et al. Enhanced Formation and Disordered Regulation of NETs in Myeloperoxidase-ANCA-Associated Microscopic Polyangiitis. J Am Soc Nephrol 2014;25(5):990–7.
47. Yuan J, et al. C5a and its receptors in human anti-neutrophil cytoplasmic antibody (ANCA)-associated vasculitis. Arthritis Res Ther 2012;14(3).
48. Chen SF, et al. Plasma complement factor H is associated with disease activity of patients with ANCA-associated vasculitis. Arthritis Res Ther 2015;17.
49. Fukui S, et al. Antineutrophilic cytoplasmic antibody-associated vasculitis with hypocomplementemia has a higher incidence of serious organ damage and a poor prognosis. Medicine (Baltimore) 2016;95(37).
50. Choi HK, Merkel PA, Niles JL. ANCA-positive vasculitis associated with allopurinol therapy. Clin Exp Rheumatol 1998;16(6):743–4.
51. Hirohama D, et al. Development of myeloperoxidase-antineutrophil cytoplasmic antibody-associated renal vasculitis in a patient receiving treatment with anti-tumor necrosis factor-alpha. Mod Rheumatol 2010;20(6):602–5.
52. Jarraya F, et al. Myeloperoxidase-antineutrophil cytoplasmic antibody-positive crescentic glomerulonephritis associated with benzylthiouracil therapy: report of the first case. Nephrol Dial Transplant 2003;18(11):2421–3.
53. Bienaime F, et al. D-Penicillamine-induced ANCA-associated crescentic glomerulonephritis in Wilson disease. Am J Kidney Dis 2007;50(5): 821–5.
54. Carmona-Rivera C, Purmalek MM, Moore E, et al. A role for muscarinic receptors in neutrophil extracellular trap formation and levamisole-induced autoimmunity. JCI Insight 2017;2(3):e89780.
55. Heijl C, Harper L, Flossmann O, et al. Incidence of malignancy in patients treated for antineutrophil cytoplasm antibody-associated vasculitis: follow-up data from European Vasculitis Study Group clinical trials. Ann Rheum Dis 2011;70(8):1415–21.
56. Guma M, et al. Frequency of antineutrophil cytoplasmic antibody in Graves' disease patients treated with methimazole. J Clin Endocrinol Metab 2003;88(5):2141–6.
57. Lenert P, Icardi M, Dahmoush L. ANA (+) ANCA (+) systemic vasculitis associated with the use of minocycline: case-based review. Clin Rheumatol 2013;32(7):1099–106.
58. Zhang AH, et al. Inhibition of oxidation activity of myeloperoxidase (MPO) by propylthiouracil (PTU) and anti-MPO antibodies from patients with PTU-induced vasculitis. Clin Immunol 2007;122(2): 187–93.
59. Denissen NH, et al. Can sulfasalazine therapy induce or exacerbate Wegener's granulomatosis? Scand J Rheumatol 2008;37(1):72–4.
60. Takahashi K, Oharaseki T, Yokouchi Y. Update on etio and immunopathogenesis of Kawasaki disease. Curr Opin Rheumatol 2014;26(1):31–6.
61. Lin YJ, et al. Genetic variants in PLCB4/PLCB1 as susceptibility loci for coronary artery aneurysm formation in Kawasaki disease in Han Chinese in Taiwan. Sci Rep 2015;5:14762.
62. Onouchi Y, et al. ITPKC and CASP3 polymorphisms and risks for IVIG unresponsiveness and coronary artery lesion formation in Kawasaki disease. Pharmacogenomics J 2013;13(1):52–9.
63. Yeung RS. Kawasaki disease: update on pathogenesis. Curr Opin Rheumatol 2010;22(5):551–60.
64. Gorelik M, et al. Plasma follistatin-like protein 1 is elevated in Kawasaki disease and may predict coronary artery aneurysm formation. J Pediatr 2012;161(1):116–19.
65. Lin IC, et al. Augmented TLR2 expression on monocytes in both human Kawasaki disease and a mouse model of coronary arteritis. PLoS ONE 2012;7(6):e38635.
66. Jia S, et al. The T helper type 17/regulatory T cell imbalance in patients with acute Kawasaki disease. Clin Exp Immunol 2010;162(1):131–7.
67. Rowley AH, et al. The transcriptional profile of coronary arteritis in Kawasaki disease. BMC Genomics 2015;16:1076.
68. Leung DY, et al. Prevalence of superantigen-secreting bacteria in patients with Kawasaki disease. J Pediatr 2002;140(6):742–6.
69. Schulte DJ, et al. Involvement of innate and adaptive immunity in a murine model of coronary arteritis mimicking Kawasaki disease. J Immunol 2009;183(8):5311–18.
70. Cid MC, et al. Immunohistochemical characterization of inflammatory cells and immunologic activation markers in muscle and nerve biopsy specimens from patients with systemic polyarteritis nodosa. Arthritis Rheum 1994;37(7):1055–61.
71. Yalcinkaya F, et al. Prevalence of the MEFV gene mutations in childhood polyarteritis nodosa. J Pediatr 2007;151(6):675–8.
72. Sansonno D, et al. Hepatitis C virus infection, cryoglobulinaemia, and beyond. Rheumatology 2007;46(4):572–8.
73. Varricchi G, et al. Interleukin-5 pathway inhibition in the treatment of eosinophilic respiratory disorders: evidence and unmet needs. Curr Opin Allergy Clin Immunol 2016;16(2):186–200.
74. Craven A, et al. ACR/EULAR-endorsed study to develop Diagnostic and Classification Criteria for Vasculitis (DCVAS). Clin Exp Nephrol 2013;17(5):619–21.
75. Luqmani R, Ponte C. ANCA associated vasculitides and polyarteritis nodosa. In: Hachulla E, Bijlsma JWJ, editors. EULAR Textbook on rheumatic diseases. BMJ Publishing Group Ltd; 2015. p. 717–53.
75a. Stahelin L, et al. Cocaine-induced midline destruction lesions with positive ANCA test mimicking Wegener's granulomatosis. Rev Bras Reumatol 2012;52(3):431–7.
76. Flossmann O, Berden A, de Groot K, et al. Long-term patient survival in ANCA-associated vasculitis. Ann Rheum Dis 2011;70(3):488–94.
77. Kemna MJ, Damoiseaux J, Austen J, et al. ANCA as a predictor of relapse: useful in patients with renal involvement but not in patients with nonrenal disease. J Am Soc Nephrol 2015;26(3):537–42.
78. Ntatsaki E, Carruthers D, Chakravarty K, et al. BSR and BHPR Standards, Guidelines and Audit Working Group BSR and BHPR guideline for the

management of adults with ANCA-associated vasculitis. Rheumatology (Oxford) 2014;53(12):2306–9.

79. Ponte C, et al. Optimisation of vasculitis disease assessments in clinical trials, clinical care and long-term databases. Clin Exp Rheumatol 2014;32(5 Suppl. 85):S-118-25.
80. Mukhtyar C, G L, Cid MC, et al. BSR and BHPR Standards, Guidelines and Audit Working Group BSR and BHPR guideline for the management. Ann Rheum Dis 2009;68:310–17.
81. Szpirt WM, Heaf JG, Petersen J. Plasma exchange for induction and cyclosporine A for maintenance of remission in Wegener's granulomatosis-a clinical randomized controlled trial. Nephrol Dial Transplant 2011;26(1):206–13.
82. Mukhtyar C, Guillevin L, Cid MC, et al. EULAR recommendations for the management of primary small and medium vessel vasculitis. Ann Rheum Dis 2009;68(3):310–17.
83. Walsh M, Casian A, Flossmann O, Westman K, et al. Long-term follow-up of patients with severe ANCA-associated vasculitis comparing plasma exchange to intravenous methylprednisolone treatment is unclear. Kidney Int 2013;84(2):397–402.
84. Flossmann O, Berden A, de Groot K, Hagen C, et al. Long-term patient survival in ANCA-associated vasculitis. Ann Rheum Dis 2011;70(3):488–94.
85. Hoffman GS, Kerr G, Leavitt RY, et al. Wegener granulomatosis: an analysis of 158 patients. Ann Intern Med 1992;116:488–98.
86. Hruskova Z, Stel V, Jayne D, et al. Characteristics and Outcomes of Granulomatosis With Polyangiitis (Wegener) and Microscopic Polyangiitis Requiring Renal Replacement Therapy: Results From the European Renal Association-European Dialysis and Transplant Association Registry. Am J Kidney Dis 2015;66(4):613–20.
87. Suppiah R, Hadden R, Batra R, et al. Peripheral neuropathy in ANCA-associated vasculitis: outcomes from the European Vasculitis Study Group trials. Rheumatology (Oxford) 2011;50(12):2214–22.
88. Martinez Del Pero M, Walsh M, Luqmani R, et al. Long-term damage to the ENT system in Wegener's granulomatosis. Eur Arch Otorhinolaryngol 2011;268(5):733–9.
89. Suppiah R, Judge A, Batra R, et al. A model to predict cardiovascular events in patients with newly diagnosed Wegener's granulomatosis and microscopic polyangiitis. Arthritis Care Res (Hoboken) 2011b;63(4):588–96.
90. Shang W, Ning Y, Xu X, et al. Incidence of Cancer in ANCA-Associated Vasculitis: A Meta-Analysis of Observational Studies. PLoS ONE 2015;10(5):e0126016.
91. Basu N, McClean A, Harper L, et al. Markers for work disability in anti-neutrophil cytoplasmic antibody-associated vasculitis. Rheumatology (Oxford) 2014;53(5):953–6.
92. Walsh M, Mukhtyar C, Mahr A, et al. Health-related quality of life in patients with newly diagnosed antineutrophil cytoplasmic antibody-associated vasculitis. Arthritis Care Res (Hoboken) 2011;63(7):1055–61.
93. Suka M, Hayashi T, Kobayashi S, et al. Improvement in health-related quality of life in MPO-ANCA-associated vasculitis patients treated with cyclophosphamide plus prednisolone: an analysis of 18 months of follow-up data from the JMAAV study. Mod Rheumatol 2012;22(6):877–84.
94. Fujimoto SWR, Kobayashi S, Suzuki K, et al. Comparison of the epidemiology of anti-neutrophil cytoplasmic antibody-associated vasculitis between Japan and the U.K. Rheumatology (Oxford) 2011;50(10):1916–20.
95. Chen M, et al. Propylthiouracil-induced antineutrophil cytoplasmic antibody-associated vasculitis. Nat Rev Nephrol 2012;8(8):476–83.
96. Makino N, Nakamura Y, Yashiro M, et al. Descriptive epidemiology of Kawasaki disease in Japan, 2011-2012: from the results of the 22nd nationwide survey. J Epidemiol 2015;25(3):239–45.

MULTIPLE-CHOICE QUESTIONS

1. Which of the following types of vasculitis are usually associated with the presence of antineutrophil cytoplasmic antibody (ANCA)?
   A. Microscopic polyangiitis
   B. Polyarteritis nodosa
   C. Goodpasture syndrome
   D. Cryoglobulinemic vasculitis
   E. Granulomatosis with polyangiitis

2. Which of the following is the *least* common form of vasculitis in Europe?
   A. Kawasaki disease
   B. Polyarteritis nodosa
   C. Granulomatosis with polyangiitis
   D. Microscopic polyangiitis
   E. Leukocytoclastic cutaneous vasculitis

3. Which of the following treatments are *not* effective for use in patients with newly presenting generalized ANCA-associated vasculitis?
   A. Colchicine
   B. Glucocorticoid therapy alone
   C. Glucocorticoids and methotrexate
   D. Rituximab and glucocorticoid therapy
   E. Plasmapheresis with cyclophosphamide and glucocorticoid therapy

59

Large-Vessel Vasculitides

Cornelia M. Weyand

Most tissues have compensatory mechanisms that allow them to sustain the damaging effects of acute and chronic inflammation, but medium and large arteries are organs without redundancy and limited regenerative capacity. Life is unsustainable unless the major arteries have uncompromised function. Accordingly, inflammatory damage to such arterial vessels leads to severe clinical consequences, immediately posing a threat for the loss of function of vital organs. When affected by inflammation, the aorta and its branches have two possible response patterns: *(i)* Inflammatory destruction of the vessel wall leads to dilatation, aneurysm formation, and rupture. Alternatively, the wall layers dissect. *(ii)* The inflammation initiates a maladaptive response to injury, resulting in luminal occlusion, disruption of blood supply, and ischemic damage of dependent organ structures.

In contrast to other vasculopathies, especially those related to atherosclerosis, vasculitides of the larger blood vessels are almost always associated with a syndrome of intense systemic inflammation.[1] Recent evidence has challenged the traditional view that systemic inflammation represents a spillover of inflammatory mediators from the vasculitic lesions. Instead, systemic activation of the innate immune system appears to be a pinnacle event that initiates the processes leading to vessel wall inflammation. The coincidence of malaise, fever, wasting, and myalgias, with signs of ischemia caused by vascular failure, remains a critical clue for the physician when diagnosing and treating large-vessel vasculitides (LVVs).

The two major forms of LVVs are giant-cell arteritis (GCA) and Takayasu arteritis (TA). In addition, aortitis can infrequently be seen in other diseases, such as infections, connective tissue diseases, sarcoidosis, and inflammatory bowel disease (IBD), and occasionally is diagnosed as an idiopathic syndrome. Polymyalgia rheumatica (PMR) is a condition closely related to GCA; it occurs in the same patient population and often precedes or follows the clinical diagnosis of GCA.[2] Patients with PMR do not have typical vascular lesions; consequently, PMR is not classified as a vasculitis. However, patients with PMR do have a systemic inflammatory syndrome indistinguishable from GCA, and about 10% of patients with PMR eventually progress to full-blown vasculitis. Similarities in the vascular lesions of GCA and TA have been interpreted as revealing parallels in immunopathogenesis. Whether similarities in histomorphology and tissue targeting reflect similarities in underlying molecular defects remains unclear. Whether the systemic inflammatory reactions accompanying GCA, TA, and PMR have disease-specific elements also remains unanswered, but this has opened the possibility of developing biomarkers that are urgently needed for clinical monitoring. Excellent progress has been made in unraveling the pathogenesis of GCA, and this will inevitably lead to improvements in diagnosis, long-term management, and broadening of the therapeutic armamentarium.

EPIDEMIOLOGY

GCA may be a very old disease, as suggested by historic evidence that more than 1000 years ago removal of the temporal artery was recommended by a physician in Baghdad. In 1932, Horton et al. at the Mayo Clinic in Minnesota recognized that GCA was an inflammatory vasculopathy when they found dense inflammation in the temporal arteries of two patients who were systemically ill and had severe headaches. The first reports of TA, or "pulseless disease," in young women surfaced in Japan in the nineteenth century. The syndrome was named after Dr. Takayasu, an ophthalmologist, who, in 1905, described peculiar optic fundus abnormalities caused by ischemia-driven collateral formation.

The strongest risk factor for GCA, TA, and PMR is age.[3,4] GCA and PMR are essentially absent in individuals <50 years of age, and their incidence climbs continuously during the seventh and eighth decades of life. TA is almost exclusively diagnosed in individuals <40 years of age, with peak incidence during the second and third decades of life. Women are affected more often by all three syndromes compared with men, with a 2 : 1 ratio in PMR and GCA and a 9 : 1 ratio in TA.[3,5]

Marked geographical variations in the incidence and prevalence of GCA, TA, and PMR have given rise to speculations about environmental exposures as key factors in disease pathogenesis. GCA is the most frequent vasculitis in the Western world, with yearly incidence rates reaching 10–20 cases per 100 000 persons >50 years of age.[3] In general, PMR is diagnosed three- to fourfold more frequently, with a prevalence of up to 1 case per 133 individuals >50 of age.[2] Iceland, Norway, Sweden, and Denmark are high-risk areas; higher incidence rates are also seen in Scandinavian immigrant populations in the United States. The risk is significantly lower in Hispanics and African Americans. Although TA can afflict all races, a predilection exists for individuals of Asian and Central and South American origins. Japan, Thailand, India, Turkey, and nations in Central and South Americas are considered high-incidence regions. TA is a rare disease with an annual incidence of 1–2 cases/million. The typical patient is a female in her 20s to 30s. In middle-aged men and women, it can be challenging to differentiate TA from rapidly progressing atherosclerotic disease, especially as both disease processes may coexist.[6]

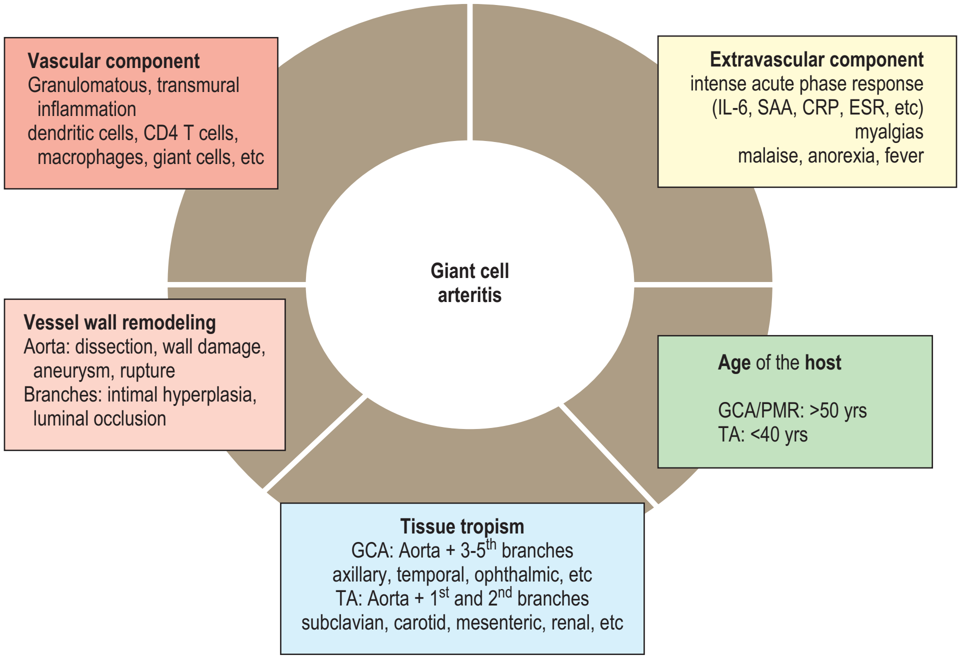

Vascular component
Granulomatous, transmural inflammation
dendritic cells, CD4 T cells, macrophages, giant cells, etc
Extravascular component
intense acute phase response (IL-6, SAA, CRP, ESR, etc)
myalgias
malaise, anorexia, fever
Giant cell arteritis
Vessel wall remodeling
Aorta: dissection, wall damage, aneurysm, rupture
Branches: intimal hyperplasia, luminal occlusion
Age of the host
GCA/PMR: >50 yrs
TA: <40 yrs
Tissue tropism
GCA: Aorta + 3-5th branches
axillary, temporal, ophthalmic, etc
TA: Aorta + 1st and 2nd branches
subclavian, carotid, mesenteric, renal, etc

FIG 59.1 Key Pathogenic Traits in Giant-Cell Arteritis (GCA). In the emerging pathogenic model, five hallmarks of disease have been identified that represent fundamental pathogenic traits and distinguish GCA from other immune-mediated disorders. Vascular GCA (granulomatous vasculitis) is now separated from extravascular GCA (systemic inflammation and intense hepatic acute phase response). Age remains the strongest risk factor with evidence for a signature immunosenescence program in patients with GCA. The stringent tissue tropism for selected vascular beds is suspected to reflect heterogeneity of vascular cells. Clinical complications are related to the patterning of vascular damage, spanning from wall destruction to luminal occlusion.

ETIOLOGY AND PATHOGENESIS

Abnormal innate and adaptive immune responses are essential pathogenic elements in medium-vessel vasculitides and LVVs. The end result is a disease process that separates GCA, PMR, and TA from other vasculitides and from other autoinflammatory and autoimmune syndromes. Progress has been made in identifying and characterizing the cellular and molecular elements that define LVVs. Recent work has delineated disease-specific signatures, with the goal to mark the pinnacle abnormality triggering the disease and to discover intersection points that lend themselves for diagnostic and therapeutic purposes.

Pathogenic studies and careful clinical observation have prompted a substantial shift in the pathogenic model (Fig. 59.1): LVVs are now understood to be chronic conditions that target restricted vascular beds and have two major disease components: *(i)* granulomatous, intramural inflammation inducing vascular wall remodeling; and *(ii)* extravascular inflammation manifesting with an intense acute phase response, myalgias, and constitutional symptoms. Emerging data suggest at least partial autonomy in the pathogenic cascades of both disease components, predicting separate biomarkers, separate pathogenic mechanisms, and separate responses to immunosuppression. Similarities in tissue tropisms and histological lesions of GCA and TA suggest overlapping disease pathways, and many aspects of immune pathogenesis may be transferrable from one syndrome to the other. The etiopathogenesis of PMR is less well understood, but experimental evidence suggests that it represents a *forme fruste* of GCA, in which inflammatory attack to the vessel wall remains below a threshold, and standard histology describes noninflamed arteries.

Distinctive features—hallmarks of disease—provide clues to the fundamental pathogenic mechanisms (see Fig. 59.1), including a stringent age cutoff, a stringent tissue tropism for selected vascular territories (aorta and major branches), an extravascular disease module defined by an abrupt and intense acute phase response, a granulomatous transmural vasculitis, and two opposing patterns of vessel wall remodeling causing aortic wall destruction versus luminal occlusion in the branch vessels.

INNATE IMMUNE SYSTEM DEFECTS

Innate immune cells make critical contributions to the pathogenesis of LVVs, but where circulating monocytes and neutrophils encounter their activating stimuli remains unknown (Fig. 59.2). Circulating monocytes and macrophages are highly activated[7] and contribute to the array of proinflammatory cytokines measurably elevated in the serum of affected patients.[8] Altered neutrophil functionality has also been described.[9] Early in the disease process, circulating neutrophils resemble those induced by a lipopolysaccharide stimulus, displaying T-cell suppressive functions. Interleukin-8 (IL-8) and IL-6 are associated with neutrophil activation. IL-6, first described in the early 1990s, is highly elevated in GCA and PMR[10] and is explicitly steroid sensitive.[11] IL-6 acts as an inducer of the acute phase response in the liver, triggering production of C-reactive protein (CRP),

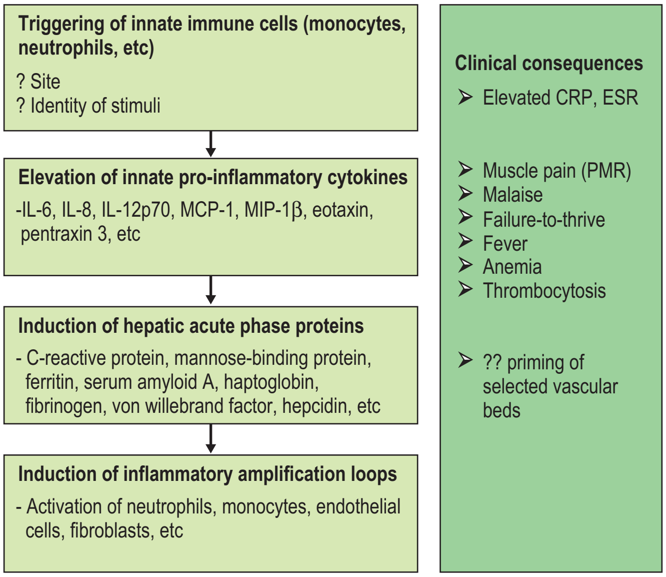

Triggering of innate immune cells (monocytes, neutrophils, etc)
? Site
? Identity of stimuli
Elevation of innate pro-inflammatory cytokines
-IL-6, IL-8, IL-12p70, MCP-1, MIP-1β, eotaxin, pentraxin 3, etc
Induction of hepatic acute phase proteins
- C-reactive protein, mannose-binding protein, ferritin, serum amyloid A, haptoglobin, fibrinogen, von willebrand factor, hepcidin, etc
Induction of inflammatory amplification loops
- Activation of neutrophils, monocytes, endothelial cells, fibroblasts, etc
Clinical consequences
➢ Elevated CRP, ESR
➢ Muscle pain (PMR)
➢ Malaise
➢ Failure-to-thrive
➢ Fever
➢ Anemia
➢ Thrombocytosis
➢ ?? priming of selected vascular beds

FIG 59.2 **Extravascular Giant-Cell Arteritis (GCA).** Circulating innate cells (monocytes, neutrophils, etc.) are highly activated and elicit a hepatic acute-phase response. Several of the acute phase proteins serve as biomarkers (C-reactive protein [CRP], erythrocyte sedimentation rate [ESR]) of disease. Extravascular GCA is explicitly sensitive to corticosteroid therapy and to cytokine blockade. Whether the extravascular competent has a direct or indirect involvement in driving/sustaining granulomatous vasculitis remains unknown. Polymyalgia rheumatic (PMR) may represent an isolated form of extravascular GCA.

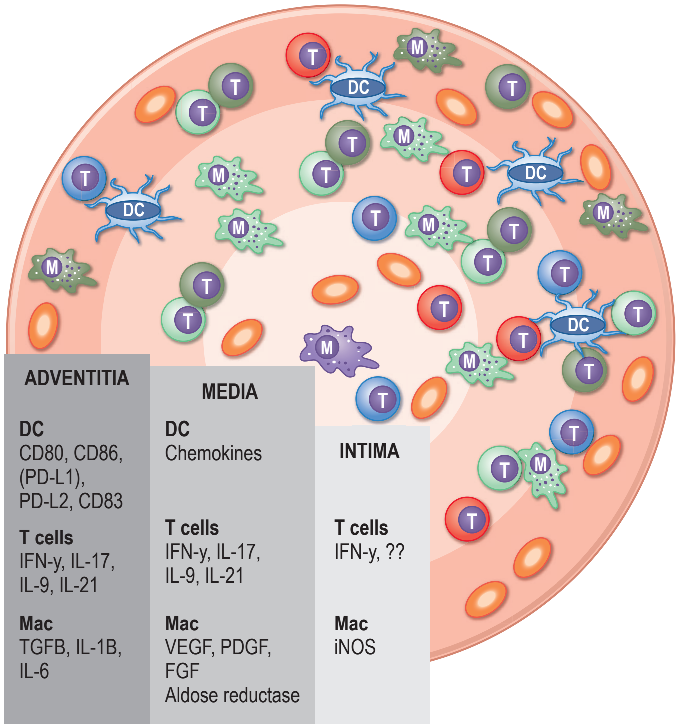

T
M
DC
ADVENTITIA
DC
CD80, CD86, (PD-L1), PD-L2, CD83
T cells
IFN-y, IL-17, IL-9, IL-21
Mac
TGFB, IL-1B, IL-6
MEDIA
DC
Chemokines
T cells
IFN-y, IL-17, IL-9, IL-21
Mac
VEGF, PDGF, FGF
Aldose reductase
INTIMA
T cells
IFN-y, ??
Mac
iNOS

FIG 59.3 **Vascular Giant-Cell Arteritis (GCA).** Adaptive immune responses in vascular GCA are supported by transmural granulomatous infiltrates. Dendritic cells (DCs), present in healthy arteries at the adventitia–media border, serve as antigen-presenting cells (APCs) and provide costimulatory signals but are deficient in the coinhibitory ligand programmed death ligand 1 (PD-L1). Wall-infiltrating T cells are multifunctional, producing a broad spectrum of effector cytokines, including interferon (IFN)-γ, interleukin (IL)-17, IL-9, and IL-21. Macrophages, some of which transform into multinucleated giant cells, supply a multitude of cytokines, metalloproteinases, reactive oxygen species (ROS), growth factors, angiogenesis factors and inducible nitric oxide synthase (iNOS). Their functional commitment depends on their geographical location in the vessel wall.

serum amyloid A (SAA), and several other acute-phase proteins. Accordingly, treatment with an IL-6 receptor–blocking antibody, effectively reduces the laboratory abnormalities measured in patients with GCA.[12] Whether IL-6 has additional functions in the disease process, particularly in the inflamed vessel wall, is uncertain. Some acute-phase proteins (*e.g.,* SAA) have proinflammatory functions by themselves, functioning as proinflammatory amplifiers.[13] Although abnormal innate immunity dominates the extravascular component of LVVs, underlying mechanisms (*e.g.,* what the original triggers are, where the activation occurs, how the systemic inflammation interfaces with vascular inflammation) are essentially unexplored.

Defects in innate immunity contributing to the early and late events in the vasculitic lesions are much better understood[14] (Fig. 59.3). A population of vessel wall–residing dendritic cells (vasDCs) have been identified in normal human arteries,[15] localized at the adventitial–medial junction, close to the adventitial microvascular network. Given their strategic positioning, they are believed to guard the artery's "backdoor." vasDCs serve a gatekeeper function in vasculitis, and stimulation of such vasDCs is a prerequisite to break the natural immunotolerance (immune privilege) of the arterial wall. Their activation presages vasculitis, as indicated by their transition from the resting state to the activated state in PMR arteries, where no vessel wall infiltrates are as yet detectable.[16] In frank vasculitis, vasDCs undergo expansion, penetrate deep into the wall, produce chemokines and cytokines, and express costimulatory ligands.[17] They have been implicated in presenting local antigen to induce T-cell clonal expansion and, through their chemokine production, shape the composition of the intramural granulomatous infiltrates. Considering that vital arteries (*e.g.,* the aorta) are protected by immune privilege, malfunction of endogenous vasDCs may, indeed, be the pinnacle defect–initiating vasculitis.

Other innate immune cells captured in the vasculitic lesions, specifically macrophages, are particularly important as inflammatory effector cells.[18] Multiple functional domains of macrophages have pathogenic relevance. Intriguingly, the functional commitment of lesional macrophages is closely linked to their geographical location in the tissue site.[19] Intima-positioned macrophages produce inducible nitric oxide synthase, participating in regulating vascular tone. Adventitial macrophages secrete cytokines (IL-1β, IL-6, transforming growth factor-β [TGF-β]), conditioning the local inflammatory environment. Granuloma formation and giant cells dominate in the media and at the media–intima border. Functional profiling of media-residing macrophages has connected them to tissue damage; they produce metalloproteinases and reactive oxygen species (ROS), provide antioxidant regulation, and supply growth factors for myofibroblasts and newly formed microvessels. Essentially they drive the hyperplasia of the intimal layer that leads to luminal obstruction

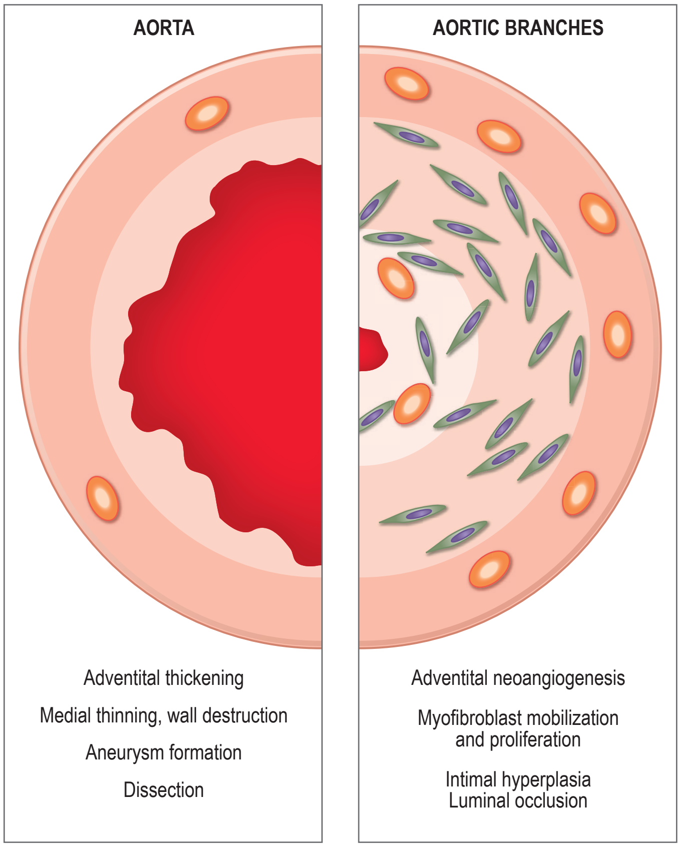


FIG 59.4 Patterns of Vessel Wall Damage in Giant-Cell Arteritis (GCA). Transmural granulomatous vasculitis results in two distinct damage patterns. In the aorta, the inflammatory infiltrates destroy the wall, inducing dissection and aneurysm formation. As a result of neoangiogenesis and fibrosis, the adventitia expands. In the aortic branches, transmural granulomatous infiltrates predominantly elicit a maladaptive healing response, characterized by the mobilization, migration, and proliferation of myofibroblasts forming hyperplastic and lumen-obstructive neointima. Formation of neotissue is associated with microvascular neoangiogenesis in the intima, media, and adventitia. New research indicates that intimal hyperplasia and neoangiogenesis are ultimately regulated through the intensity of the T-cell response in the vasculitic lesions.

and tissue ischemia. Whether macrophage effector functions are fundamentally different in aortitis, compared with branch vasculitis, requires further exploration. Differences of tissue damage patterns (wall destruction versus luminal occlusion) suggest at least substantial variances (Fig. 59.4).

ADAPTIVE IMMUNE SYSTEM DEFECTS

The hallmark lesions in the vessel wall are granulomatous infiltrates, in essence a mixture of highly activated macrophages, giant cells, and surrounding lymphocytes (see Fig. 59.3). The overwhelming majority of these lymphocytes are memory CD4 T cells. CD8 T cells are infrequent, and B cells are rare. Multinucleated giant cells are present in about 50% of patients, often localized close to the intima–media border adjacent to the lamina elastica interna. Fragmentation of that elastic membrane is a hallmark of GCA.

Analysis of clonal CD4 populations has yielded identical T-cell receptors (TCRs) in independent temporal artery biopsies from the same patient, strongly supportive for antigen driving T-cell behavior in the tissue microenvironment.[20] Wall-embedded vasDCs serve as antigen-presenting cells (APCs); the nature of the antigen remains unknown. Reports about infectious agents isolated from temporal arteries have notoriously not been confirmed in independent studies. The older adult host is expected to accumulate viral and bacterial microorganisms, and locally deposited pathogens may not indicate causal involvement but, rather, represent the vascular microbiota. Carefully designed experiments need to probe for a specific role of pathogens isolated from temporal artery tissue.

T helper 1 (Th1) cells are a key pathogenic element in GCA. Interferon-γ (IFN-γ), the Th1 marker cytokine, has multiple disease-relevant functions: activating endothelial cells and macrophages, regulating wall remodeling, inducing microvascular neoangiogenesis, and driving intimal-layer expansion.[14] IFN-γ production is predominantly produced in the adventitia, where T cells are instructed by vasDCs. Th1 cells are present in lesions from untreated patients and persist despite prolonged corticosteroid therapy,[21] suggesting independence of vasculitogenic Th1 cells from the steroid-sensitive acute-phase response and emphasizing their role as stabilizers of chronic-persistent lesions. During early vasculitis, Th1 cells are accompanied by Th17 cells. However, Th17 cells are highly sensitive to steroid therapy and chronic lesions are Th17 depleted.[21] Comparative analysis of sequential temporal artery biopsy tissues harvested from patients before and after treatment has demonstrated that adaptive immune responses evolve over the course of the disease and that different subpopulations of lesion-residing T cells are differentially responsive to immunosuppression.

Recent data have emphasized that the T-cell infiltrate in GCA-affected arteries is typically multifunctional and composed of a broad spectrum of functional lineages.[22] Besides Th1 and Th17 cells, Th9, and IL-21–producing T follicular helper (Tfh) cells are part of the infiltrate. Little is known about precise effector functions of these T cells. However, overall the T-cell population recruited to and retained in the vessel wall is functionally highly diverse, raising the possibility that a mixed set of antigens with vasculitogenic potential. Alternatively, the defect leading to granulomatous wall inflammation is the sequel of antigen-nonspecific defects.[22]

T cells have now been placed at the top of the artery's maladaptive response to injury, but precise effector pathways are not understood (see Fig. 59.4). Chronic aortitis leads to wall destruction and aneurysm formation. Dissection is increasingly recognized as a sequel of aortic wall inflammation. In rare cases, LVV results in the fatal complication of aortic rupture. In TA, direct cytotoxic function of CD8 T cells, natural killer (NK) cells, and γδ T cells has been implicated in local tissue injury.[23] Conversely, in the aortic branches, LVV typically causes luminal stenosis/occlusion as a result of formation of lumen-obstructive neotissue. Proliferating myofibroblasts lay down matrix and build hyperplastic intima. Newly formed microvessels appear in the adventitia and intima to supply oxygen and nutrients to the thickening wall. How T cells instruct vascular cells to this maladaptive wound healing response and whether heterogeneity of such vascular cell populations imposes the disease's tissue tropism are the subjects of ongoing research. So far, the few wall-infiltrating B cells have not been assigned to a specific pathogenic mechanism.

KEY CONCEPTS

Large Vessel Vasculitis

- Medium and large arteries in humans have multiple wall layers and a wall structure substantial enough to be the target of transmural inflammation.
- Vasculitis causes the rapid and concentric growth of hyperplastic intima, leading to luminal occlusion and ischemia of dependent tissues. Intramural inflammation of the aorta can result in wall damage followed by aneurysm formation and rupture.
- Because of the vital function and nonregenerative nature of large human arteries, the threshold for the induction and persistence of innate and adaptive immune responses in the wall structures of such arteries must be explicitly high.
- Inflammatory infiltrates typical for granulomatous vasculitis enter the vessel from the "back door," the adventitia, and not from the lumen.
- Besides their critical role in securing blood flow, medium and large blood vessels also possess immunoregulatory functions mediated by dendritic cells (DCs) indigenous to the vascular wall. DCs in each vascular territory express a distinctive pattern of Toll-like receptors (TLRs), giving each vessel its own immunological identity.

Defective T Regulatory Cells and Insufficient Immune Checkpoints in Giant Cell Arteritis

Recent data have brought to the forefront that antigen-nonspecific immunoregulatory pathways may have disease relevance in LVVs, specifically in enabling unopposed and lasting adaptive immune responses to induce and sustain vessel wall inflammation.

Like in many chronic inflammatory lesions, GCA's arterial wall infiltrates lack sufficient antiinflammatory T regulatory cells (Tregs). This may be a consequence of the proinflammatory milieu, but recent studies have identified a novel mechanism of Treg-dependent immunosuppression that is nonfunctional in GCA. In healthy individuals, secondary lymphoid tissue are occupied by a population of CD8 Tregs that effectively control clonal expansion of CD4 T cells, and thus the overall size of the CD4 T-cell compartment.[24] Such CD8 Tregs suppress activation and expansion of neighboring CD4 T cells by the directed transfer of nicotinamide adenine dinucleotide phosphate (NADPH) oxidase 2 (Nox2)–containing membrane particles. The population of Nox2$^+$ CD8$^+$ Tregs is distinctly low in patients with GCA and remains low with antiinflammatory therapy, suggesting a preexisting abnormality in the patients' immune system. Implicating Nox2 expressed on CD8 T cells in the threshold setting of CD4 T-cell immunity opens an entirely new perspective in vasculitis research and redirects the focus away from antigenic triggers that induce protracted macrophage activity.

A second immunoregulatory defect weakening proper control of CD4 T-cell immunity has been localized to the programmed death 1 (PD-1) immunoinhibitory checkpoint.[25] Inhibitory checkpoints are designed to protect tissue from excessive immune reactivity. PD-1 is expressed on activated T cells, the ligand PD-L1 is expressed on APCs. Triggering of PD-1 sends negative signals into T cells and stops their proliferation and polarization. DCs in patients with GCA, both lesional vasDCs and circulating DCs, are PD-L1^{low}. GCA T cells thus lack negative signaling and fail to undergo clonal contraction. Indeed, in inflamed temporal arteries PD-1$^+$ T cells are strongly enriched. In human artery–chimeric mice, blocking PD-1 with antibody treatment exacerbates vasculitis, and wall-infiltrating PD-1$^+$ T cells produce a spectrum of effector cytokines (IFN-γ, IL-17, IL-9, IL-21, etc.). Most importantly, PD-1–PD-L1 interactions in the artery have direct impact on the density of newly formed microvessels and the thickness of the hyperplastic intimal layer.[25] Hyperactive checkpoints are now targeted therapeutically to unleash T-cell immunity in patients with cancer. A role for PD-1–derived negative signaling in vasculitis predicts that such patients are at risk to develop drug-induced vasculitis.

CLINICAL FEATURES IN GIANT CELL ARTERITIS

CLINICAL PEARLS

Clinical and Epidemiological Clues in Giant-Cell Arteritis

- Patient older than 50 years of age
- Female
- Northern European heritage
- Laboratory findings of a highly activated acute-phase response
- Insidious onset of nonspecific symptoms (weight loss, night sweats, malaise, fever)
- Ischemia of ocular structures, cranial muscles, scalp, or upper extremities

Clinical manifestations of GCA reflect the combination of a systemic inflammatory syndrome with vascular insufficiency (Table 59.1).[2,26] Increased sensitivity of vascular imaging methods and longer survival of affected individuals have made it clear that in most patients medium and large arteries are involved. In some patients, the disease preferentially targets more peripheral branches of the aorta (*e.g.*, cranial arteries, such as the temporal artery). In others, the aorta and its proximal branches (*e.g.*, subclavian, axillary arteries) are involved to a major extent. In a subpopulation of patients, the clinical consequences of arterial inflammation are minimal, and these consequences come to clinical attention because of failure to thrive.

In cranial GCA, symptoms result from vascular stenosis of the neck and head arteries, most prominently in the branches of the external carotid artery. Arteritis of the scalp arteries leads to the typical presentations of headaches and scalp tenderness. Patients report difficulties with wearing glasses or combing their hair. The headaches are often intense and unresponsive to standard analgesics. Headaches are a nonspecific clinical symptom, yet in an older individual with other findings of an inflammatory syndrome, physicians need to rule out GCA. Insufficient blood flow to the masseter muscles and the tongue causes jaw or tongue claudication, elicited by prolonged chewing and talking. Although this type of claudication is present in <30% of patients, it is clinically helpful because it rarely occurs outside of GCA. Similarly, painful dysphagia can be a useful clinical clue.

The orbits and the optic nerve are strictly dependent on blood supply from the external carotid system, particularly the ophthalmic artery. GCA in branches of the ophthalmic artery, specifically the posterior ciliary arteries, leads to anterior ischemic optic neuropathy, presenting as sudden and painless vision loss. Typically, patients lose vision in the early-morning hours or wake up blind. Involvement of one eye may be followed by visual loss in the partner eye if the disease is not diagnosed and treated promptly. Besides anterior optic neuropathy, GCA can cause a number of ischemic complications in the orbits and along the visual axis, which may present as diplopia or partial vision loss.[18] If recognized and treated immediately, vasculitis-associated sight

TABLE 59.1 **Clinical Features of Giant-Cell Arteritis (GCA), Polymyalgia Rheumatica (PMR), and Takayasu Arteritis (TA)**

Organ System	Clinical Features	FREQUENCIES		
		GCA	PMR	TA
Vascular	Headaches	+++		*
	Limb claudication	+		***
	Scalp tenderness	**		
	Jaw claudication	**		
	Absent or asymmetrical pulses	*		***
	Asymmetrical blood pressure readings	*		***
	Bruit	*		***
	Tongue claudication	*		
	Tissue gangrene	*		
	Abdominal angina			*
	Cough (dry, nonproductive)	*		*
Constitutional	Malaise	**	**	***
	Failure to thrive	*	**	*
	Weight loss	**	**	**
	Fever	*	*	*
Central nervous	Ocular symptoms	**		*
	Stroke/transient ischemic attack	*		*
Peripheral nervous	Peripheral neuropathy	*		
Cardiac	Aortic dilatation and regurgitation	*		*
	Myocardial infarction	*		*
	Congestive heart failure			*
Musculoskeletal	Proximal stiffness/muscle pain	**	***	
	Synovitis of peripheral joints		*	
Others	Intense acute-phase response	***	***	***
	Normochromic or hypochromic anemia	**	*	**

*** = high frequency (>70%); ** = moderate frequency (20–70%); * = low frequency (<20%).

loss is preventable, which indicates that GCA should be considered an ophthalmological emergency.

Chronic nonproductive cough can be related to arteritis in bronchial artery branches. If the vertebral and basilar arteries develop vasculitic stenosis, ischemia of the central nervous system manifests with transient ischemic attacks or frank stroke.

In patients with large-vessel GCA, cranial symptoms may be minimal, and temporal artery biopsy can be negative.[27] Instead, vascular insufficiency is focused on the upper-extremity vessels and the aorta. In rare cases, lower extremities are affected. Typically, patients have asymmetrical blood pressure readings or experience total loss of upper-extremity blood pressure and pulse. The underlying lesions are occlusions in the distal subclavian arteries, often extending into the axillary sections (Fig. 59.5). Patients with subclavian GCA are on average about 10 years younger at disease onset compared with those with dominant cranial manifestations. Diagnosis of large-vessel GCA can be delayed as symptoms are nonspecific and the systemic inflammatory component is less pronounced. Ischemic pain in the hands when using the arms is often combined with coolness and bluish discoloration. Gangrene of the fingertips is rare. Disability can be significant, as patients have difficulties with activities of daily living. With stenotic lesions of the subclavian arteries, blood pressure readings can be unreliable or not possible, requiring alternative strategies for blood pressure monitoring. Although carotid involvement is considered infrequent, it can be challenging to distinguish atherosclerotic disease and vasculitic disease. Patients with carotid GCA are at high risk for cerebral ischemic events. Aortic involvement preferentially targets the thoracic aorta and infrequently the abdominal aorta (Fig. 59.6). Dilation of the aortic root can lead to aortic insufficiency. Aortic aneurysms are often clinically silent. The diagnosis may first be made from tissue obtained surgically during aortic aneurysm repair. In extreme cases, the aortic wall ruptures.

The response pattern of arteries to inflammation may not include intimal hyperplasia, thus eluding luminal compromise and vascular failure. In such patients, the systemic inflammatory component dominates the clinical presentation. Fever, fatigue, malaise, weight loss, and depression are often intense enough to prompt a workup for a malignancy. GCA needs to be a differential diagnosis in all cases of fever of unknown origin, particularly in older individuals. Patients with cranial GCA have abnormally thick and tender temporal arteries exhibiting nodularity and loss of pulses, whereas clinical findings in nonstenosing GCA can be unremarkable. Temporal artery biopsy should be pursued even if clinical examination does not suggest the diagnosis.

CLINICAL FEATURES IN POLYMYALGIA RHEUMATICA

PMR is diagnosed in patients presenting with pronounced stiffness and pain in the shoulder and pelvic girdle muscles (see Table 59.1).[2] Laboratory testing reveals a systemic inflammatory syndrome; arterial biopsy is negative for arteritis. It is estimated that about 10% of patients with PMR without any signs of vascular inflammation will eventually develop full-blown vasculitis. Notably, PMR often occurs in patients with GCA and is present in about 40% of patients with GCA at disease onset. Tapering of immunosuppressive therapy in GCA is frequently associated with new or remittent PMR symptoms.[26] Complaints are focused on muscle pain and stiffness, classically affecting the neck, shoulders, and pelvic girdle. The muscles of the torso may be involved. Peripheral areas of the arms and legs are spared. Muscle pain is most intense in the early morning and improves during the day. Inability to get out of bed, stand up from a chair, or get off the toilet seat should alert the physician to consider PMR. Some patients with PMR have synovitis or bursitis in their shoulder and hip joints,[2,28] making it difficult to distinguish these patients from those with seronegative polyarthritis. No diagnostic procedure that allows for the diagnosis of PMR is available; the syndrome remains an exclusion diagnosis in cases of myalgia combined with laboratory signs of systemic inflammation. On clinical examination, passive motion of shoulder and hip joints is seen to be maintained, but active motion is restricted because of pain. Muscle strength is often normal. Careful evaluation of the temporal arteries is warranted to avoid missing fully developed GCA.

CLINICAL FEATURES IN TAKAYASU ARTERITIS

The clinical manifestations of TA are diverse and depend on the affected vascular territory (see Table 59.1; Table 59.2).[29] Initial

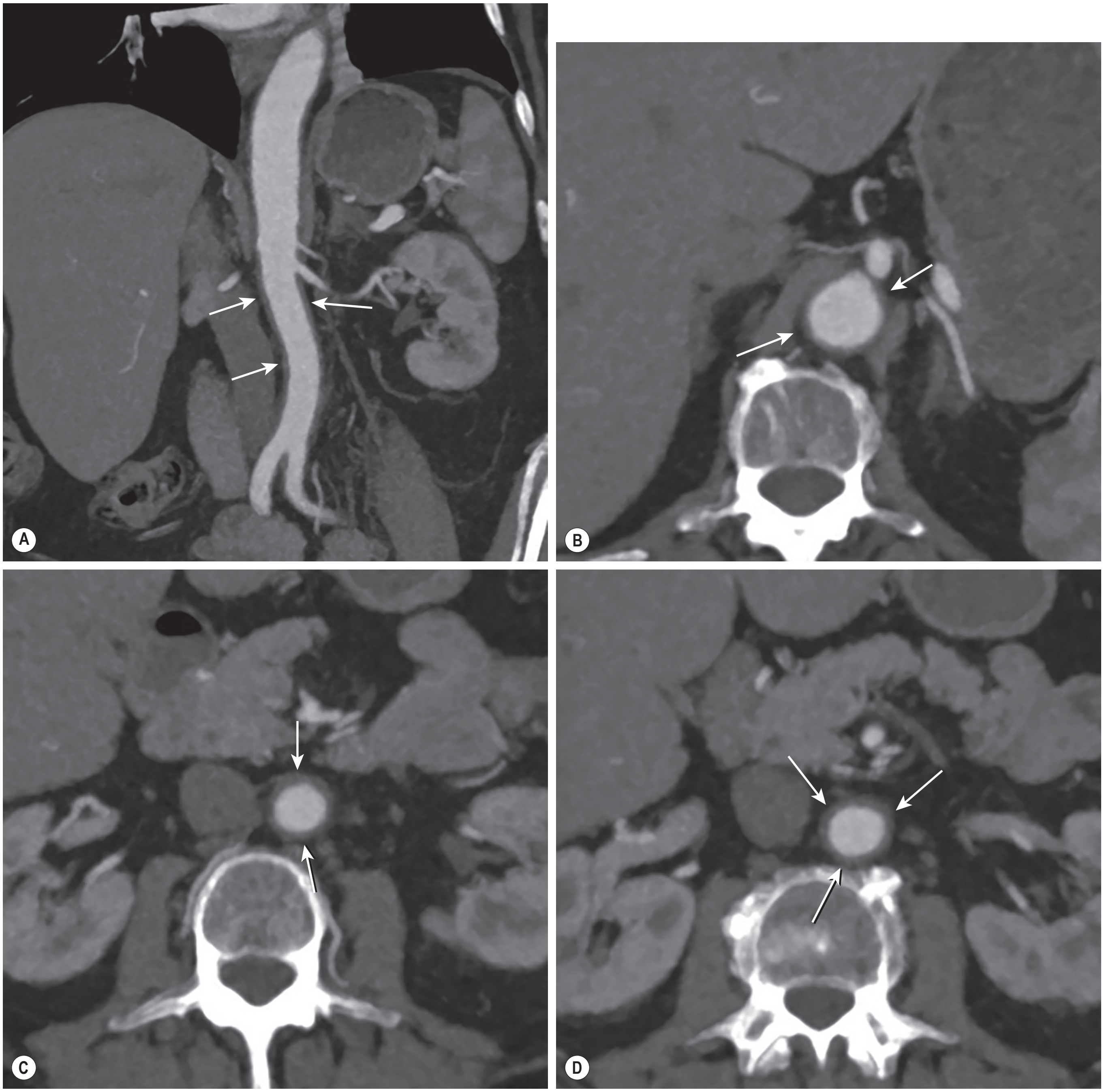


FIG 59.5 **Diagnostic Imaging in Giant-Cell Arteritis (GCA): Computed Tomographic Angiography (CTA).** Contrast-enhanced CTA in a 74-year-old female with a temporal artery biopsy positive for GCA. CTA imaging shows diffuse, circumferential mural thickening along the entire descending aorta and the abdominal aorta (A). Arrows mark the thickened aortic wall. Axial images (B, C, and D) reveal the circumferential distribution *(arrows)* of the wall thickening from the distal arch to the intrarenal portion of the aorta. Thickness measurements can be used to monitor disease burden over time. The aortic diameter is within normal limits, indicating that the aortitis has not yet resulted in aneurysm formation. [Images were generated by Dr. D. Fleischmann, Department of Radiology, Stanford University.]

symptoms are usually nonspecific and include fever, cough, malaise, weight loss, night sweats, myalgias, and arthralgias. Signs of vascular deficiency develop later in the disease course and generally are ischemic in nature. Geographical variations in disease pattern have been reported, likely reflecting the interplay between host risk genes and dysfunctional immunity. In North American, Japanese, and Korean patients, the aortic arch and its primary cervical and upper extremity branches are preferentially targeted, giving rise to aortic insufficiency, cerebral ischemia, face and neck pain, ocular ischemia, and the typical presentation

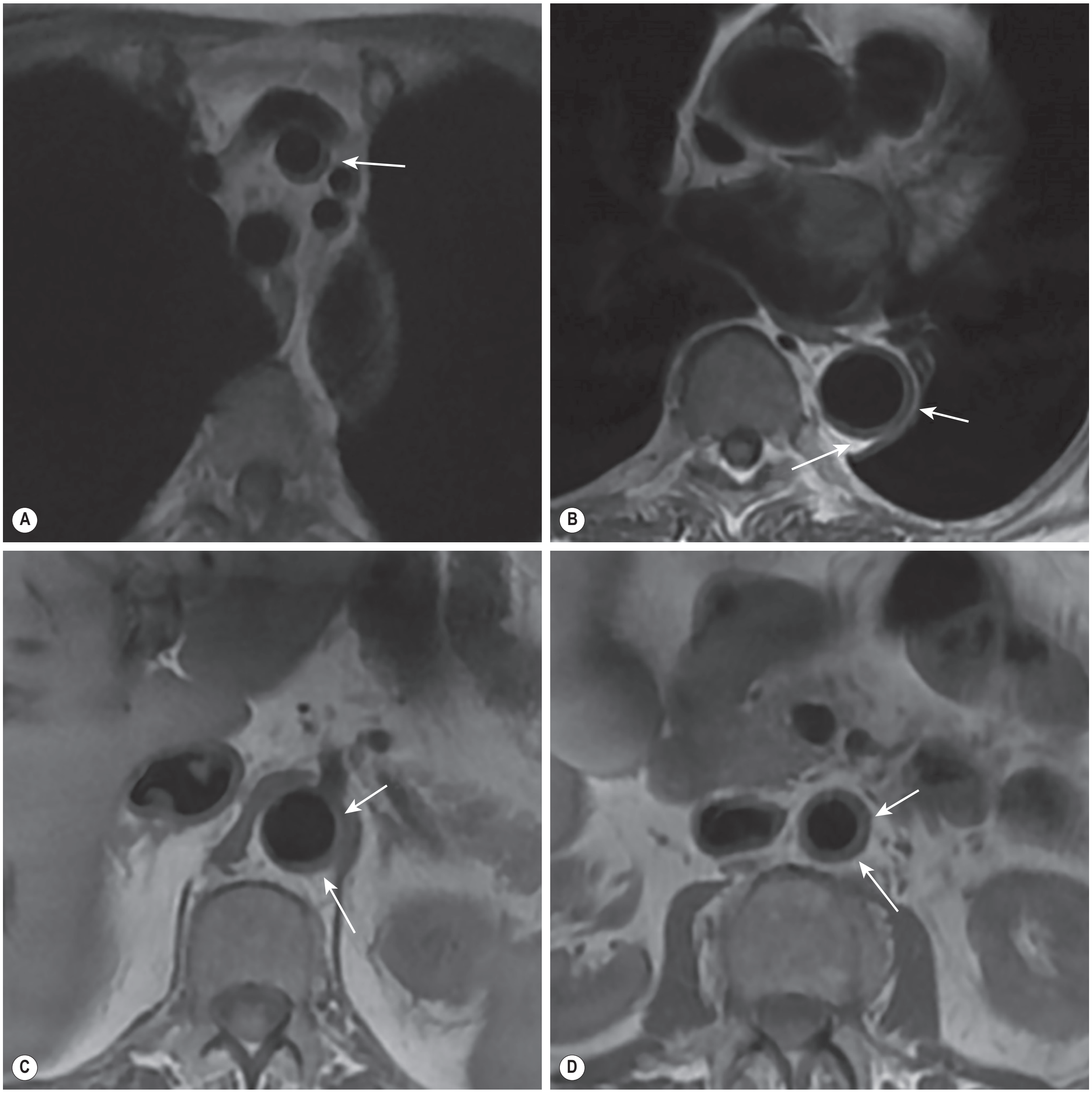


FIG 59.6 **Diagnostic Imaging in Giant-Cell Arteritis (GCA): Magnetic Resonance Angiography (MRA).** Contrast-enhanced MRA of the chest and abdomen in the patient whose computed tomographic angiography (CTA) images are presented in Fig. 59.5. Double-inversion recovery magnetic resonance imaging (MRI) scans in the axial plane demonstrate diffuse, contiguous mural thickening of the great vessels (A, brachiocephalic trunk) and the aorta (B–D). Arrows are placed to mark the circumferential distribution of the mural thickening and compare luminal diameter and wall thickness at different levels of the vessels. [Images courtesy of Dr. D. Fleischmann, Department of Radiology, Stanford University.]

of "pulseless disease" (Fig. 59.7). In patients in India, the abdominal aorta and renal arteries are more commonly affected, causing renovascular hypertension and the long-term risk of cardiac failure (Fig. 59.8).[30]

Nonspecific complaints of headaches, syncope, and face and neck pain are often misinterpreted as stress-related problems, particularly in young women. Consequently, the diagnosis can be missed for months. Only a few patients come to clinical attention because of catastrophic neurological symptoms related to brain ischemia. Helpful clues are differences in blood pressure, loss of pulses, and vascular bruits heard on clinical examination. Retinal neoangiogenesis, induced by hypoperfusion of the eye,

TABLE 59.2 Takayasu Arteritis: Relationship Between Clinical Symptoms and Affected Vascular Territories

Vascular Bed Involvement	Approximate Frequency (%)	Predominant Clinical Symptoms
Subclavian	90	Arm claudication, pulselessness
Common carotid	60	Visual defects, stroke, transient ischemic attack, syncope
Abdominal aorta	45	Claudication, hypertension, abdominal angina
Renal	35	Hypertension
Aortic arch/root	35	Aortic insufficiency, congestive heart failure
Vertebral	35	Dizziness, visual impairment
Celiac axis	20	Abdominal angina
Superior mesenteric	20	Abdominal angina
Iliac	20	Claudication
Pulmonary	10	Dyspnea, chest pain
Coronary	10	Myocardial infarction, angina

is now relatively rare, but fleeting visual abnormalities may indicate transient ischemic attacks. Signs of aortic insufficiency are unlikely to be encountered in early disease, but continuous monitoring for aortic dilation is an essential part of follow-up care. Coronary artery stenosis in a young patient must prompt the physician to rule out TA. In a subset of patients, the origins of mesenteric arteries are involved by stenosing vasculitis. Clinical consequences include weight loss, nausea, vomiting, diarrhea, and abdominal claudication, typically elicited by the increased intestinal blood demand following a meal.

Renal artery stenosis may be clinically silent and is often noticed in routine screening. Correct measurement of blood pressure can represent a pressing clinical problem if the upper-extremity arteries are affected. Involvement of the infrarenal aorta can lead to lower-extremity claudication. Musculoskeletal examinations are usually unrevealing, although joint and muscle pains are common complaints.

DIAGNOSIS

Classification criteria have been developed for GCA and TA to differentiate patients with LVV from those with other vasculitic entities (Tables 59.3 through 59.5).[31-33] Age at disease onset and the pattern of arteritis are clearly important for establishing the diagnosis and distinguishing between these two related vasculopathies. Diagnostic criteria for PMR remain a challenge (see Table 59.3) as they rely on nonspecific symptoms, such as muscle pain and stiffness and elevated erythrocyte sedimentation rate (ESR), all of which can occur in many other diseases.[2] No specific laboratory test is currently available to diagnose PMR. Therapeutic responsiveness of patients with PMR to low-dose corticosteroids is clinically helpful but stresses the need for objective diagnostic criteria.

Laboratory Tests

In all three conditions—GCA, PMR, and TA—the laboratory findings indicate an intense acute-phase response in a vast majority of patients.[2,26,34] Generally this is captured by measuring ESR or CRP. It is important to note, however, that a subset of patients with GCA has normal ESR readings, even before initiation of immunosuppressive therapy. A normal ESR or CRP reading is not sufficient to exclude the diagnosis, and further diagnostic workup is required if clinical presentation is suspicious for vasculitis. Other acute-phase proteins, such as fibrinogen and SAA, have been reported to be elevated as well. IL-6 is a potent inducer of acute-phase proteins in the liver and has been found to be a sensitive marker of continuous systemic inflammation.[10,11]. Other laboratory abnormalities, such as elevation of alkaline phosphatase, thrombocytosis, and anemia, are in line with a robust acute-phase response.

Autoantibody measurements are not helpful beyond excluding differential diagnoses, such as rheumatoid arthritis (RA), systemic lupus erythematosus (SLE), or antineutrophil cytoplasmic antibody (ANCA)–related vasculitides. No disease-specific autoantibodies for GCA and TA have been discovered, emphasizing that B-cell immunity may not contribute to vascular inflammation.

Tissue Biopsy

In patients with TA, tissue biopsies are rarely available unless the patient had to undergo vascular reconstructive surgery. In most patients, the diagnosis is made on the basis of imaging procedures revealing luminal and wall abnormalities in affected blood vessels.

In contrast, arterial biopsy remains a critical diagnostic approach in managing patients with GCA. Temporal arteries are easily accessible, and a segment of these arteries can be removed in an outpatient setting. Recommendations include harvesting 2–3 cm of the temporal artery, starting at the most symptomatic side. Frozen tissue sections can lead to a quick diagnosis of granulomatous vasculitis. Whether the second side should be biopsied during the same surgical procedure remains a matter of debate. In cohorts that included several hundred patients, vasculitis was detected in 2–3% of tissue samples from the second side if the first side was negative. If the clinical suspicion is strong, biopsy confirmation can be sought from a second-side biopsy immediately after the first biopsy or after careful monitoring of the patient for several weeks. Negative findings on temporal artery biopsy do not exclude the diagnosis of GCA. In a retrospective cohort study, about half of all patients with subclavian GCA had no evidence of vasculitis in the temporal arteries, emphasizing that the disease may display clear preference for certain vascular territories.[35] There has been a recent trend toward considering negative biopsies as "false negative" and classifying patients with a negative biopsy as having GCA. This may lead to unnecessary immunosuppressive therapy, and the patient may not thus obtain a proper diagnosis. Findings that prompt biopsies (headaches, elevated acute-phase reactants) are notoriously nonspecific, and both physicians and patients are anxious to avoid treatable blindness, biasing toward overtreatment. The temporal artery biopsy remains a powerful diagnostic tool and the major tool that allows unequivocal classification of the disease process. A technically proper temporal artery biopsy will detect vasculitis in the vast majority of patients.

Corticosteroid therapy does not eradicate pathological findings of vascular wall infiltrates, and biopsy can still be valuable in making the diagnosis in patients on steroids.[21] Nevertheless it is possible that treatment with steroids leads to a false-negative biopsy result in some patients.

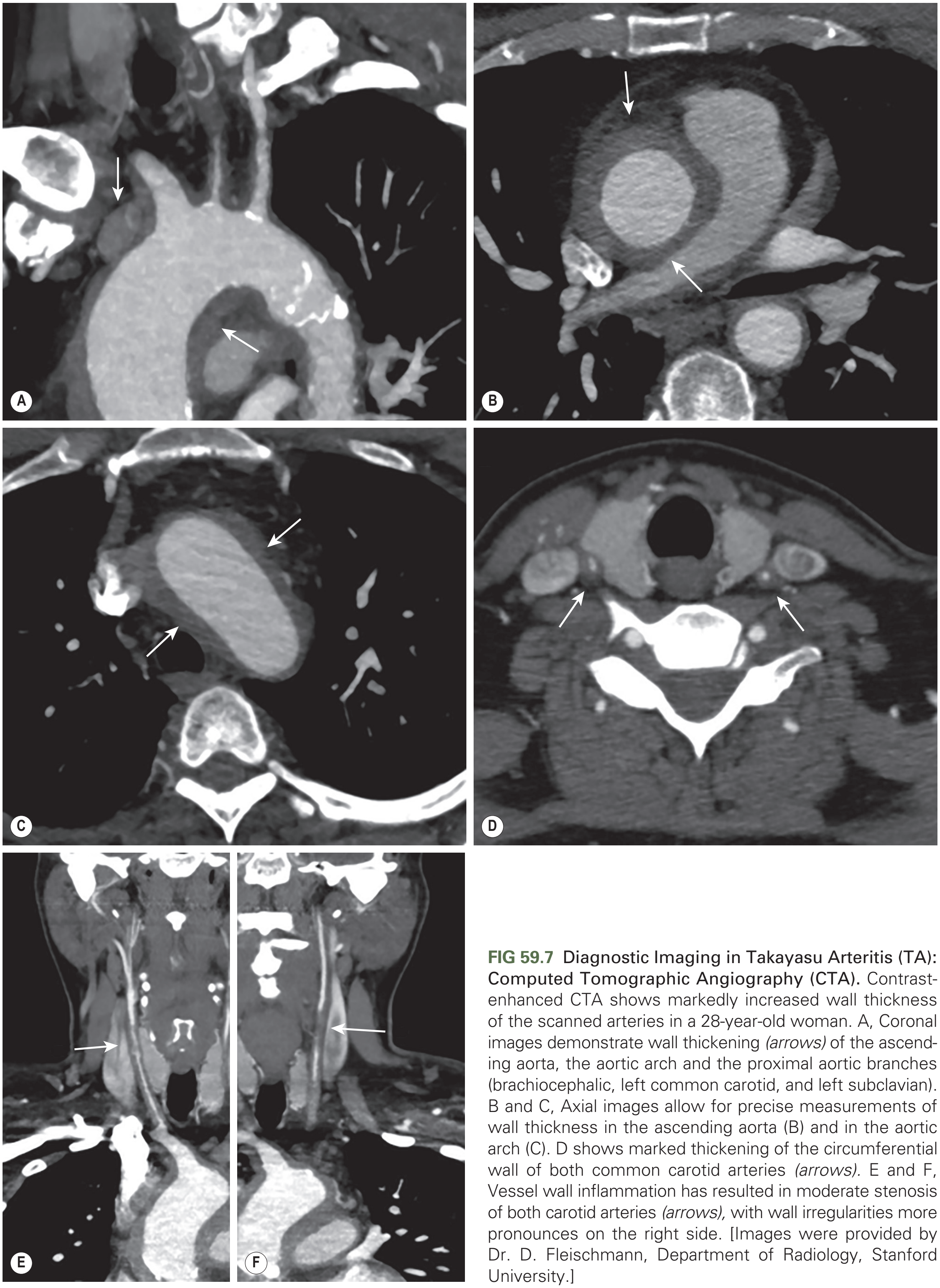


FIG 59.7 **Diagnostic Imaging in Takayasu Arteritis (TA): Computed Tomographic Angiography (CTA).** Contrast-enhanced CTA shows markedly increased wall thickness of the scanned arteries in a 28-year-old woman. A, Coronal images demonstrate wall thickening *(arrows)* of the ascending aorta, the aortic arch and the proximal aortic branches (brachiocephalic, left common carotid, and left subclavian). B and C, Axial images allow for precise measurements of wall thickness in the ascending aorta (B) and in the aortic arch (C). D shows marked thickening of the circumferential wall of both common carotid arteries *(arrows).* E and F, Vessel wall inflammation has resulted in moderate stenosis of both carotid arteries *(arrows),* with wall irregularities more pronounces on the right side. [Images were provided by Dr. D. Fleischmann, Department of Radiology, Stanford University.]

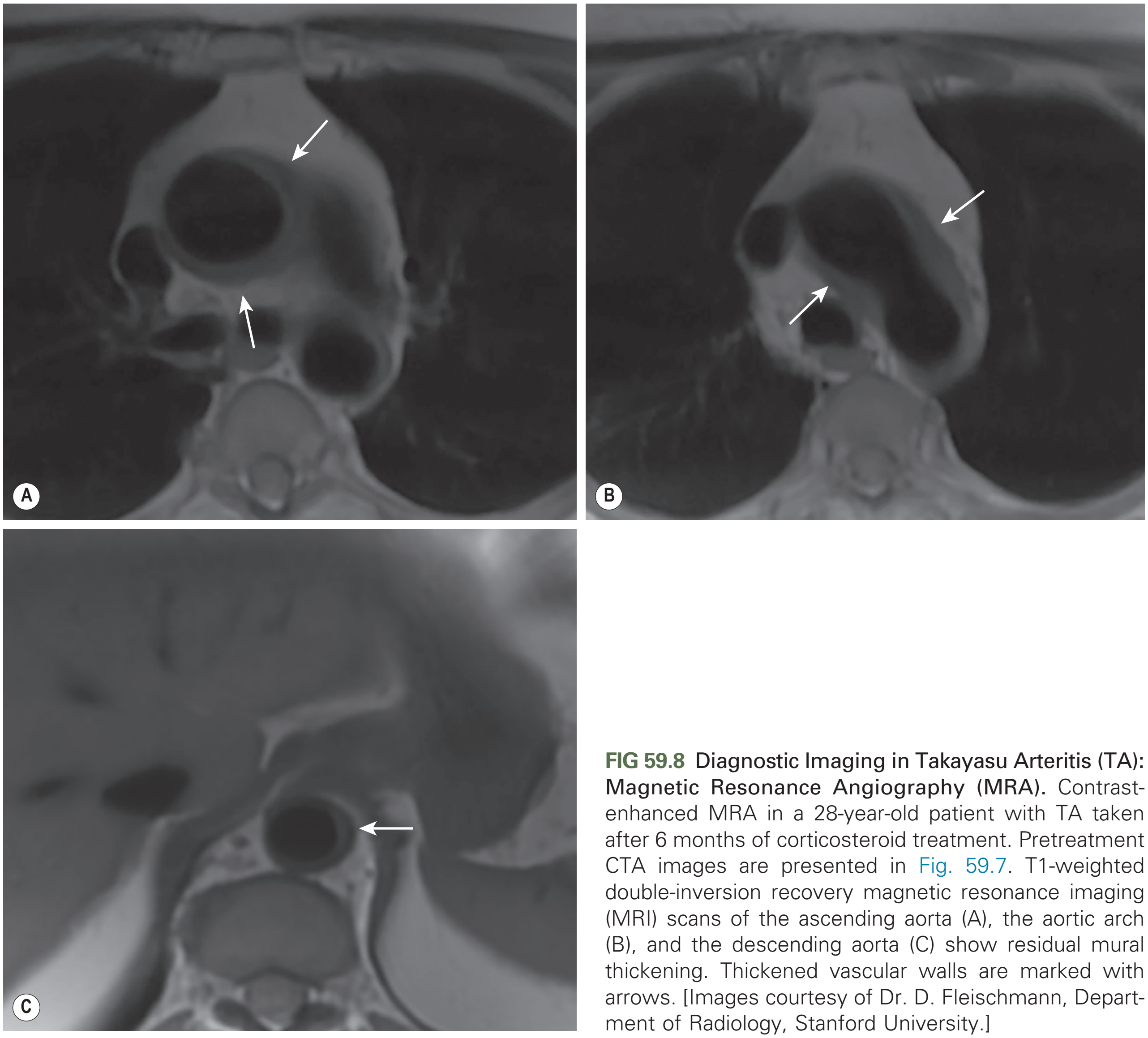


FIG 59.8 Diagnostic Imaging in Takayasu Arteritis (TA): Magnetic Resonance Angiography (MRA). Contrast-enhanced MRA in a 28-year-old patient with TA taken after 6 months of corticosteroid treatment. Pretreatment CTA images are presented in Fig. 59.7. T1-weighted double-inversion recovery magnetic resonance imaging (MRI) scans of the ascending aorta (A), the aortic arch (B), and the descending aorta (C) show residual mural thickening. Thickened vascular walls are marked with arrows. [Images courtesy of Dr. D. Fleischmann, Department of Radiology, Stanford University.]

TABLE 59.3 American College of Rheumatology 1990 Classification Criteria for Giant-Cell Arteritis[a] and Polymyalgia Rheumatica

Age at disease onset ≥50 years
New-onset or new type of headache
Temporal artery tenderness or decreased artery pulse
Elevated erythrocyte sedimentation rate (≥50 mm/hr)
Histological incidence of arteritis (characterized by a predominance of mononuclear cell infiltrates or a granulomatous process with multinucleated giant cells)

[a] A patient is classified as having giant-cell arteritis if at least three of the five criteria are present.

Reprinted from Hunder GG, Bloch DA, Michel BA, et al. The American College of Rheumatology 1990 criteria for the classification of giant cell arteritis. Arthritis Rheum 1990; 33: 1122–1128, with permission of Wiley-Liss, Inc., a subsidiary of John Wiley & Sons, Inc. ©1990.

TABLE 59.4 Provisional Classification Criteria for Polymyalgia Rheumatica

Age ≥50 years	required
Bilateral shoulder ache	required
Abnormal C-reactive protein (CRP) and/or erythrocyte sedimentation rate (ESR)	required
Morning stiffness >45 minutes	2
Hip pain or limited range of motion	1
Absence of rheumatoid factor (RF) or anti-citrullinated protein antibody (ACPA)	2
Absence of other joint involvement	1
A score of ≥4 is categorized as polymyalgia rheumatica.	

Reprinted from Dasgupta B, et al. Provisional classification criteria for polymyalgia rheumatica: A European League Against Rheumatism/American College of Rheumatology collaborative initiative. Arthritis Rheum 2012; 64: 943–954.

Histomorphological reports describe mononuclear cell infiltrates penetrating through all layers of the vessel wall (Fig. 59.9).[1] Recent discussions have focused on the diagnostic relevance of isolated inflammatory cell clusters in the adventitia or perivascular lymphocytes limited to small blood vessels.[36] These findings may not be sufficient to indicate arteritis. Multinucleated giant cells may or may not be found. They tend to lie along the internal elastic lamina, at the junction between the media and the intima. Media destruction is not unusual, but findings of fibrinoid necrosis should prompt a search for a different vasculitic entity. The vessel lumen is more or less compromised by hyperplastic intima formed from proliferating fibroblasts, smooth-muscle cells, and deposition of acid mucopolysaccharides.

The histology of TA is similar to that of GCA, making it difficult to dissect both syndromes in tissue samples derived from the aorta or its primary branches. Lymphocytes and plasma cells accumulate around vasa vasorum and form transmural infiltrates. Marked wall thickening with inflammatory tissue extending into perivascular structures is typical for TA (Fig. 59.10). Destruction of elastic membranes is often extensive and combined with patchy areas of media necrosis. Weakening of the vessel wall can lead to aneurysm formation. Inflammatory lesions may be arranged in a "skipped" pattern, with normal vessel wall segments alternating with stretches of intense destructive inflammation.

TABLE 59.5 American College of Rheumatology 1990 Criteria[a] for the Classification of Takayasu Arteritis

- Disease onset at ≤40 years
- Claudication of an extremity
- Decreased brachial artery pulse
- >10 mm Hg difference in systolic blood pressure between arms
- Bruit over the subclavian arteries or the aorta
- Arteriographic evidence of narrowing or occlusion of the entire aorta, its primary branches, or large arteries in the proximal upper or lower extremities

[a]For purposes of classification, a patient is classified as having Takayasu arteritis if more than three of the six criteria are fulfilled.

Reprinted from Arend WP, Michel BA, Bloch DA, et al. The American College of Rheumatology 1990 criteria for the classification of Takayasu arteritis. Arthritis Rheum 1990; 33: 1129–1134, with permission of Wiley-Liss, Inc., a subsidiary of John Wiley & Sons, Inc. ©1990.

Physicians may encounter morphological findings of granulomatous aortitis in patients undergoing aortic aneurysm repair without any prior diagnosis of vasculitis. Detailed workup of these patients is necessary to identify those with undiagnosed PMR, GCA, or TA. Rare causes of aortitis, including inflammatory bowel disease, sarcoidosis, syphilis, relapsing polychondritis, and connective tissue disease, should be ruled out. Isolated granulomatous aortitis is diagnosed as idiopathic aortitis. The pathogenesis and prognosis of this condition are essentially unknown.

Diagnostic Imaging

Modern imaging modalities have fundamentally changed the diagnostic approach to LVV.[26] Indeed, diagnosing TA mostly depends on identifying vascular lesions in typical distribution by imaging.[37]

Conventional angiography still has its place in preoperative planning and can be combined with intravascular interventions. It provides ideal visualization of the vascular lumen not only for large but also for medium-sized arteries, such as the axillary and brachial arteries (see Fig. 59.5). Ultrasound (US)–based methods are extremely useful for screening carotid arteries, but they have also emerged as the method of choice for initial assessment of the distal subclavian arteries, vertebral arteries, renal arteries, and femoral arteries. US examination is also the optimal method for long-term monitoring of vessel bypasses in patients who have undergone revascularization surgical procedures. Magnetic resonance imaging (MRI), magnetic resonance angiography (MRA),

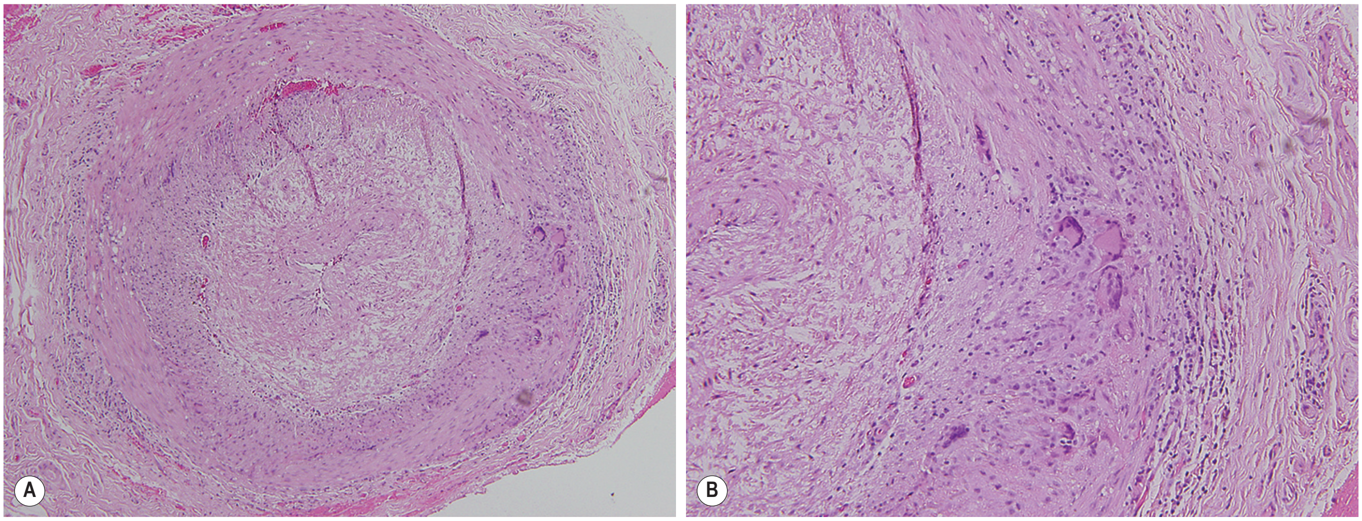


FIG 59.9 Histomorphology of Giant-Cell Arteritis (GCA). (A) Temporal artery cross-section with mononuclear infiltrates throughout all wall layers. The adventitia is infiltrated by round cells with cuffing of vasa vasorum by lymphocytes. The vessel lumen is occluded by intimal hyperplasia. (B) Higher magnification showing intense granulomatous inflammation with multinucleated giant cells in the proximal media and at the media–intima junction.

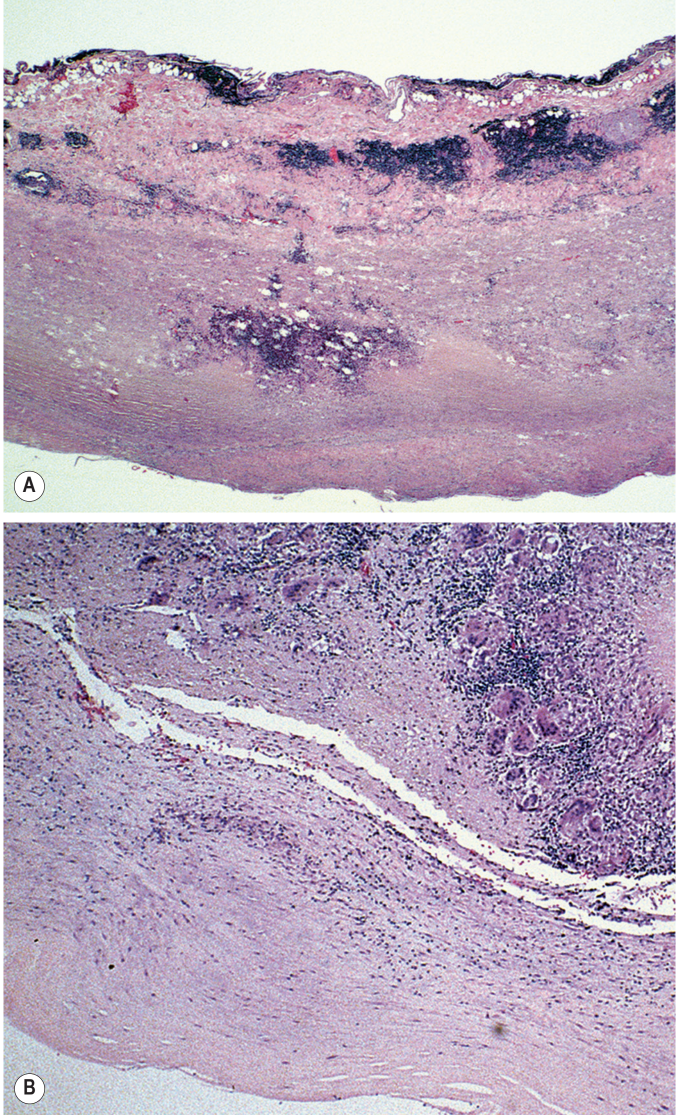


FIG 59.10 **Histopathology of Takayasu Arteritis (TA).** (A) Full-thickness section of the aortic wall shows dense mononuclear infiltrates in the adventitia and media. The intima is thickened and wavy; hematoxylin and eosin (H&E). (B) Florid granulomatous inflammation along the media–intima junction with numerous giant cells; H&E.

and computed tomography (CT) are now widely used for evaluating the vascular tree. These methods provide excellent information on abnormalities of the vascular lumen and wall, with increasing resolution, now capturing abnormalities in the more peripheral arterial branches. CT imaging is fast, well tolerated by patients with claustrophobia, and allows excellent assessment of the aorta and its wall (see Fig. 59.7). However, it has the disadvantages of contrast loading and radiation exposure. With its inherent multiplanar imaging capabilities, magnetic resonance is used to examine neck vessels, the aorta, and its primary branches (see Figs. 59.6 and 59.8). Great hope was placed on its potential to measure wall edema and intramural vascularity, which would make MR useful for estimating disease burden and responses to therapy. However, a carefully conducted study comparing imaging results with laboratory parameters of inflammation to results from surgically harvested vessel biopsy specimens has been disappointing, cautioning one that edema-weighted MR should not be used as a sole means of measuring disease activity and therapeutic responsiveness.[38] Both, CT angiography and MRA are now used routinely to monitor progression or regression of vascular involvement and have an important place in managing the chronic phase of GCA and TA.

THERAPEUTIC MANAGEMENT

THERAPEUTIC PRINCIPLES

Treatment of Large Vessel Vasculitides

- To prevent vision loss, patients with giant-cell arteritis (GCA) require immediate treatment. Similarly, with the threat of catastrophic cerebral ischemia in Takayasu arteritis (TA), prompt initiation of therapy is imperative.
- Corticosteroids are the immunosuppressive drug of choice for large-vessel vasculitides (LVVs). Often, the drugs must be given over a period of several years but may be clinically effective at very low doses.
- Clinical trials have failed to show convincing steroid-sparing effects for either methotrexate or tumor necrosis factor (TNF)-α blockade in GCA.
- A phase 3 clinical trial has demonstrated steroid-sparing effect of treatment with tocilizumab.
- Molecular studies of the inflammatory infiltrate in GCA have shown that early and untreated disease is characterized by two functional T-cell lineages; T helper (Th)1 and Th17 cells. Th17 cells are rapidly responsive to corticosteroids, whereas Th1 cells persist and promote chronic, smoldering vasculitis.
- It is not known whether the smoldering activity persisting beyond the acute phase of the disease requires immunosuppressive therapy or whether the benefits of chronic immunosuppressive therapy outweigh the potential risks.
- Clinical experience (not evidence-based therapeutic trials) suggests that a combination of methotrexate, mycophenolate mofetil, or TNF-α–blocking agents with corticosteroids may be beneficial in controlling disease in some patients with TA.
- Close monitoring for diabetes, hypertension, and hyperlipidemia combined with bone-saving therapy should be part of the treatment regime in patients with LVVs on long-term corticosteroids.

With increasing knowledge of the disease process and refinement of diagnosis and long-term treatment, the prognosis for patients with LVVs has significantly improved. Life expectancy of patients with GCA is preserved. Follow-up studies of Japanese patients with TA have suggested good control of disease activity in about 75% of patients, with only 25% experiencing serious complications and cardiac manifestations dominating long-term outcome. Whether vasculitis predisposes patients to accelerated atherosclerotic disease, given the combination of chronic inflammation and injury to vessel wall structures, is still being discussed. It is not known whether progression of atherosclerosis and its complications require a different management approach or whether standard vasoprotective measures (treating hypertension and hyperlipidemia, smoking cessation, etc.) are sufficient.[6]

Pathogenic studies have pointed out that the traditional view of GCA as a self-limiting disease is incorrect.[11] To the contrary, granulomatous vasculitis has shown surprising resistance to immunosuppression, with vessel wall infiltrates persisting for >12 months in almost 50% of patients despite appropriate immunosuppressive therapy. Based on examination of serial temporal artery biopsy specimens from patients before and after treatment, it is now clear that arteritis persists, albeit sustained

ON THE HORIZON

- Medium and large arteries in humans sense danger signals through wall-embedded cells; changing the understanding of how the immune system interacts with the vascular system.
- The multilineage nature of vasculitic T cells, which display differential therapeutic responsiveness, almost certainly will require more complex therapies, adapted to disease stage and immune status of the host.
- The immune system changes profoundly with age. Age-appropriate management of each patient and avoidance of overtreatment of older adults are important.
- Current therapies in large-vessel vasculitis induce partial remission. Appropriately designed studies are required to explore whether partial remission is sufficient and whether the risk/benefit ratio is maintained if complete remission is attempted.

by an immune network distinct from that in untreated patients.[21] It is currently not known whether this persistent smoldering process needs to be treated and what the risk/benefit ratio is for the older patient population affected by GCA. Unchanged life expectancy in GCA suggests adequacy of current management. Whether intensification of immunosuppressive therapy or chronic maintenance therapy can prevent long-term complications, such as aortic aneurysm/dissection from GCA aortitis, and improve the overall prognosis is unknown. The ultimate decision depends on the cost/benefit analysis comparing the risk from smoldering disease with the risks imposed by long-standing immunosuppression. In that context, it is important to remember the profound impact of the immune aging process, which leaves the patient with an impaired immune system and amplifies the risk of immunosuppression.[39]

Induction Therapy

In newly diagnosed patients with GCA, TA, and PMR, the immunosuppressants of choice are corticosteroids. Patients with GCA are started on a daily prednisone dose of 40–60 mg (about 1 mg/kg body weight). In patients with PMR, a daily dose of 20 mg prednisone is sufficient in almost all patients. The response is usually dramatic, with improvements within 24–48 hours. The promptness of clinical improvement is so exceptional that it has been suggested as a diagnostic criterion for PMR. However, the promptness of response may be limited to extravascular LVVs. Myalgias, fever, malaise, and headaches improve swiftly, in parallel with a fast reduction of acute-phase reactants (CRP, IL-6, ESR). Emerging data suggest that the vascular component is much more resistant to immunosuppression and may require entirely new therapeutic strategies.

Once the condition is stabilized, steroid tapering is guided by close monitoring of the clinical presentation as well as laboratory markers of inflammation. In general, steroids should be reduced by 10–20% every 2 weeks. Monthly monitoring of ESR and CRP is mandatory to adjust therapy. Patients frequently return with signs or symptoms of recurrent disease as immunosuppression is lowered. Fortunately, disease exacerbations causing vision loss are infrequent. Disease flare-ups typically present with PMR symptoms or nonspecific manifestations of malaise and failure to thrive. In most patients, a transient small increase in the steroid dose reinstates disease control.

Much effort has been invested in identifying steroid-sparing agents. In a small study, treatment with pulse corticosteroids appeared to have long-term beneficial effects, reducing the overall steroid requirement and the rate of disease flare-ups.[40] After pulses of 1000 mg methylprednisolone were administered for 3 consecutive days, therapy was continued with oral prednisone, and daily doses were swiftly tapered. Compared with the control arm, patients who received three initial steroid pulses had lower likelihoods of disease flare-ups. Particularly, once they reached prednisone doses close to 10 mg/day, these patients could tolerate steroid withdrawal significantly better, and most were taking 5 mg/day prednisone at 36 weeks.[40] The benefit from initial pulse therapy continued over subsequent months, suggesting the potential benefit of intense immunosuppression during early disease.

Several biological agents have been explored or are currently undergoing testing in clinical trials.[41] Tumor necrosis factor-α (TNF-α) inhibitors may have a role in TA but had no steroid-sparing effect in GCA.[42] Preliminary study results suggest that targeting T-cell costimulation with abatacept may prevent disease relapses in GCA. Ustekinumab, a monoclonal antibody (mAb) targeting IL-12 and IL-23, was reported to have potential efficiency in refractory GCA in a small open-label study. The IL-6 receptor blocker tocilizumab has been explicitly effective in reducing acute phase reactants (CRP, ESR) and is being explored in GCA and TA and a phase 3 double-blind trial of tocilizumab given weekly or every other week demonstrated substantial steroid-sparing effect over a one-year period.[12]

Maintenance Therapy

With a major shift in the pathogenic concept about LVVs, especially the realization that the disease process has two, partly independent components (extravascular, vascular) and that vessel wall infiltrates persist chronically, the therapeutic needs for maintenance therapy have become the dominant issue for the treating physician. Patients with PMR are often managed successfully with low-dose corticosteroids (prednisone 5 mg daily) and typically are highly responsive to transient and very small dose increases (1–2 mg prednisone/day). Long-term management of patients with GCA and TA relies on low-dose corticosteroids as well unless there is objective evidence for progressive vascular wall inflammation. Unfortunately no reliable biomarkers can separate the extravascular and vascular disease components, and no evidence that suppressing acute phase response in the periphery will ultimately restrict transmural vasculitis has been presented.

Methotrexate is considered to have mild-to-moderate steroid-sparing potential in GCA and PMR[43] but is more frequently used in TA.[44] When given to human artery–severe combined immunodeficiency (SCID) chimeras, acetylsalicylic acid (aspirin) has marked antiinflammatory activities, with suppression of IFN-γ in vascular lesions. Clinical trials are needed to test whether this immunosuppressive action can translate into corticosteroid sparing. Because arteries are the primary targets of LVVs, the use of aspirin as an antiplatelet agent should be routinely recommended.

There is no evidence that immunosuppressants, such as azathioprine and cyclophosphamide, lower steroid needs, prevent vascular complications, or shorten the duration of steroid use. Whether any of the biological agents described above has a place to effectively suppress vessel wall inflammation and change the course of chronic disease is currently unknown.

An integral part of chronic immunosuppression with prednisone is regular monitoring for diabetes and hypertension. Patients should be encouraged to increase physical activity, as steroid-induced myopathy occurs frequently. A major issue of chronic steroid treatment, particularly in older individuals, is the risk of excessive bone loss, possibly resulting from increased bone resorption and impaired bone formation. Several effective and safe therapies for osteopenia/osteoporosis are available.

Calcium and vitamin D supplementation should be part of the therapeutic regimen.

In many, but not all, patients, immunosuppressive treatment can be discontinued 18–24 months after diagnosis. Markers of systemic inflammation may remain elevated, and continuous monitoring for aortic involvement and recurrence of cranial arteritis is recommended.

Most patients with PMR are sufficiently treated with an initial dose of 20 mg of prednisone per day. In some patients, 10 mg of prednisone can induce and sustain a clinical response. Steroids should be titrated to minimally needed doses to avoid side effects; tapering usually needs to be slow, over many months. In TA, long-term management should be tailored to individual patient conditions.[44] It has been argued that patients should be maintained on a low dose of corticosteroids, such as 5–7 mg prednisone daily, even after successful control of active disease.

Given the age at disease onset in TA, preventive measures to counteract accelerated atherosclerosis and optimize blood pressure control are important aspects of management.

It has been suggested that up to 50% of patients with TA may require a second immunosuppressive agent.[44] Steroid-sparing effects of methotrexate have been reported for some patients. Similarly, mycophenolate mofetil may have clinical efficiency, although only published data on a small patient cohort are available. Empirically, azathioprine may have a place in maintenance therapy of patients with TA. Finally, there may be a place for agents blocking TNF-α in patients with persistent disease activity. Results from well-designed placebo-controlled treatment trials testing the efficiency of such immunosuppressive drugs are awaited.

Detecting and treating hypertension is an essential component of caring for patients with TA. Untreated hypertension leads to acceleration of atherosclerosis and cardiac insufficiency. In patients with upper-extremity involvement, obtaining accurate blood pressure measurements is a challenge and requires education of the patient and caregivers.

Revascularization Procedures

Besides pharmacological therapy, revascularization procedures—including both surgical and endovascular interventions—have vastly broadened therapeutic options in patients with TA and large-vessel GCA.[45] To minimize the risk of complications, such as rapid reocclusion, an effort should be made to suppress vascular wall inflammation, ideally before subjecting the patients to revascularization treatment. Conventional bypass grafts are still considered the method of choice. Percutaneous transluminal angioplasty can be useful in managing renal artery stenosis or other short-segment lesions. Bypass surgery is needed in patients with cerebrovascular ischemia in whom catastrophic strokes may be prevented by bypassing critical stenosis of cervical vessels with grafts originating from the aortic arch. Reestablishing flow in the upper- and lower-extremity arteries can be complicated by multiple and long-segment stenosis, and arterial reconstructions with prosthetic graft materials or veins may be the only alternative to obtain long-term patency. Placing of conventional stents can be complicated by eliciting rapid restenosis, and it is not known whether outcomes can be improved by drug-eluting stents. Occlusive disease of the coronary arteries usually represents a challenging clinical scenario, and most physicians opt for conventional bypass surgery. Depending on symptoms, patients with aortic regurgitation may require repair of the weakened aortic wall.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Weyand CM, Goronzy JJ. Medium- and large-vessel vasculitis. N Engl J Med 2003;349(2):160–9.
2. Salvarani C, Cantini F, Boiardi L, et al. Polymyalgia rheumatica and giant-cell arteritis. N Engl J Med 2002;347(4):261–71.
3. Nordborg E, Nordborg C. Giant cell arteritis: epidemiological clues to its pathogenesis and an update on its treatment. Rheumatology (Oxford) 2003;42(3):413–21.
4. Vanoli M, Bacchiani G, Origg L, et al. Takayasu's arteritis: a changing disease. J Nephrol 2001;14(6):497–505.
5. Isobe M. Takayasu arteritis revisited: current diagnosis and treatment. Int J Cardiol 2013;168(1):3–10.
6. Numano F, Kishi Y, Tanaka A, et al. Inflammation and atherosclerosis. Atherosclerotic lesions in Takayasu arteritis. Ann N Y Acad Sci 2000;902:65–76.
7. Wagner AD, Goronzy JJ, Weyand CM. Functional profile of tissue-infiltrating and circulating CD68+ cells in giant cell arteritis. Evidence for two components of the disease. J Clin Invest 1994;94(3):1134–40.
8. Baldini M, Maugeri N, Ramirez GA, et al. Selective up-regulation of the soluble pattern-recognition receptor pentraxin 3 and of vascular endothelial growth factor in giant cell arteritis: relevance for recent optic nerve ischemia. Arthritis Rheum 2012;64(3):854–65.
9. Nadkarni S, Dalli J, Hollywood J, et al. Investigational analysis reveals a potential role for neutrophils in giant-cell arteritis disease progression. Circ Res 2014;114(2):242–8.
10. Roche NE, Fulbright JW, Wagner AD, et al. Correlation of interleukin-6 production and disease activity in polymyalgia rheumatica and giant cell arteritis. Arthritis Rheum 1993;36(9):1286–94.
11. Weyand CM, Fulbright JW, Hunder GG, et al. Treatment of giant cell arteritis: interleukin-6 as a biologic marker of disease activity. Arthritis Rheum 2000;43(5):1041–8.
12. Stone JH, Tuckwell K, Dimonaco S, et al. Trial of tocilizumab in giant-cell arteritis. N Engl J Med 2017;377:317–28.
13. O'Neill L, Rooney P, Molloy D, et al. Regulation of inflammation and angiogenesis in giant cell arteritis by acute-phase serum amyloid A. Arthritis Rheumatol 2015;67(9):2447–56.
14. Weyand CM, Goronzy JJ. Immune mechanisms in medium and large-vessel vasculitis. Nat Rev Rheumatol 2013;9(12):731–40.
15. Pryshchep O, Ma-Krupa W, Younge BR, et al. Vessel-specific Toll-like receptor profiles in human medium and large arteries. Circulation 2008;118(12):1276–84.
16. Ma-Krupa W, Jeon MS, Spoerl S, et al. Activation of arterial wall dendritic cells and breakdown of self-tolerance in giant cell arteritis. J Exp Med 2004;199(2):173–83.
17. Krupa WM, Dewan M, Jeon MS, et al. Trapping of misdirected dendritic cells in the granulomatous lesions of giant cell arteritis. Am J Pathol 2002;161(5):1815–23.
18. Weyand CM, Liao YJ, Goronzy JJ. The immunopathology of giant cell arteritis: diagnostic and therapeutic implications. J Neuroophthalmol 2012;32(3):259–65.
19. Weyand CM, Wagner AD, Bjornsson J, et al. Correlation of the topographical arrangement and the functional pattern of tissue-infiltrating macrophages in giant cell arteritis. J Clin Invest 1996;98(7):1642–9.
20. Weyand CM, Schonberger J, Oppitz U, et al. Distinct vascular lesions in giant cell arteritis share identical T cell clonotypes. J Exp Med 1994;179(3):951–60.
21. Deng J, Younge BR, Olshen RA, et al. Th17 and Th1 T-cell responses in giant cell arteritis. Circulation 2010;121(7):906–15.
22. Watanabe R, Hosgur E, Zhang H, et al. Pro-inflammatory and anti-inflammatory T cells in giant cell arteritis. Joint Bone Spine 2016;pii: S1297-319X(16)30124-5. doi:10.1016/j.jbspin.2016.07.005. [Epub ahead of print].

23. Seko Y, Minota S, Kawasaki A, et al. Perforin-secreting killer cell infiltration and expression of a 65-kD heat-shock protein in aortic tissue of patients with Takayasu's arteritis. J Clin Invest 1994;93(2):750–8.
24. Wen Z, Shimojima Y, Shirai T, et al. NADPH oxidase deficiency underlies dysfunction of aged CD8+ Tregs. J Clin Invest 2016;126(5):1953–67.
25. Zhang H, Watanabe R, Berry GJ, et al. Immunoinhibitory checkpoint deficiency in medium and large vessel vasculitis. Proc Natl Acad Sci USA 2017;114(6):E970–9.
26. Weyand CM, Goronzy JJ. Giant-cell arteritis and polymyalgia rheumatica. Ann Intern Med 2003;139(6):505–15.
27. Brack A, Rittner HL, Younge BR, et al. Glucocorticoid-mediated repression of cytokine gene transcription in human arteritis-SCID chimeras. J Clin Invest 1997;99(12):2842–50.
28. Salvarani C, Cantini F, Olivieri I, et al. Proximal bursitis in active polymyalgia rheumatica. Ann Intern Med 1997;127(1):27–31.
29. Caspary L. Inflammatory diseases of the aorta. Vasa. 2016;45(1):17–29.
30. Kobayashi Y, Numano F. 3. Takayasu arteritis. Intern Med 2002;41(1):44–6.
31. Hunder GG, Bloch DA, Michel BA, et al. The American College of Rheumatology 1990 criteria for the classification of giant cell arteritis. Arthritis Rheum 1990;33(8):1122–8.
32. Dasgupta B, Cimmino MA, Kremers HM, et al. 2012 Provisional classification criteria for polymyalgia rheumatica: a European League Against Rheumatism/American College of Rheumatology collaborative initiative. Arthritis Rheum 2012;64(4):943–54.
33. Arend WP, Michel BA, Bloch DA, et al. The American College of Rheumatology 1990 criteria for the classification of Takayasu arteritis. Arthritis Rheum 1990;33(8):1129–34.
34. Kerr GS, Hallahan CW, Giordano J, et al. Takayasu arteritis. Ann Intern Med 1994;120(11):919–29.
35. Achkar AA, Lie JT, Hunder GG, et al. How does previous corticosteroid treatment affect the biopsy findings in giant cell (temporal) arteritis? Ann Intern Med 1994;120(12):987–92.
36. Jia L, Couce M, Barnholtz-Sloan JS, et al. Is all inflammation within temporal artery biopsies temporal arteritis? Hum Pathol 2016;57:17–21.
37. Steeds RP, Mohiaddin R. Takayasu arteritis: role of cardiovascular magnetic imaging. Int J Cardiol 2006;109(1):1–6.
38. Hartlage GR, Palios J, Barron BJ, et al. Multimodality imaging of aortitis. JACC Cardiovasc Imaging. 2014;7(6):605–19.
39. Goronzy JJ, Weyand CM. Understanding immunosenescence to improve responses to vaccines. Nat Immunol 2013;14(5):428–36.
40. Mazlumzadeh M, Hunder GG, Easley KA, et al. Treatment of giant cell arteritis using induction therapy with high-dose glucocorticoids: a double-blind, placebo-controlled, randomized prospective clinical trial. Arthritis Rheum 2006;54(10):3310–18.
41. Koster MJ, Matteson EL, Warrington KJ. Recent advances in the clinical management of giant cell arteritis and Takayasu arteritis. Curr Opin Rheumatol 2016;28(3):211–17.
42. Hoffman GS, Cid MC, Rendt-Zagar KE, et al. Infliximab for maintenance of glucocorticosteroid-induced remission of giant cell arteritis: a randomized trial. Ann Intern Med 2007;146(9):621–30.
43. Mahr AD, Jover JA, Spiera RF, et al. Adjunctive methotrexate for treatment of giant cell arteritis: an individual patient data meta-analysis. Arthritis Rheum 2007;56(8):2789–97.
44. Liang P, Hoffman GS. Advances in the medical and surgical treatment of Takayasu arteritis. Curr Opin Rheumatol 2005;17(1):16–24.
45. Mason JC. Takayasu arteritis: surgical interventions. Curr Opin Rheumatol 2015;27(1):45–52.

MULTIPLE CHOICE QUESTIONS

1. A 27-year old female presents with a 6-weeks history of intermittent fever, arthralgias and weight loss. She was admitted after she fainted in the emergency room. On physical examination she has no palpable peripheral pulse in her left upper extremity, and Takayasu arteritis (TA) is added as a differential diagnosis.
   Which of the following statements is true:
   A. She should be scheduled for temporal artery biopsy, to rule out giant cell arteritis.
   B. Laboratory tests, e.g. sedimentation rate will be of no use, as they can be non-specifically abnormal.
   C. A pregnancy test should be performed.
   D. A detailed physical examination will not be helpful.
   E. If a diagnosis of TA is made, it will be a self-limited disease.

2. Giant cell arteritis and Takayasu arteritis have a stringent tissue tropism. Which of the following vascular beds can be involved? Identify all correct answers:
   A. Aortic arch
   B. Carotid artery
   C. Subclavian artery
   D. Renal capillaries
   E. Arteries in the skin

3. Extravascular GCA may result in the following laboratory abnormalities. Identify the incorrect answer:
   A. Elevated C-reactive protein
   B. Autoantibodies to nuclear antigens
   C. Thrombocytosis
   D. Elevated serum amyloid A
   E. Elevated alkaline phosphatase

60

Systemic Autoinflammatory Syndromes

Catharina M. Mulders-Manders, Jeroen C.H. van der Hilst,
Jos W.M. van der Meer, Anna Simon

Autoinflammatory diseases, which are also known as *periodic fever syndromes,* encompass a group of rare disorders characterized by recurrent or persistent inflammation. *Autoinflammation* is a term that has been used since the late 1990s to illustrate the difference between autoimmune disorders and diseases characterized by exuberant inflammation. Typically, autoinflammatory diseases do not show features of excess adaptive immune system activation, and autoantigens or auto-antigen specific T-cells are not present in these diseases. It is now recognized that autoinflammation and autoimmunity form two ends of a spectrum of inappropriate immune system activation and share several common features.

Located at the autoinflammatory end of this spectrum are the classic monogenic autoinflammatory diseases: familial Mediterranean fever (FMF), cryopyrin-associated periodic syndrome (CAPS), mevalonate kinase deficiency (MKD; also known as *hyperimmunoglobulin D and periodic fever syndrome* [HIDS]), and tumor necrosis factor (TNF) receptor–associated periodic syndrome (TRAPS). The number of autoinflammatory diseases is increasing rapidly. New monogenic autoinflammatory diseases have been identified in the last decades. For many of the recently described autoinflammatory diseases, no genetic cause has been found yet.

It has also become clear that autoinflammation is at least partially involved in the pathogenesis of other, more common diseases, such as gout, Crohn disease, and ulcerative colitis.

As it is impossible to discuss all autoinflammatory diseases in detail here, the classic monogenic diseases FMF, CAPS, TRAPS, and MKD have been selected as the main focus of this chapter. Their pathophysiological mechanisms are understood to a much higher degree than in many newer autoinflammatory diseases, and their clinical presentations have been described precisely.

In addition, two other autoinflammatory diseases are discussed, one with relatively high prevalence and the other because of its interesting pathophysiological mechanism: *(i)* periodic fever, aphthous stomatitis, pharyngitis, and adenitis (PFAPA) syndrome and *(ii)* Schnitzler syndrome.

The cornerstone of diagnosing an autoinflammatory disease is the clinical assessment of the patient. This includes a detailed medical and family history and direct observation of an inflammatory episode. The first step in the diagnostic process is to exclude other more common causes of recurrent inflammation, including infections, malignancy and paraneoplastic phenomena, and autoimmune disease.[1] A first differential diagnosis can be made on the basis of age of onset, associated signs and symptoms, duration of inflammation, family history, and ethnic background, (Table 60.1), and this can guide targeted diagnostic testing.

KEY CONCEPTS

Autoinflammation Versus Autoimmunity

- Common features:
  - Inflammation due to excessive immune activation
  - Phenotypes characterized by exacerbations and remissions
- Distinctive features:
  - Autoinflammation: dysregulation of innate immunity, no high-titer autoantibodies or autoantigen-specific T cells
  - Autoimmunity: dysregulation of adaptive immunity, defect in lymphocyte function, autoantibodies may be present.
- Autoinflammation and autoimmunity form two ends of a continuous spectrum of excessive immune system activation.
- Many diseases show overlapping features between autoinflammation and autoimmunity.

EPIDEMIOLOGY

It is important to realize that the incidence of specific diseases varies widely among ethnic groups. With more than 100 000 patients worldwide, FMF is the most prevalent monogenic autoinflammatory disease. It is most common in individuals originating from around the Mediterranean basin, such as Turks, Jews (primarily non-Ashkenazi), Arabs, and in Armenians. In these selected populations, the carrier frequency of mutations in the *MEFV* gene can be as high as one in three individuals.[1] This may indicate a survival benefit for carriers of heterozygous mutations, possibly through protection against certain unknown infectious agents.

The first patients with MKD were described in 1984 in The Netherlands[2] (then referred to as HIDS). Over 200 patients have now been identified, most of Western European and Caucasian ancestry. This could be partly explained by increased awareness for this disease among physicians in that part of the world. An alternative explanation is a common founder effect with clustering of carriers, illustrated by a carrier rate of 1:153 for the most common mutation in the mevalonate kinase gene (*MVK*) (V377I) in Dutch newborns.[3,4]

TRAPS is seen in patients from around the world, although most patients originate from northwestern Europe. A few dozen families and over 200 sporadic cases have been reported.

The exact prevalence of the CAPS is unknown, but over 130 cases have been recognized. Disease awareness and recognition among clinicians have improved because of the availability of effective treatment for this disease.

PFAPA syndrome was first reported at the end of the 1980s.[5] It is difficult to estimate the incidence of PFAPA, as the level of

TABLE 60.1 Hereditary Autoinflammatory Syndromes

	FMF	CAPS	TRAPS	MKD
Mode of inheritance	Autosomal recessive	Autosomal dominant	Autosomal dominant	Autosomal recessive
Age of onset (years)	<20	Generally <1, in MWS/FCAS <20 possible	Variable, most <10	<1
Main ethnic distribution	Turks, Arabs, Jews, Armenians	Europeans	All	Northwestern Europeans (Dutch, French)
Gene involved	*MEFV*	*NLRP3*	*TNFRSF1A*	*MVK*
Protein involved	Pyrin	NLRP3	TNF receptor type 1	Mevalonate kinase
Duration of typical attack	2–3 days	Variable; hours–days or continuous inflammation	Days–weeks	HIDS: 4–6 days; MA: continuous, flares possible
Distinguishing symptoms	Peritonitis, pleuritis, erysipelas-like skin lesions	Aseptic meningitis; sensorineural deafness; bone lesions, dysmorphic features May be cold-induced	Severe myalgia, periorbital edema	HIDS: lymphadenopathy, attacks induced by vaccination MA: joint contractions, growth and developmental delay
Risk of amyloidosis[a]	Up to 75%	Up to 33%	25%	<5%
Treatment	Colchicine, combination with IL-1 inhibition when resistant	IL-1 inhibition	Mild disease: NSAIDs; Severe disease: IL-1 inhibition	IL-1 inhibition

[a] In patients with long-term uncontrolled inflammation.

CAPS, cryopyrin-associated periodic syndrome; FMF, familial Mediterranean fever; HIDS, hyperimmunoglobulin D and periodic fever syndrome; IL, interleukin; MA, mevalonic aciduria; MKD, mevalonate kinase deficiency; TRAPS, tumor necrosis factor (TNF) receptor–associated periodic syndrome.

awareness of this disease seems to vary among clinicians. A single pediatric center in the United States reported 122 patients fulfilling the criteria for PFAPA in 10 years,[6] making it more common than any of the monogenic autoinflammatory diseases (with the exception of FMF in certain populations). In most patients with PFAPA, symptoms cease before or during adolescence. The cause of this spontaneous resolution is unknown. The characteristic combination of symptoms of PFAPA has been described in adults, but it remains a matter of debate whether these patients suffer from true PFAPA syndrome.

Schnitzler syndrome, first described by the French dermatologist Schnitzler in 1972,[7] is an acquired autoinflammatory disorder with a median age of onset of 51 years. Over 160 cases have been reported worldwide.

SIGNS AND SYMPTOMS

Familial Mediterranean Fever

FMF is an autosomal recessive disease. Over 90% of patients become symptomatic within the first two decades of life. Typically, attacks are characterized by abrupt onset of high fever, peaking soon after onset and lasting for 12 hours to 3 days. Subsequently, the fever subsides rapidly. Painful serositis accompanies the fever. Serositis can also be present without fever. Over 95% of patients experience abdominal pain, which lasts up to 3 days. The pain, which is caused by sterile peritonitis, may initially be focal and progress to more diffuse pain. Before being diagnosed with FMF, a significant proportion of patients will have undergone exploratory abdominal surgery under suspicion of appendicitis. At surgery, intraabdominal adhesions, a result of recurrent peritonitis, may be found. Pelvic adhesions can reduce fertility in female patients. Pleuritis, presenting as thoracic pain, is experienced by approximately 40% of patients. Synovitis with monoarthritis of knee, ankle, or wrist occurs in one-half to three-quarters of patients. An arthritic attack may have a more protracted course compared with nonarthritic FMF, with fever lasting up to a week. Joint pain may persist when fever has already subsided. Synovitis usually resolves completely without joint destruction. The skin can be affected. Erysipelas-like skin lesions overlying the shins are very characteristic of FMF, but are only seen in 30% of patients. Less frequent symptoms of FMF include vasculitis, orchitis, aseptic meningitis, and myalgia. Pericarditis is rare in FMF.

There are no consistent triggers for FMF attacks. Emotional stress or menstruation may increase the frequency of attacks; some patients are able to report specific triggers for their attacks. Attack frequency varies greatly among patients and during an individual patient's life. Attacks may be as frequent as 2–3 times each month and as rare as less than once a year.

The literature on FMF is replete with genotype–phenotype studies. The most consistent finding is that carriers of the M694V/M694V genotype have more severe disease, with earlier onset and higher frequency of arthritis and long-term complications.

Life expectancy of patients with FMF depends on timely initiation of appropriate treatment to prevent amyloidosis. Without amyloidosis, FMF patients have normal life expectancy.

Cryopyrin-Associated Periodic Syndrome

CAPS is autosomal dominantly inherited. Originally, three separate clinical syndromes, all with their own clinical features, were distinguished: familial cold autoinflammatory syndrome (FCAS), Muckle Wells syndrome (MWS), and neonatal onset multisystem inflammatory disease (NOMID), which is also known as chronic infantile neurological, cutaneous and articular (CINCA) syndrome. With the discovery of *NLRP3* mutations in all three diseases, it has become clear that the clinical phenotype of CAPS is a continuous spectrum of severity, instead of distinct diseases. There is no genotype–phenotype association, suggesting a role for other yet undiscovered disease modifying factors.

CAPS often manifests clinically soon after birth or in early childhood. It is characterized by recurrent urticaria-like rash, arthralgia, myalgia, headache, and fever. Ocular symptoms, including conjunctivitis and uveitis, are common. Some patients develop sensorineural hearing loss during adolescence or adulthood. At the severe end of the clinical spectrum, central nervous system (CNS) symptoms, including chronic aseptic meningitis that is characterized by chronic headache, increased intracranial pressure, hydrocephalus, mental retardation, and seizures, are common. Papilledema with optic nerve atrophy can lead to loss of vision. In severely affected patients, arthropathy, with distinct radiographic findings of premature patellar and epiphyseal long bone ossification and osseous overgrowth, develops early in life. If left untreated, this leads to growth retardation, severe joint contractures, and persisting disability.

The duration of attacks is variable and ranges from hours to days. At the severe end of the spectrum, patients have continuous inflammation. Attacks may be triggered by exposure to cold, minor trauma, or emotional stress.

In the past, patients with severe CAPS often died in childhood. This changed after the introduction of anti–interleukin-1 (IL-1) therapy, which is very effective in treating CAPS. Overall, patients without neurological involvement are now believed to have a normal life expectancy.

Tumor Necrosis Factor Receptor–Associated Periodic Syndrome

TRAPS is inherited in an autosomal dominant fashion. Age of onset varies widely. Many patients become symptomatic within the first years of life, with a median age of onset of 3 years, but adult onset is also possible. The usual duration of fever in TRAPS is considerably longer than in the other classic autoinflammatory syndromes: attacks persist for a minimum of 3 days but can last for several weeks. The interval between attacks in a single patient can vary substantially.

Localized myalgia, a deep cramping, and often severely disabling pain in a single limb resulting from monocytic fasciitis and associated with fever is found in virtually all patients. The affected limb may show local erythema, which may migrate to the distal part of the extremity (Fig. 60.1). Almost all patients have abdominal pain, often accompanied by vomiting, constipation, and bowel obstruction. Arthralgia and monoarthritis involving hips, knees, or ankles are present in 25% of patients at some point. Chest pain is frequent and can be caused by pleuritis or may be musculoskeletal in origin. Ocular symptoms range from conjunctivitis and periorbital pain to severe uveitis and iritis. Periorbital edema with conjunctival injection is a distinctive, but infrequent, feature of TRAPS. Other less frequently observed symptoms are pericarditis and lymphadenopathy.

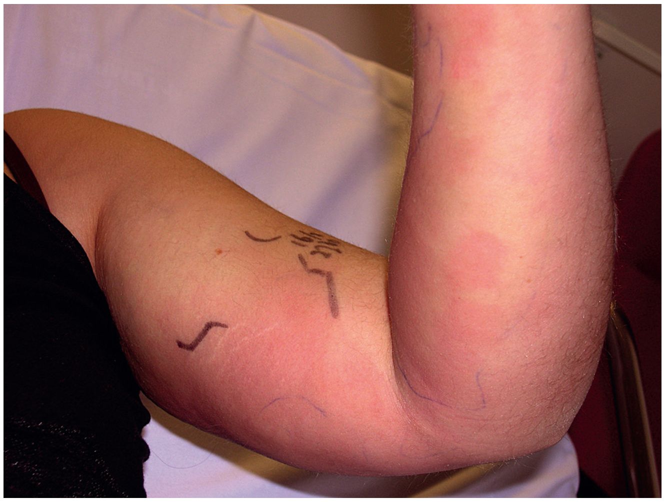


FIG 60.1 Migratory erythematous macular rash during an inflammatory attack in a patient with tumor necrosis factor receptor–associated periodic syndrome (TRAPS).

Mevalonate Kinase Deficiency

Before the discovery of the underlying genetic defect in the mevalonate kinase gene, two distinct autosomal recessive diseases were distinguished, which are now known to form two ends of a continuous spectrum: HIDS at the less severe end and mevalonic aciduria (MA) at the most severe end.

HIDS is characterized by recurrent fever attacks that last for 4–6 days, starting in early childhood. The inflammatory attacks occur on average every 4–6 weeks. Attack frequency varies in a single patient and among patients and tends to decrease later in life. Attacks often start with chills, followed by a rapid rise in temperature. Factors that can provoke an attack are infections, trauma, vaccination, and both physical and emotional stress, although a clear trigger is often absent. Characteristic for HIDS is the first attack being triggered by childhood vaccination.

Fever is accompanied by cervical lymphadenopathy and abdominal pain with vomiting and diarrhea. The skin may show erythema, papules, urticarial rash, or exanthema. The majority of patients suffer from large-joint arthralgia or arthritis. Oral or genital aphthous ulcers may be present during attacks. Hepatosplenomegaly has been reported. Patients with HIDS appear to have normal life expectancy and experience no complications.

MA is located at the severe end of the MKD spectrum. This severe disease is present from birth and is characterized by psychomotor retardation, ataxia, failure to thrive, cataracts, and facial dysmorphia. Episodic fever or inflammation are present in MA. Many patients die in early childhood.

In recent years, it has become clear that the spectrum of MKD comprises more than only these two diseases. Mutations in *MVK* have been found in the absence of typical MKD features in patients with retinitis pigmentosa and early-onset ulcerative colitis. Mutations in *MVK* have also been found in patients with the skin diseases disseminated superficial actinic porokeratosis and porokeratosis of Mibelli, cyclic neutropenia, and macrophage activation syndrome, but there is no evidence of decreased mevalonate kinase activity in these patients.[8] Mutations in these diseases may overlap with mutations that may cause MA and HIDS.

Periodic Fever, Aphthous Stomatitis, Pharyngitis, and Adenitis Syndrome

PFAPA is primarily a childhood disease with a usual onset before the age of 5 years. Patients suffer from recurring episodes of fever that generally last for 3–6 days and recur with great regularity. Additional symptoms include pharyngitis, cervical adenitis, and aphthous stomatitis. Other symptoms may include headache, vomiting and mild abdominal pain, arthralgia, and myalgia. Between fever episodes, patients are symptom-free. In most patients, attacks cease after several years, often before or during adolescence.

Schnitzler Syndrome

A typical feature of Schnitzler syndrome is its late onset, at a median age of 51 years. Patients typically present with chronic recurrent and mostly nonpruritic urticarial rash. This can be accompanied by fever, arthralgia or arthritis, and bone pain (Table 60.2). Symptoms progress over years. The presence of monoclonal paraproteinemia, typically of immunoglobulin M (IgM), is characteristic for Schnitzler syndrome. Presence of monoclonal IgG is less common and is sometimes referred to as *variant Schnitzler syndrome.* Onset of symptoms can precede paraproteinemia for

TABLE 60.2 **Diagnostic Criteria for Schnitzler Syndrome**[a]

Major Criteria (≥1 present)
(Chronic) urticarial rash
Monoclonal immunoglobulin M (IgM; or IgG: variant type)
Minor Criteria (≥2 present)
Intermittent fever
Arthralgia or arthritis
Bone pain
Lymphadenopathy
Hepatomegaly and/or splenomegaly
Elevated erythrocyte sedimentation rate (ESR) and/or leukocytosis
Bone abnormalities (on radiological or histological examination)

[a] Schnizler syndrome can be diagnosed only after exclusion of other causes.
From: de Koning et al. Diagnostic criteria for Schnitzler syndrome. Semin Arthritis Rheum 2007:37:137–148.

years, but only little is known about this. Symptom severity is unrelated to the level and type of paraproteinemia.

An important long-term complication of Schnitzler syndrome is Waldenström macroglobulinemia, which, in a single study, had an incidence of 15% 10 years after diagnosis. Patients with Schnitzler syndrome have a normal life expectancy.

PATHOGENESIS

The common pathophysiological feature of most autoinflammatory diseases is overproduction of the proinflammatory cytokine IL-1β. This protein is produced as an inactive proform (pro–IL-1β), which must be cleaved to become activated. The most common pathway of cleavage is by caspase-1. Like IL-1β, caspase-1 is transcribed as an inactive proform (procaspase-1), and it, too, must be cleaved by a multiprotein complex called the *inflammasome.* Several inflammasomes have been identified, of which the nucleotide-binding oligomerization domain (NOD)–like receptor P3 (NLRP3) inflammasome has been studied in greatest detail.

The NLRP3 inflammasome consists of the central protein NLRP3, the adapter protein ASC, and the effector protein procaspase-1. Upon activation of the inflammasome, procaspase-1 is converted into mature caspase-1, which is then able to cleave inactive pro–IL-1β to its active form (Fig. 60.2).

Familial Mediterranean Fever

FMF is caused by mutations in the Mediterranean fever gene (*MEFV*), which encodes the protein pyrin, primarily expressed in peripheral blood leukocytes, especially neutrophils and monocytes. Pyrin is a member of the pyrin-domain (PYD)–containing proteins, which are able to bind to the PYD domain of other proteins, including apoptosis-associated speck-like protein containing a C-terminal caspase recruitment domain (ASC). Binding of pyrin to ASC leads to activation of ASC, with consequent recruitment and activation of procaspase-1. Pyrin can also bind to the PYD domains of other proteins that are able to initiate apoptosis or activate nuclear factor (NF)-κB, including caspase-8 (Fig. 60.3). These complexes are called *pyrin inflammasomes.*

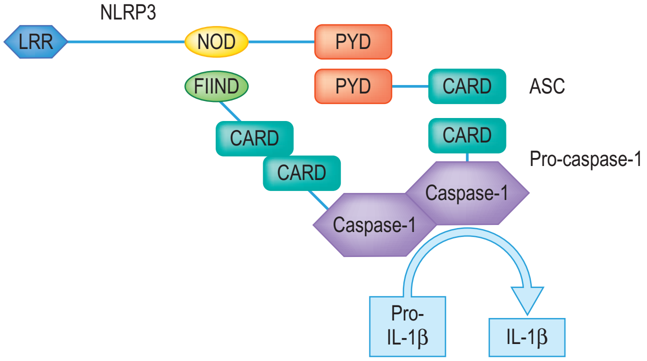


FIG 60.2 The Nucleotide-Binding Oligomerization Domain (NOD)–Like Receptor P3 (NLRP3) Inflammasome. NLRP3 is the central component of this inflammasome. NLRP3 contains three domains: a pyrin domain (PYD), NOD, and a domain of leucine-rich repeats (LRR). Cryopyrin binds apoptosis-associated speck-like protein containing a C-terminal caspase recruitment domain (ASC) through its PYD domain and through the NOD domain. The association of these proteins ultimately leads to the release of active caspase-1, which, in turn, activates interleukin-1β (IL-1β) through the cleavage of pro–IL-1β.

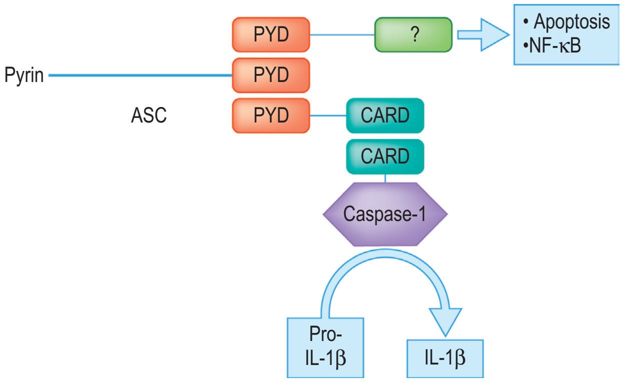


FIG 60.3 Mechanism of Action of Pyrin. Pyrin contains a pyrin domain (PYD) that is able to bind to apoptosis-associated speck-like protein containing a C-terminal caspase recruitment domain (ASC). ASC can recruit caspase-1 via its CARD domain leading to production of mature interleukin-1β (IL-1β). Pyrin is also able to bind to the PYD domain of other proteins involved in inflammation and apoptosis.

So far, over 300 sequence variants in the *MEFV* gene have been reported in the central online Infevers registry (http://fmf.igh.cnrs.fr/ISSAID/infevers/), most of which are clustered in exon 10 of the gene. The six most prevalent mutations (M694V, V726A, M680I, M694I, V694I, E148Q) cause approximately 80% of cases. Recently, autosomal dominant mutations in *MEFV* were described in four families with a colchicine-responsive FMF-like phenotype.[9]

Cryopyrin-Associated Periodic Syndrome

CAPS is caused by mutations in the gene encoding NLRP3. Previous to its discovery at the beginning of this century, the protein was unknown. It was named *cryopyrin,* in analogy with pyrin in FMF and to illustrate the influence of cold exposure in some patients with CAPS. Literature on this gene can be confusing, as the gene previously has also been referred to as *NALP3, PYPAF1,* and *CIAS1.* Mutations associated with CAPS are gain-of-function (GOF) mutations, leading to increased NLRP3 activity.

Tumor Necrosis Factor Receptor–Associated Periodic Syndrome

Mutations in the gene *TNFRSF1A* are responsible for TRAPS. This gene encodes the TNF-receptor superfamily 1A (TNFRSF1A), the main cell surface receptor for TNF. This receptor consists of three domains: an extracellular ligand-binding domain, a transmembrane domain, and an intracellular effector domain. So far, over 100 TNFRSF1A sequence variants have been described, and all TRAPS-associated mutations are located within the extracellular domain of the protein. Upon ligand binding by the extracellular receptor domain, the TNFR forms trimers, triggering the recruitment of intracellular adaptor proteins, which initiate a downstream signaling cascade, leading to NF-κB and mitogen-activated protein kinase (MAPK) activation and caspase-induced apoptosis. When the receptor is activated, the extracellular domain of the TNFR is shed from the membrane. These shed receptors form an extracellular pool of soluble TNFRs, retain their affinity for binding TNF, and are therefore able to mitigate the immune response. Initially, it was hypothesized that TRAPS-associated mutations would lead to defective shedding of TNFR1 receptors, but this hypothesis was discarded as the major pathogenetic mechanism for TRAPS after *in vitro* experiments showed misfolding and intracellular accumulation of mutated proteins. These aggregated receptors retain their normal signaling function and can induce ligand-independent MAPK signalling and production of reactive oxygen species (ROS), resulting in inflammation. (Fig. 60.4)

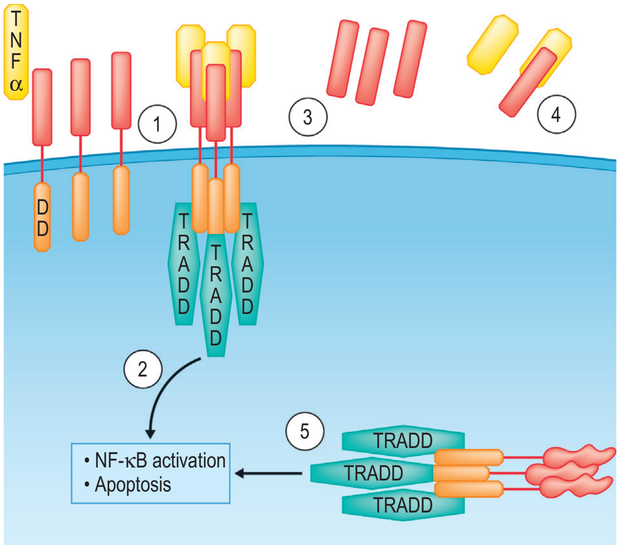


FIG 60.4 **Pathophysiology of Tumor Necrosis Factor Receptor–Associated Periodic Syndrome (TRAPS).** (1) Tumor necrosis factor (TNF) binds to the TNF receptor on the surface of inflammatory cells (2). After receptor triggering, TNF receptor type 1–associated DEATH domain (TRADD) is recruited, inducing a signaling cascade leading to apoptosis and production of proinflammatory cytokines (3). Receptors are shed from the surface, leading to a pool of receptors that dampen immune responses (4). Mutated TNF receptors form aggregates and are retained intracellularly. These aggregated receptors are capable of binding TRADD (5) and stimulate ligand-independent cytokine production.

There are two exceptional *TNFRSF1A* mutations: R92Q and P46L. These mutations do not lead to receptor misfolding and are present in low frequency in the general population. They may, however, cause a mild inflammatory phenotype.

Mevalonate Kinase Deficiency

The genetic defect in MKD is located in *MVK*. Mevalonate kinase is a key enzyme in the isoprenoid pathway and is located directly downstream from 3-hydroxy-3methylglutaryl-coenzyme A reductase (HMG-coA-reductase). The end products of the mevalonate kinase pathway are cholesterol and a number of nonsterol isoprenoids, which are essential compounds in various cellular functions. Mutations in *MVK* lead to reduced mevalonate kinase enzyme activity. In patients with mild disease, residual mevalonate kinase activity is generally 5–15% of healthy controls and is even lower in patients with the severe phenotypes.

The mechanistic link between reduced mevalonate kinase activity and autoinflammation is thought to be defective protein prenylation. Prenylation is a posttranscriptional modification, in which nonsterol isoprenoids are coupled to proteins, influencing protein–protein and protein–membrane interactions.

In a human cellular model of MKD, deficiency of certain isoprenoids were shown to lead to defective prenylation of RhoA, with consequent activation of Rac1 and PKB, which are able to induce IL-1β secretion.[10-12] Defective prenylation with inactivation of RhoA also impairs mitochondrial function. Mitochondria from patients with MKD are elongated and unstable.[12–14] Normally, these abnormal mitochondria would be cleared from the cytosol by autophagy, but in MKD, they accumulate in the cytosol. Abnormal mitochondria release excessive amounts of ROS, and mitochondrial DNA may directly activate NLRP3.[13,14]

Periodic Fever, Aphthous Stomatitis, Pharyngitis, and Adenitis Syndrome

Little is known about the pathophysiology of PFAPA. No genetic defect for PFAPA has been discovered as yet, and this is in agreement with the absence of a clear hereditary pattern. It may be linked to a complex genetic trait. A positive family history has been described, although not all of the patients with a family history were screened for other autoinflammatory diseases.[15,16]

During PFAPA flare-ups, upregulation of complement genes and genes in the IFN–IL-1 pathway are seen. Isolated peripheral blood mononuclear cells (PBMCs) and monocytes of patients with PFAPA show increased IL-1β production without induction of transcription of IL-1β RNA or caspase-1 induction upon lipopolysaccharide (LPS) stimulation. This increased inflammatory response can be abolished by a pan-caspase inhibitor, indicating the important role of the inflammasome in this disease.[17] Spontaneous apoptosis of polymorphous mononuclear cells is significantly lower in patients with PFAPA compared with healthy controls, and during fever episodes, increased production of ROS has been observed *in vitro* in patients with PFAPA.[18]

Schnitzler Syndrome

The etiology of Schnitzler syndrome remains unknown. Involvement of autoreactive antibodies has been suggested, but this finding could not be reproduced.[19] A central role for IL-1β is illustrated by the high efficacy of anti–IL-1β therapy in patients with Schnitzler syndrome.[19,20]

No causative genetic defect for Schnitzler syndrome has been found, but somatic mosaicism of two different *NLRP3* mutations restricted to myeloid cells has been described recently in two patients with Schnitzler syndrome.[21] Strikingly, these patients had the most severe phenotype in this cohort. Somatic mosaicism may be an explanation for the late onset of Schnitzler syndrome, and low-grade mosaicism may not be picked up by routine gene sequencing. Other genetic defects have also been described in very small numbers of patients, but their role in the pathogenesis of Schnitzler syndrome remains unclear.

Schnitzler syndrome is regarded as a paraneoplastic syndrome by some because of its association with Waldenström macroglobulinemia.

LABORATORY TESTS

In classic monogenetic autoinflammatory diseases, an explicit acute-phase response, with elevated inflammatory markers (cross-reactive protein [CRP], erythrocyte sedimentation rate [ESR], serum amyloid A [SAA]) and leukocytosis, is invariably present during symptomatic periods. In other autoinflammatory diseases, this may be less evident.

During clinical disease remission, persistent subclinical inflammation can be found. Cold agglutinins, antinuclear autoantibodies (ANAs), or cryoglobulins are usually not present but may be found in low titers. Proteinuria (over 0.5 g of protein per 24 hours) is highly suggestive for secondary AA amyloidosis.

Elevated serum IgD can be present in MKD. It is discussed in detail in the respective section on the diagnosis of this disease. Similarly, the paraproteins encountered in Schnitzler syndrome are discussed in the section on the symptoms of Schnitzler syndrome.

DIAGNOSIS

Familial Mediterranean Fever

FMF is a clinical diagnosis. A validated set of clinical criteria for the diagnosis of FMF has been defined (The Tel Hashomer criteria, Table 60.3). These criteria have high positive predictive value and negative predictive value in populations with high pretest probability, but their diagnostic accuracy is lower in other populations.

TABLE 60.3 Tel Hashomer Criteria for the Diagnosis of Familial Mediterranean Fever

Major Criteria (≥1 present)
Typical attack[a] with abdominal symptoms
Typical attack[a] with pleural symptoms
Typical attack[a] with monoarthritis
Typical attack[a] with only fever
Incomplete attack[b] with abdominal symptoms
Minor Criteria (≥2 present)
Favorable response to colchicine
Incomplete attack with monoarthritis
Exertional leg pain

[a] Typical attacks are defined as at least three attacks with fever >38°C.

[b] Incomplete attacks are painful and recurrent attacks not meeting the criteria for typical.

The sensitivity and specificity of these criteria for the diagnosis of FMF are >95% and >97%, respectively.

From: Livneh et al. Criteria for the diagnosis of familial Mediterranean fever. Arthritis Rheum 1997;40:1879–85

For experienced physicians, it is not difficult to diagnose FMF in patients who have a medical history compatible with FMF and who come from an ethnic group with high prevalence of FMF. In countries with a low incidence of FMF, several years of diagnostic delay are not unusual, as typical attacks remain unrecognized.

When FMF is suspected on clinical grounds, treatment with colchicine should be initiated immediately. A positive effect of colchicine is confirmatory for the diagnosis FMF.

In populations with a low prevalence of FMF and in atypical cases, sequencing of the *MEFV* gene can be helpful in the diagnostic workup.

Cryopyrin-Associated Periodic Syndrome

CAPS is diagnosed on the basis of typical clinical features, sometimes supported by a positive family history reflecting autosomal dominant inheritance. Detection of mutations in the *NLRP3* gene will confirm the diagnosis in most cases, but cases with “mutation-negative” CAPS have been described. Some of these patients may have somatic mosaicism. Reaction to treatment is the same for these patients as for patients with proven *NLRP3* mutation.

Tumor Necrosis Factor Receptor–Associated Periodic Syndrome

A set of clinical criteria for TRAPS has been proposed (Table 60.4) but has not been validated. The cornerstone of the diagnosis of TRAPS is detection of mutations in the *TNFRSF1A* gene.

Mevalonate Kinase Deficiency

MKD can be suspected when characteristic clinical findings are present in combination with persistent elevated serum levels of IgD >100 international unit per milliliter (IU/mL). Elevation of serum IgD is not pathognomonic for MKD, as it may also occur in other inflammatory conditions, including FMF and PFAPA. Furthermore, in the very young, IgD may be normal, and in some of the affected individuals, IgD is never elevated. Elevated IgD is accompanied by elevated IgA in 80% of patients. The serum level of IgD does not correlate with disease severity, and elevated concentrations are present during asymptomatic periods.

Clinical suspicion of MKD can be confirmed by sequencing of the *MVK* gene. During attacks, traces of mevalonic acid may be found both in urine and serum and can be measured with special techniques, which are, however, not commonly available. Measurement of mevalonate kinase enzyme activity is usually only done in the research setting and requires cell cultures.

TABLE 60.4 Diagnostic Criteria for Tumor Necrosis Factor Receptor–Associated Periodic Syndrome (TRAPS)

- Recurrent episodes of inflammatory symptoms spanning a period of more than 6 months’ duration. Inflammatory symptoms include fever, abdominal pain, myalgia, rash, conjunctivitis/periorbital edema, chest pain, and arthralgia or monoarticular synovitis
- Episodes lasting ≥5 days
- Responsiveness to glucocorticosteroids, but not colchicine
- Affected family members (not always present)
- Any ethnicity may be affected

From: Hull et al. The TNF receptor-associated periodic syndrome (TRAPS): emerging concepts of an autoinflammatory disorder. Medicine (Baltimore) 2002;81:349–68

TABLE 60.5 Diagnostic Criteria for Periodic Fever, Aphthous Stomatitis, Pharyngitis, and Adenitis (PFAPA) Syndrome

I Regularly recurring fevers with an early age of onset (<5 years of age)
II Constitutional symptoms in the absence of upper respiratory infection, which include at least one of the following:
  a) Aphthous stomatitis
  b) Cervical lymphadenitis
  c) Pharyngitis
III Exclusion of cyclic neutropenia
IV Completely asymptomatic interval between episodes
V Normal growth and development

From: Thomas et al. Diagnostic criteria used for PFAPA Periodic fever syndrome in children. J Pediatr 1999;135:15–21

Periodic Fever, Aphthous Stomatitis, Pharyngitis, and Adenitis Syndrome

There is currently no diagnostic test to prove PFAPA. It is diagnosed based on clinical signs and symptoms. The modified criteria by Thomas et al. are used (Table 60.5) and may be supplemented by numeric limits, such as minimum number of attacks and duration of fever. Exclusion of other causes, including other autoinflammatory syndromes, is important.

Schnitzler Syndrome

Schnitzler syndrome is diagnosed on the basis of clinical criteria (see Table 60.2). Exclusion of other causes, particularly monoclonal gammopathy of undetermined significance (MGUS) and chronic idiopathic urticaria, is of paramount importance.

Autoinflammation of Unknown Origin

Regardless of the increasing number of autoinflammatory diseases recognized and the increasing insights in the mechanisms of autoinflammation, an increasing number of patients presenting with an autoinflammatory phenotype cannot be assigned to one of the known autoinflammatory diseases. These patients are considered to suffer from *autoinflammation of unknown origin*. It is likely that further research and newly developed diagnostic techniques will identify new proteins, genetic defects, and pathways in these patients, leading to recognition of new diseases. In patients with autoinflammation of unknown origin, anti–IL-1 therapy can be tried both diagnostically and therapeutically.[21a]

TREATMENT

Colchicine

Colchicine is the treatment of first choice for FMF. Although it has been used since the 1970s, the mechanism of action of colchicine in FMF is still unknown. It is highly effective in preventing attacks. Response to colchicine has been used as a diagnostic criterion for FMF.

The average dose used is 1.0–1.5 mg/day. If tolerated, the dose can be increased up to 3 mg/day in patients with insufficient response. There are few patients with FMF that are unresponsive to colchicine. Others may not be able to tolerate an effective dose of colchicine because of side effects. These patients may benefit from IL-1 inhibition.[22]

As colchicine is the only drug with proven protection against the development of amyloidosis, all patients with FMF should be prescribed colchicine, regardless of disease severity and attack frequency. When anti–IL-1 treatment is started, expert opinion recommends continuation of colchicine in the highest tolerable dose for the prevention of amyloidosis.

The most common side effects of colchicine are diarrhea and abdominal pain. These are dose dependent. In patients with persistent diarrhea, dose reduction can be tried to reduce severity. Myopathy, neuropathy, and leukopenia are very rare, but severe side effects occur primarily in patients with abnormal kidney or liver function or because of interaction with other drugs.

High-dose colchicine has been shown to be teratogenic in animals. However, multiple cohort studies have shown that colchicine can be used safely during pregnancy and breastfeeding. In therapeutic doses, colchicine does not have a negative effect on sperm number and quality, and it has no negative effect on male or female fertility.

As a general rule, there is no place for colchicine in the treatment of autoinflammatory syndromes other than FMF. As an exception to this rule, patients with autoinflammation of unknown origin may experience a favorable effect, especially if their disease shares characteristics with FMF. A small trial has demonstrated decreased attack frequency during colchicine treatment for PFAPA.[23]

Inhibition of Interleukin-1

KEY CONCEPTS

Interleukin-1β (IL-1β)

- IL-1β is a very potent proinflammatory cytokine, tightly regulated at multiple levels.
- The majority of autoinflammatory diseases is caused by dysregulation of IL-1β.
- Measurement of IL-1 β serum levels has no value in the diagnosis of autoinflammation or in assessing disease severity.
- IL-1 inhibition is the treatment of first choice for many autoinflammatory diseases, and it has greatly improved quality of life in patients.

The detection of *NLRP3* mutations in CAPS has illustrated the importance of IL-1β in the pathogenesis of autoinflammation.

The first inhibitor of IL-1 developed was the recombinant human IL-1 receptor antagonist anakinra, which is still the most commonly available IL-1 inhibitor. It competitively binds to the IL-1 receptor, completely inhibiting the actions of both IL-1α and IL-1β. Anakinra has a short half-life and needs to be given as a once-daily subcutaneous injection.

The selective anti-IL-1β monoclonal antibody canakinumab has a longer half-life and is also injected subcutaneously. Standard injection frequency is once per 8 weeks, but shorter intervals may often be necessary in severe disease.

Rilonacept is a construct of two extracellular chains of the IL-1 receptor complex fused to the Fc-portion of IgG. It is given as a weekly subcutaneous injection.

Currently, there is evidence for the effectiveness of anti–IL-1 therapy in many autoinflammatory diseases, including FMF, CAPS, TRAPS, MKD, PFAPA, and Schnitzler syndrome. Typically, anti–IL-1 treatment leads to instant abortion of inflammation with clinical response within the first hours to days after the first injection. This vivid response to treatment is so characteristic that IL-1 inhibition can serve as a diagnostic tool in the diagnosis of these autoinflammatory disorders. Anakinra and canakinumab have both been approved for the treatment of CAPS by the US American Food and Drug Association (FDA) and the European Medical Association (EMA). These agents were given "orphan" status for the treatment of TRAPS by the EMA. In 2016, the FDA and EMA approved the use of canakinumab in patients with HIDS, TRAPS and colchicine-resistant FMF. Rilonacept is approved for the treatment of CAPS by the FDA but is not commonly used in Europe.[24]

In patients with mild MKD and periodic symptoms with long symptom-free intervals, anakinra can be used on demand. In these cases it can be started at the first signs of an attack and only be continued for a few days.[25]

Side effects of IL-1 inhibition include painful injection-site reactions, which are most commonly seen with anakinra, and increased frequency of infections, mostly mild upper respiratory tract infections. More severe infections are rarely seen.

Inhibition of Interleukin-6

Tocilizumab, a monoclonal antibody (mAb) against the IL-6 receptor, has been registered for the treatment of rheumatoid arthritis and systemic-onset juvenile idiopathic arthritis (SoJIA) and is used with increasing frequency in autoinflammatory diseases. It is available as an intravenous infusion or subcutaneous injection. The most common dose is 8 mg/kg in children and adults, or 10 to 12 mg/kg in children with body weight under 30 kg, with an interval of 2–4 weeks. Side effects of tocilizumab are increased susceptibility to infections, most commonly upper respiratory tract infections, elevated liver enzymes, and hematological abnormalities. Bowel perforations have been reported. As IL-6 induces the production of CRP in the liver, anti–IL-6 therapy always normalizes CRP, making it impossible to use it as a marker for disease activity. Tocilizumab has been shown to be effective in patients with anakinra-resistant Schnitzler syndrome, MKD, and TRAPS.

Inhibition of Tumor Necrosis Factor

Three widely used inhibitors of TNF are the mAbs infliximab and adalimumab and the recombinant soluble TNF receptor etanercept. The most common side effect of TNF inhibition is increased risk of serious infections.

Originally, TNF inhibition was regarded the treatment of first choice in patients with TRAPS unresponsive to nonsteroidal antiinflammatory drugs (NSAIDs). However, TNF inhibition induces complete response only in a minority of patients with TRAPS and is therefore far less effective than anti–IL-1 treatment.[26]

Anti-TNF treatment may also be effective in MKD, but responses are mostly partial. It may be tried in patients who show unsatisfactory response to anti–IL-1 treatment.[24,26]

Corticosteroids

Corticosteroids are very effective in PFAPA. Prednisone 1 mg/kg is most commonly used during attacks, although lower doses may also be effective. Use of corticosteroids may increase attack frequency in PFAPA.[26] Mild TRAPS can be treated with a short course of steroids (30 mg daily for 7 days), and patients experiencing more severe attacks may respond to higher doses. The beneficial effect of corticosteroids may decrease over time, necessitating dose escalation.[24,26] Short-term corticosteroids may also be effective in patients with mild MKD.[24,26]

Simvastatin

The finding of dysregulation of the mevalonate kinase pathway, with cholesterol as its major end product, has led to the hypothesis that blockage of this pathway by HMG-CoA reductase inhibitors (statins) would possible be beneficial in patients with MKD. Small trials and case reports showed a statistically significant, but clinically negligible, effect in some patients. Statins have, therefore, been abandoned as treatment for MKD in clinical practice.

Other Immunosuppressive Drugs

In the past, numerous immunosuppressive drugs were tried in a trial-and-error approach to find an effective treatment for the disabling symptoms of autoinflammation. The results have been mostly disappointing, and there is no strong evidence to support the use of other immunosuppressants in autoinflammation. When other treatments are ineffective, some patients may, however, benefit from immunosuppressive drugs.

Other Treatments

It remains a matter of debate whether adenotonsillectomy effectively resolves symptoms in PFAPA. Few meta-analyses have focused on this subject; one of them found two small randomized controlled trials of unclear risk–benefit ratio.[27] A case cohort study on the use of the histamine type 2 receptor antagonist cimetidine (150 mg twice daily) in PFAPA showed that this drug is able to resolve fever episodes in 25% of patients.[6] There are no clinical trials supporting this observation.

Several case reports on the positive effects of hematopoietic stem cell transplantation (HSCT) in patients with severe MKD have been published.[28-30] Because of the severe side effects, HSCT should be considered the last resort.

AMYLOIDOSIS

Secondary or type AA amyloidosis is a serious complication of all autoinflammatory syndromes. It is caused by tissue deposition of insoluble degradation products of the inflammatory protein SAA. Kidneys are most commonly affected. Since SAA is an acute-phase reactant, there is a close relationship between the duration and level of inflammation and the development of amyloidosis. The incidence of AA amyloidosis varies among autoinflammatory diseases. Patients with FMF are at highest risk, with an incidence of up to 75% before the introduction of colchicine treatment for this disease. There is a strong correlation between ethnicity and risk of amyloidosis in FMF, with increased risk among Sephardi Jews.

Up to 25% of patients with TRAPS will develop amyloidosis if left untreated. There seems to be a strong familial predilection. In CAPS, approximately one-third of patients develop amyloidosis in the absence of treatment. Patients with MKD and Schnitzler syndrome have a relatively small risk of amyloidosis, with only a few patients with these diseases and amyloidosis known worldwide.[19] It is unclear why some patients with the same level of inflammation may develop or never develop amyloidosis. This may be related to single nucleotide polymorphisms (SNPs) in the *SAA* gene or certain genotypes.

As proteinuria is often the first sign of AA amyloidosis, patients with autoinflammation should be screened for it with regular urine sampling. Amyloidosis can be confirmed by Congo red staining of the biopsy specimen of affected tissue. This will show

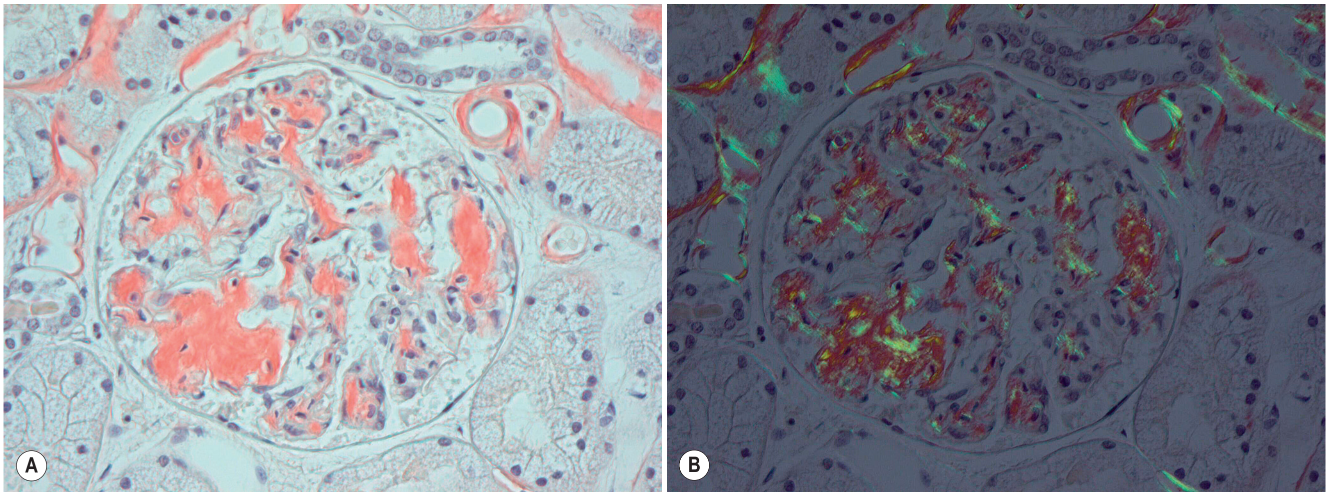


FIG 60.5 Renal Biopsy From a Patient With Amyloid a (AA) Amyloidosis. Amyloid deposits are visualized by staining with Congo red (A). Under polarized light microscopy amyloid deposits show typical apple-green birefringence (B).

an apple-green birefringence under polarized light microscopy (Fig. 60.5). Progression of amyloidosis is strongly dependent on the ability to control the underlying inflammation. If the SAA concentration can be kept under 10 mg/L, progression of amyloidosis can be halted in many cases. Some patients even show regression of amyloidosis during treatment.[31-34] A cohort mainly consisting of patients diagnosed with amyloidosis before the availability of targeted therapy showed a median survival of 19 years after the diagnosis of amyloidosis.[33]

CLINICAL PEARLS

A typical case of autoinflammation

- An 18-year-old Western European woman was seen because of recurrent episodes of fever and skin rash. These attacks started 2 hours after birth when a systemic maculopapular skin rash appeared. The rash remained present on a daily basis but was not triggered or exacerbated by cold exposure.
- At the age of 2 years, the patient developed episodes of fever, approximately 3 days each week. Typically, she had a single daily fever spike up to 39°C, which usually occurred in the evening. There was no hearing loss, and arthralgia/arthritis, dysmorphia, meningitis, myalgias, abdominal pain, and lymphadenopathy were also absent. The family history was negative for autoinflammatory diseases.
- On the basis of the clinical presentation, cryopyrin-associated periodic syndrome (CAPS) was suspected as a cause, and targeted gene analysis was performed. This showed an R260W mutation of the *NLRP3* gene. Thus this patient was diagnosed with CAPS.
- After the diagnosis was made, the patient was successfully treated with anakinra, 100 mg once daily. When canakinumab, the long-acting interleukin-1 (IL-1) inhibitor, became available, she was switched to 150 mg canakinumab once every 2 months; she remains symptom-free.

CONCLUSIONS

ON THE HORIZON

- Identification of new (hereditary) autoinflammatory syndromes in patients with episodic inflammation of unknown origin
- Wider availability and application of interleukin-1 (IL-1) inhibitors and development of new IL-1 inhibitors to improve treatment.

Since the introduction of the term *autoinflammation* at the end of the twentieth century, the clinical characteristics of the classic monogenic autoinflammatory diseases FMF, CAPS, MKD, and TRAPS have been described in more detail. Many new diseases have been identified. Research focused on their pathophysiological mechanism has revealed new genes and pathways and provided further insight into the mechanism of inflammation in general. The development of IL-1–targeting drugs has led to better quality of life for patients and to increased life expectancy in some of these diseases by prevention of complications, such as secondary amyloidosis. Nevertheless, many patients with autoinflammation remain undiagnosed.

For the near future, further study of the mechanisms of inflammation in patients with autoinflammatory diseases will likely provide a deeper insight into the workings of innate immunity and will lead to the identification of even more diseases in the autoinflammatory spectrum. The search for new therapies, preferably oral drugs, for the treatment of autoinflammation will continue.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Bodar EJ, Drenth JP, van der Meer JW, et al. Dysregulation of innate immunity: hereditary periodic fever syndromes. Br J Haematol 2009;144:279–302.
2. van der Meer JW, Vossen JM, Radl J, et al. Hyperimmunoglobulinaemia D and periodic fever: a new syndrome. Lancet 1984;1:1087–90.
3. Houten SM, van Woerden CS, Wijburg FA, et al. Carrier frequency of the V377I (1129G>A) MVK mutation, associated with Hyper-IgD and periodic fever syndrome, in the Netherlands. Eur J Hum Genet 2003;11:196–200.
4. Simon A, Mariman EC, van der Meer JW, et al. A founder effect in the hyperimmunoglobulinemia D and periodic fever syndrome. Am J Med 2003;114:148–52.
5. Marshall GS, Edwards KM, Butler J, et al. Syndrome of periodic fever, pharyngitis, and aphthous stomatitis. J Pediatr 1987;110:43–6.

6. Feder HM, Salazar JC. A clinical review of 105 patients with PFAPA (a periodic fever syndrome). Acta Paediatr 2010;99:178–84.
7. Schnitzler L. Lésions urticariennes chroniques permanentes (érythème pétaloïde ?) Cas cliniques, n° 46 B. J Dermatol d'Angers 1972.
8. Mulders-Manders CM, Simon A. Hyper-IgD syndrome/mevalonate kinase deficiency: what is new? Semin Immunopathol 2015;37:371–6.
9. Stoffels M, Szperl A, Simon A, et al. MEFV mutations affecting pyrin amino acid 577 cause autosomal dominant autoinflammatory disease. Ann Rheum Dis 2014;73:455–61.
10. Mandey SHL, Kuijk LM, Frenkel J, et al. A role for geranylgeranylation in interleukin-1 beta secretion. Arthritis Rheum 2006;54:3690–5.
11. Liao YH, Lin YC, Tsao ST, et al. HMG-CoA reductase inhibitors activate caspase-1 in human monocytes depending on ATP release and P2X7 activation. J Leukoc Biol 2013;93:289–99.
12. van der Burgh R, Pervolaraki K, Turkenburg M, et al. Unprenylated RhoA contributes to IL-1beta hypersecretion in mevalonate kinase deficiency model through stimulation of Rac1 activity. J Biol Chem 2014;289:27757–65.
13. van der Burgh R, Nijhuis L, Pervolaraki K, et al. Defects in mitochondrial clearance predispose human monocytes to interleukin-1beta hypersecretion. J Biol Chem 2014;289:5000–12.
14. Tricarico PM, Kleiner G, Valencic E, et al. Block of the mevalonate pathway triggers oxidative and inflammatory molecular mechanisms modulated by exogenous isoprenoid compounds. Int J Mol Sci 2014;15:6843–56.
15. Wurster VM, Carlucci JG, Feder HM Jr, et al. Long-term follow-up of children with periodic fever, aphthous stomatitis, pharyngitis, and cervical adenitis syndrome. J Pediatr 2011;159:958–64.
16. Cochard M, Clet J, Le L, et al. PFAPA syndrome is not a sporadic disease. Rheumatology 2010;49:1984–7.
17. Kolly L, Busso N, von Scheven-Gete A, et al. Periodic fever, aphthous stomatitis, pharyngitis, cervical adenitis syndrome is linked to dysregulated monocyte IL-1beta production. J Allergy Clin Immunol 2013;131:1635–43.
18. Sundqvist M, Wekell P, Osla V, et al. Increased intracellular oxygen radical production in neutrophils during febrile episodes of periodic fever, aphthous stomatitis, pharyngitis, and cervical adenitis syndrome. Arthritis Rheum 2013;65:2971–83.
19. de Koning HD, Bodar EJ, van der Meer JW, et al. Schnitzler syndrome: beyond the case reports: review and follow-up of 94 patients with an emphasis on prognosis and treatment. Semin Arthritis Rheum 2007;37:137–48.
20. de Koning HD, Schalkwijk J, van der Meer JW, et al. Successful canakinumab treatment identifies IL-1beta as a pivotal mediator in Schnitzler syndrome. J Allergy Clin Immunol 2011;128:1352–4.
21. de Koning HD, van Gijn ME, Stoffels M, et al. Myeloid lineage-restricted somatic mosaicism of NLRP3 mutations in patients with variant Schnitzler syndrome. J Allergy Clin Immunol 2015;135:561–4.
21a. Harrison SR, McGonagle D, Nizam S, et al. Anakinra as a diagnostic challenge and treatment option for systemic autoinflammatory disorders of undefined etiology. JCI Insight 2016;1(6):e86336.
22. van der Hilst J, Moutschen M, Messiaen P, et al. Efficacy of anti-IL-1 treatment in familial Mediterranean fever: a systematic review of the literature. Biologics 2016;10:75–80.
23. Butbul Aviel Y, Tatour S, Gershoni Baruch R, et al. Colchicine as a therapeutic option in periodic fever, aphthous stomatitis, pharyngitis, cervical adenitis (PFAPA) syndrome. Semin Arthritis Rheum 2016;45:471–4.
24. ter Haar NM, Oswald M, Jeyaratnam J, et al. Recommendations for the management of autoinflammatory diseases. Ann Rheum Dis 2015;74:1636–44.
25. Bodar EJ, Kuijk LM, Drenth JP, et al. On-demand anakinra treatment is effective in mevalonate kinase deficiency. Ann Rheum Dis 2011;70:2155–8.
26. Ter Haar N, Lachmann H, Ozen S, et al. Treatment of autoinflammatory diseases: results from the Eurofever Registry and a literature review. Ann Rheum Dis 2013;72:678–85.
27. Burton MJ, Pollard AJ, Ramsden JD. Tonsillectomy for periodic fever, aphthous stomatitis, pharyngitis and cervical adenitis syndrome (PFAPA). Cochrane Database Syst Rev 2010;(9):CD008669.
28. Arkwright PD, Abinun M, Cant AJ. Mevalonic aciduria cured by bone marrow transplantation. N Engl J Med 2007;357:1350.
29. Neven B, Valayannopoulos V, Quartier P, et al. Allogeneic bone marrow transplantation in mevalonic aciduria. New Engl J Med 2007;356:2700–3.
30. Chaudhury S, Hormaza L, Mohammad S, et al. Liver transplantation followed by allogeneic hematopoietic stem cell transplantation for atypical mevalonic aciduria. Am J Transplant 2012;12:1627–31.
31. van der Hilst JC. Recent insights into the pathogenesis of type AA amyloidosis. ScientificWorldJournal 2011;11:641–50.
32. Lane T, Gillmore JD, Wechalekar AD, et al. Therapeutic blockade of interleukin-6 by tocilizumab in the management of AA amyloidosis and chronic inflammatory disorders: a case series and review of the literature. Clin Exp Rheumatol 2015;33:S46–53.
33. Lane T, Loeffler JM, Rowczenio DM, et al. AA amyloidosis complicating the hereditary periodic fever syndromes. Arthritis Rheum 2013;65:1116–21.
34. Yilmaz S, Cinar M, Simsek I, et al. Tocilizumab in the treatment of patients with AA amyloidosis secondary to familial Mediterranean fever. Rheumatology 2015;54:564–5.

MULTIPLE-CHOICE QUESTIONS

1. Autoinflammatory diseases are characterized by:
   - **A.** Mutations in the NLRP3 inflammasome
   - **B.** Spontaneous recurrent or persistent inflammation
   - **C.** The presence of autoantibodies

2. What is the MOST common monogenic autoinflammatory disease worldwide?
   - **A.** Cryopyrin-associated periodic syndrome (CAPS)
   - **B.** Familial Mediterranean fever (FMF)
   - **C.** Mevalonate kinase deficiency (MKD)
   - **D.** Periodic fever, aphthous stomatitis, pharyngitis, and adenitis syndrome (PFAPA)
   - **E.** Schnitzler syndrome
   - **F.** Tumor necrosis factor receptor–associated periodic syndrome (TRAPS)

3. What is the treatment of first choice for FMF?
   - **A.** Acetaminophen
   - **B.** Colchicine
   - **C.** Interleukin (IL)-1 inhibition
   - **D.** Tonsillectomy

4. What is the MOST common long-term complication of the monogenic autoinflammatory diseases?
   - **A.** Amyloid A (AA) amyloidosis
   - **B.** Chronic meningitis
   - **C.** Joint contractures
   - **D.** Malignancy

61

Antiphospholipid Syndrome

Thomas L. Ortel, Doruk Erkan, Michael D. Lockshin

Diagnosis of the antiphospholipid syndrome (APS) requires that a patient has *both* a clinical event (thrombosis and/or pregnancy loss) *and* persistent antiphospholipid antibody (aPL), documented by a solid phase serum assay (anticardiolipin [aCL] or anti β_2-glycoprotein-I [β_2GPI] enzyme-linked immunosorbent assay) or an inhibitor of phospholipid-dependent clotting (lupus anticoagulant [LA] test) or both (Table 61.1).[1] Antiphospholipid syndrome occurs as an isolated diagnosis, referred to as primary APS, or is associated with other systemic rheumatic diseases such as systemic lupus erythematosus (SLE).

The primary antigen to which aPL binds is β_2GPI (apolipoprotein H), a phospholipid-binding plasma protein. β_2GPI is normally present at a concentration of 200 µg/mL and is a member of the complement control protein family. An octapeptide in the fifth domain of the protein and critical cysteine bonds are necessary for both phospholipid binding and antigenicity; a first domain site is implicated in pathogenicity. *In vivo,* β_2GPI binds, primarily as a dimer, to phosphatidylserine on activated or apoptotic cell membranes, including those of trophoblast, platelets, and endothelial cells, undergoing conformational change and becoming antigenic. This binding may initiate cell activation, clearance of apoptotic cells by macrophages, and/or coagulation.[2]

KEY CONCEPTS

- Antiphospholipid antibodies (aPL) exist as a family of autoantibodies directed against phospholipid-binding plasma proteins, most commonly β_2-glycoprotein-I.
- The origin of aPL is unknown but is hypothesized to be an incidental exposure to environmental agents inducing aPL in susceptible individuals.
- In humans, although cross-sectional and prospective cohort studies demonstrate that aPL can predict future thrombosis, the pathogenic mechanism is unknown; more than one mechanism may be involved.
- Concomitant prothrombotic risk factors may promote clotting in an additive manner in aPL-positive patients.

Autoimmune aPL binds β_2GPI (β_2GPI-dependent aPL), which in turn binds negatively charged phospholipids. Drugs (such as chlorpromazine, procainamide, quinidine, and phenytoin), malignancies (such as lymphoproliferative disorders), and infectious agents (such as syphilitic and nonsyphilitic *Treponema, Borrelia burgdorferi,* human immunodeficiency virus, *Leptospira,* or parasites) induce β_2GPI-independent transient aPL. β_2GPI-independent aPLs are usually made up of low-titer aCLs, which may bind directly to phospholipids and are rarely associated with thrombosis.

EPIDEMIOLOGY

Low-titer aCLs occur in <10% of normal blood donors, and moderate- to high-titer aCLs and/or a positive LA test occurs in <1%. The prevalence of positive aPL tests increases with age; because the differential diagnosis of vascular occlusion is broader than it is in young adults, particular care is necessary in diagnosing APS in older patients. Thirty to 40% of SLE patients and approximately one-fifth of rheumatoid arthritis patients have positive tests for aPL.

The strength of association between aPL and clinical events varies among studies. A study of healthy male physicians followed prospectively for 3 years showed that those with moderate to high titers of IgG aCL have a risk for venous thrombosis or pulmonary embolus that is eight times higher than that for men with negative tests. Although a number of studies have tried to estimate the annual thrombosis risk in asymptomatic aPL-positive patients, most of these studies have included predominantly patients with SLE. Despite the lack of well-controlled studies, in aPL-positive individuals with no other systemic autoimmune diseases or risk factors for thrombosis, the annual risk of first thrombosis is probably very low (<1%/year). However, aPL-positive patients with other systemic autoimmune diseases such as SLE are at increased annual risk for first thrombosis (<4%/year).

Approximately 10% of first-stroke victims have aPL,[3] especially those who are young, as do up to 21% of women who have suffered three or more consecutive fetal losses.[4]

ETIOPATHOGENESIS

Antiphospholipid antibodies exist as a family of autoantibodies directed against phospholipid-binding plasma proteins, most commonly β_2GPI. Such proteins bind negatively charged phospholipids (cardiolipin, phosphatidylglycerol, phosphatidylserine, phosphatidylinositol) but not zwitterionic or neutral phospholipids (phosphatidylethanolamine, phosphatidylcholine). Other phospholipid-binding plasma proteins are prothrombin, thrombomodulin, protein C, protein S, and annexins I and V. *In vivo,* the likely relevant phospholipid to which these proteins bind is phosphatidylserine, which is normally sequestered on the inner cell membrane but is exteriorized during cell activation and apoptosis. Annexins bind to exposed phosphatidylserine and create a “remove me” signal on cells.

TABLE 61.1 **Revised Sapporo Classification Criteria for the Antiphospholipid Syndrome**[1]

Clinical Criteria

1. Vascular thrombosis[a]
    - (a) One or more clinical episodes[b] of arterial, venous, or small-vessel thrombosis,[c] in any tissue or organ.
2. Pregnancy morbidity:
    - (a) One or more unexplained deaths of a morphologically normal fetus at or beyond the 10th week of gestation, *or*
    - (b) One or more premature births of a morphologically normal neonate before the 34th week of gestation because of eclampsia, severe preeclampsia, or recognized features of placental insufficiency[d] *or*
    - (c) Three or more unexplained consecutive spontaneous abortions before the 10th week of gestation, with maternal anatomical or hormonal abnormalities and paternal and maternal chromosomal causes excluded.

Laboratory Criteria[e]

1. Lupus anticoagulant present in plasma, on two or more occasions at least 12 weeks apart, detected according to the guidelines of the International Society on Thrombosis and Haemostasis.
2. Anticardiolipin antibody of IgG and/or IgM isotype in serum or plasma, present in medium or high titer (*i.e.,* >40 GPL or MPL, or > the 99th percentile), on two or more occasions, at least 12 weeks apart, measured by a standardized enzyme-linked immunosorbent assay (ELISA).
3. Anti-β_2 glycoprotein-I antibody of IgG and/or IgM isotype in serum or plasma, (in titer > the 99th percentile) present on two or more occasions, at least 12 weeks apart, measured by a standardized ELISA.

Definite antiphospholipid syndrome (APS) is present if at least one of the clinical criteria and one of the laboratory criteria are met. Classification of APS should be avoided if less than 12 weeks or more than 5 years separates the positive antiphospholipid antibody (aPL) test and the clinical manifestation. In studies of populations of patients who have more than one type of pregnancy morbidity, investigators are strongly encouraged to stratify groups of subjects according to a, b, or c above.

[a] Coexisting inherited or acquired factors for thrombosis are not a reason for excluding patients from APS trials. However, two subgroups of APS patients should be recognized, according to (a) the presence, and (b) the absence of additional risk factors for thrombosis. Indicative (but not exhaustive), such cases include age (>55 in men and >65 in women) and the presence of any of the established risk factors for cardiovascular disease (hypertension, diabetes mellitus, elevated LDL or low HDL cholesterol, cigarette smoking, family history of premature cardiovascular disease, body mass index ≥30 kg/m^2, microalbuminuria, estimated glomerular filtration rate (GFR) <60 mL/min), inherited thrombophilias, oral contraceptives, nephritic syndrome, malignancy, immobilization, and surgery. Thus patients who fulfill criteria should be stratified according to contributing causes of thrombosis.

[b] A thrombotic episode in the past could be considered as a clinical criterion, provided that thrombosis is proved by appropriate diagnostic means and that no alternative diagnosis or cause of thrombosis is found.

[c] Superficial venous thrombosis is not included in the clinical criteria.

[d] Generally accepted features of placental insufficiency include (i) abnormal or nonreassuring fetal surveillance test(s), *e.g.,* a nonreactive nonstress test, suggestive of fetal hypoxemia; (ii) abnormal Doppler flow velocimetry waveform analysis suggestive of fetal hypoxemia, *e.g.,* absent end-diastolic flow in the umbilical artery; (iii) oligohydramnios, *e.g.,* an amniotic fluid index of 5 cm or less; or (iv) a postnatal birth weight less than the 10th percentile for the gestational age.

[e] Investigators are strongly advised to classify APS patients in studies into one of the following categories: I, more than one laboratory criteria present (any combination); IIa, LA present alone; IIb, aCL antibody present alone; IIc, anti-β_2-glycoprotein-I antibody present alone.

TABLE 61.2 **Possible Mechanisms of Antiphospholipid Antibody–Induced Thrombosis**

Endothelial Cells–Antiphospholipid Antibody (aPL) Interaction

Endothelial cell damage or activation (via increased expression of adhesion molecules)
Coexisting antiendothelial antibodies
aPL-induced monocyte adhesion to endothelial cells
Increased tissue factor expression

Platelet–aPL Interaction

Platelet activation
Stimulation of thromboxane production

Coagulation System–aPL Interaction

Inhibition of activation of protein C by the thrombomodulin–thrombin complex
Inhibition of activation of protein C via its cofactor protein S
Interaction between aPL and substrates of activated protein C such as factors Va and VIIIa
Interaction between aPL and an annexin V anticoagulant shield

Complement Activation

The commonly accepted hypothesis regarding the origin of aPL states that an incidental exposure to environmental agents, the antigens of which contain β_2GPI-like peptides, induces aPL in susceptible individuals via molecular mimicry.[5] In experimental animal models, passive or active immunization with viral peptides, bacterial peptides, and heterologous β_2GPI induces polyclonal aPL and clinical events associated with APS. β_2GPI polymorphisms influence the generation of aPL in individuals but have only a weak relationship to the occurrence of aPL-related clinical events. Persons congenitally lacking β_2GPI as well as β_2GPI knock-out mice appear normal.

In humans, although cross-sectional and prospective cohort studies demonstrate that aPL can predict future thrombosis, the pathogenic mechanism remains unknown; more than one mechanism may be involved (Table 61.2). Because high-titer antibody can persist for years in asymptomatic persons and because positive aPL tests can precede symptoms for years, it is likely that vascular injury and/or endothelial cell activation will immediately precede thrombosis in persons bearing the antibody. Platelet activation followed by binding of aPL to platelet membrane phospholipid-bound annexins may initiate platelet adhesion and thrombosis. Antiphospholipid antibodies can inhibit phospholipid-dependent reactions in the coagulation cascade, such as protein C and protein S activation. Interaction between aPL and an annexin A5 anticoagulant shield is another potential mechanism.[2] Recent studies also demonstrate that aPLs induce cellular activation through receptors such as annexin A2 and apoER2, and induce release of microparticles from endothelial cells.[6]

In experimental animal models, aPLs cause fetal resorption (a proxy for recurrent fetal loss) and increase size and duration of trauma-induced venous and arterial thrombi. Inhibiting complement activation prevents experimental aPL-induced fetal death, and C5 knock-out mice carry pregnancies normally despite the presence of aPL, implying that a complement-mediated effector mechanism is a requirement for fetal death to occur.[7]

In the "second-hit hypothesis," a second trigger event (such as oral contraceptives or surgical procedures) may be necessary for an asymptomatic aPL-positive individual to develop thrombosis. Acquired and heritable risk factors for thrombosis may increase the risk of thromboembolic events in aPL patients.

A proposed pathogenesis for aPL-mediated thrombosis and placental injury begins with activation or apoptosis (by unknown triggers, possibly infectious or traumatic) of platelets, endothelial cells, or trophoblasts. Negatively charged phosphatidylserine migrates from the inner to the outer cell membrane. Circulating β_2GPI binds to phosphatidylserine, which is followed by aPL binding to a β_2GPI dimer, activating a complement and, through C5a, initiating a signaling cascade that induces cell surface tissue factor (TF) expression and adhesion molecules (*e.g.*, ICAM-1). Other components of the innate immune system may also be activated, leading to a milieu that promotes platelet activation and thrombosis. In addition, aPL adversely affects formation of a trophoblast syncytium, placental apoptosis, and trophoblast invasion, all processes required for the normal establishment of placental function. *In vitro,* pathogenic aPL induces adhesion molecules and enhances adherence of leukocytes to cultured endothelial cells.

DIAGNOSIS

Clinical Manifestations

The clinical manifestations associated with aPL represent a spectrum from asymptomatic to catastrophic APS (Table 61.3). The principal manifestations are venous or arterial thromboses and pregnancy losses. Except for their severity, the youth of affected patients, and unusual anatomical locations (Budd-Chiari syndrome, and sagittal sinus and upper extremity thromboses), thromboses in APS do not clinically differ from thromboses attributable to other causes. Stroke and transient ischemic attack are the most common presentation of arterial thrombosis, whereas deep vein thrombosis and pulmonary embolism are the most common venous manifestations of APS. Glomerular capillary endothelial cell injury or thrombosis of renal vessels (thrombotic microangiopathy) causes proteinuria without celluria or hypocomplementemia and may lead to severe hypertension and/or renal failure.

Many patients have livedo reticularis (a lattice-like pattern of superficial skin veins) (Fig. 61.1), cardiac valve disease (vegetations, valve thickening and dysfunction), or other nonthrombotic manifestations described in several studies but not included in the revised Sapporo criteria[1] due to nonspecificity or rarity (Table 61.4). These manifestations by themselves do not classify a patient as having APS for clinical studies, but they add information to the diagnosis of individual patients. The pathogenesis of cardiac valve disease in APS is unknown; valve replacement may be necessary. A putative association of aPL and increased risk of atherosclerosis may exist, but this is controversial; in studies of atherosclerosis in patients with SLE, aPL protected against atherosclerosis.[8] Some patients develop nonfocal neurological symptoms such as lack of concentration, forgetfulness, and dizzy spells. Small hyperintense lesions can be seen on magnetic resonance imaging (MRI), primarily in the periventricular white matter, but do not correlate well with clinical symptoms (Fig. 61.2).

Pregnancy losses in patients with aPL typically occur after 10 weeks' gestation (fetal loss), but early losses also occur (preembryonic or embryonic losses).[9] Patients with antiphospholipid syndrome may develop severe early preeclampsia and HELLP (*h*emolysis, *e*levated *l*iver enzymes, *l*ow *p*latelets) syndrome. Although placental infarction may be a cause of fetal growth restriction or death, nonthrombotic mechanisms of placental dysfunction are likely more important.

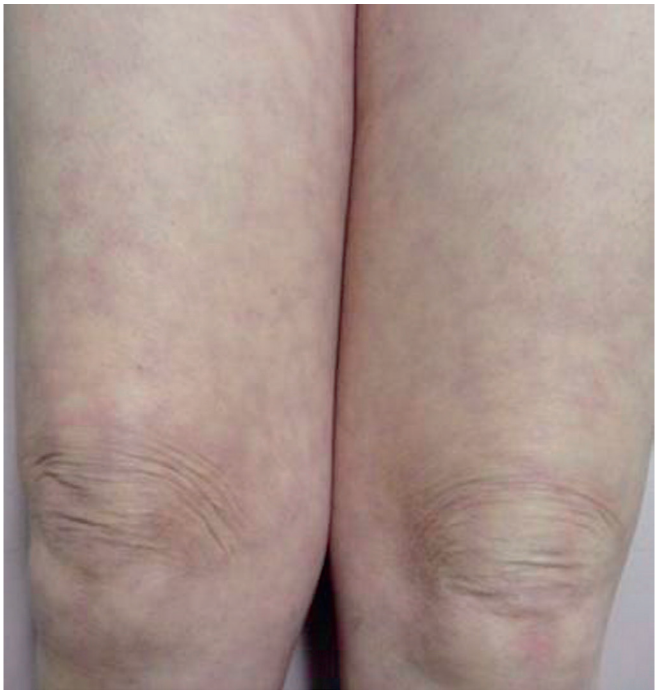


FIG 61.1 Livedo reticularis in a patient with primary antiphospholipid syndrome.

TABLE 61.3 The Clinical Spectrum of Antiphospholipid Antibodies

Asymptomatic[a] antiphospholipid antibody (aPL)–positivity
Antiphospholipid syndrome with vascular events
Antiphospholipid syndrome with only pregnancy morbidity
Asymptomatic[a] aPL-positivity with noncriteria aPL manifestations
Catastrophic antiphospholipid syndrome

[a] No history of thrombosis or pregnancy morbidity as per the Sapporo Criteria.[1]

TABLE 61.4 Noncriteria Features of the Antiphospholipid Syndrome

Type	Features
Clinical	Livedo reticularis Cardiac valve disease Autoimmune hemolytic anemia Thrombocytopenia (usually 50 000–100 000/mm^3) Multiple sclerosis–like syndrome and myelopathy Nonfocal neurological symptoms Chorea
Laboratory	IgA anticardiolipin and anti-β_2-glycoprotein-I antibodies Antiphosphatidylserine, phosphatidylinositol, phosphatidylglycerol, phosphatidylethanolamine antibodies Antiprothrombin antibodies Antiphosphatidylserine-prothrombin antibodies

Catastrophic APS (CAPS) is a rare, abrupt, life-threatening presentation. It consists of multiple thromboses of medium and small arteries occurring over a period of days, causing stroke; cardiac, hepatic, adrenal, renal, and intestinal infarction; and peripheral gangrene. Proposed diagnostic criteria for CAPS

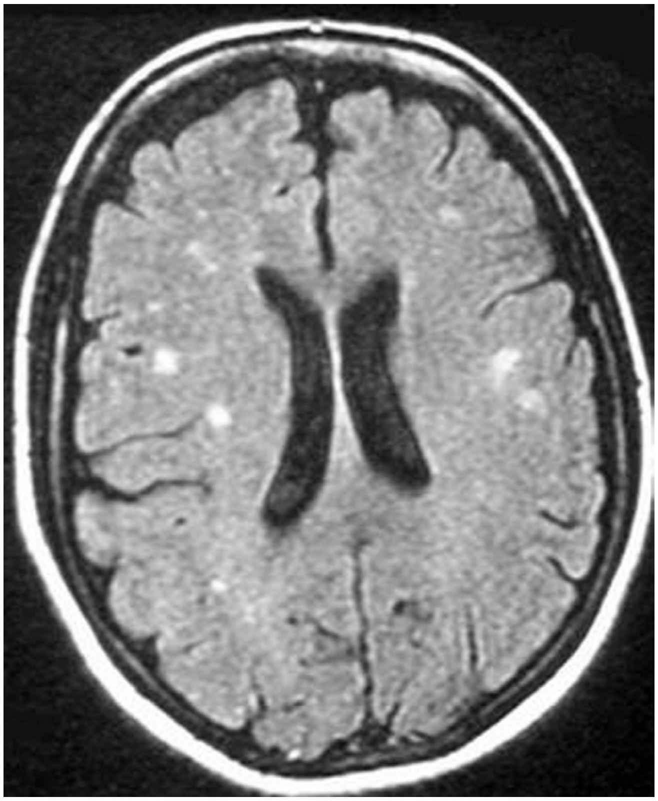


FIG 61.2 Magnetic resonance imaging demonstrating multiple periventricular white matter hyperintense lesions.

TABLE 61.5 Preliminary Criteria for the Classification of Catastrophic Antiphospholipid Syndrome

1. Evidence of involvement of three or more organs, systems, and/or tissues[a]
2. Development of manifestations simultaneously or in less than 1 week
3. Confirmation by histopathology of small-vessel occlusion in at least one organ or tissue[b]
4. Laboratory confirmation of the presence of antiphospholipid antibody (aPL)[c]

Definite Catastrophic Antiphospholipid Syndrome (APS):

- All 4 criteria

Probable Catastrophic APS:

- Criteria 2–4 and two organs, systems, and/or tissues involved;
- Criteria 1–3, except no aPL confirmation 6 weeks apart due to the early death of a patient not tested before catastrophic episode;
- Criteria 1, 2, 4; or
- Criteria 1, 3, 4, and development of a third event more than 1 week but less than 1 month after first despite anticoagulation.

[a] Usually, clinical evidence of vessel occlusions, confirmed by imaging techniques when appropriate. Renal involvement is defined by a 50% rise in serum creatinine, severe systemic hypertension, and/or proteinuria.

[b] For histopathological confirmation, significant evidence of thrombosis must be present, although vasculitis may coexist occasionally.

[c] If the patient had not been previously diagnosed as having APS, laboratory confirmation requires that the presence of aPL must be detected on two or more occasions at least 12 weeks apart (not necessarily at the time of the event), according to the proposed preliminary criteria for the classification of APS.

are shown in Table 61.5. Early diagnosis can be a challenge, especially in patients with no history of APS or aPL-positivity; diagnostic algorithms are available to guide physicians for timely diagnosis.[10] Acute adrenal failure may be the initial clinical event, heralded by unexplained back pain and vascular collapse. Patients with CAPS often have moderate thrombocytopenia; fragmented erythrocytes can be seen, but they are less frequent than in hemolytic–uremic syndrome or in thrombotic thrombocytopenic purpura. Renal failure and pulmonary hemorrhage may occur. Tissue biopsies show noninflammatory vascular occlusions involving both small- and medium-sized vessels.

Antibodies to prothrombin sometimes accompany aPLs and may cause hemorrhage by depleting prothrombin (lupus anticoagulant hypoprothrombinemia syndrome).

Laboratory Tests

In the presence of characteristic clinical events, APS is diagnosed when patients have persistent aPLs, including moderate- to high-titer IgG and/or IgM aCL, moderate- to high-titer IgG and/or IgM aβ_2GPI, and/or positive LA test.[1] Approximately 80% of patients with positive LA tests have aCL, but only 20% of patients positive for aCL have positive LA tests. Some patients are only aβ_2GPI positive. Patients with positive LA tests have higher risk for thrombosis than patients with negative tests. Several studies have shown that patients with positive aβ_2GPI tests are also at higher risk for thrombosis, but this has not been consistently observed.

LAs are detected in coagulation assays that detect the ability of aPL to interfere with phospholipid-dependent coagulation reactions. Guidelines of the International Society on Thrombosis and Haemostasis for diagnosis of LA[11] include the following:

Demonstration of a prolonged phospholipid-dependent coagulation screening test such as activated partial thromboplastin time (aPTT) or dilute Russell viper venom time (dRVVT);

Failure to correct the prolonged screening test by mixing the patient plasma with normal platelet-poor plasma, demonstrating the presence of an inhibitor;

Shortening or correcting the prolonged screening test by addition of excess phospholipid, demonstrating phospholipid dependency; and

Exclusion of other coagulopathies.

A positive screening coagulation test without confirmatory steps is not a positive LA test. Patients on anticoagulation may have false-positive or false-negative results for the LA test; thus when possible, the LA test should be ordered when the patient is not receiving anticoagulation therapy.[11]

Interpretation of positive tests should take into account the following observations: moderate- to high-titer (>40 U) aCL or aβ_2GPI is more strongly associated with clinical events than is low-titer; LA is a more specific but less sensitive predictor of thromboses than other aPL tests[12]; multiple positive aPL tests impart a worse prognosis than does any single type of test[13]; and positive aPL tests require a repeat test after 12 weeks to exclude transient aPL.[1]

Laboratory variability in these assays is moderately high: Assay stability in individual patients occurs over time in most patients. We showed that aPL results remain stable for at least three-quarters of subsequent tests during a mean follow-up of 2.4 years for the LA test, 3.5 years for the aCL test, and 1.0 year for the anti-β_2GPI test.[14] Based on same-day specimens, the consistency of aCL results among different commercial laboratories ranges from 64% to 88%, with moderate agreement for IgG and IgM, but marginal agreement for aCL IgA.[14]

Antiphospholipid antibody tests developed based on phosphatidylserine, phosphatidylinositol, phosphatidylethanolamine, or prothrombin are not well standardized or widely accepted; their clinical significance is unknown. IgA aCL can occur rarely as the only aPL in patients with APS. When positive, an IgA aCL may justify a diagnosis of APS in LA- and aCL-IgG/IgM test-negative patients with clinically typical disease. A false-positive test for syphilis is not diagnostic for APS.

Antinuclear and anti-DNA antibodies occur in approximately 45% of patients with APS who are not diagnosed as having SLE. Thrombocytopenia occurs in APS and is usually modest (>50,000/mm^3); proteinuria and renal insufficiency occur in patients with thrombotic microangiopathy. Erythrocyte sedimentation rate and hemoglobin and leukocyte count are usually normal in patients with uncomplicated APS, except during acute thrombosis. Complement levels are usually normal or only modestly low.

Imaging Studies

MRI studies show vascular occlusion and infarction consistent with clinical symptoms, without special characteristics, except that multiple otherwise unexplained cerebral infarctions in a young person suggest the syndrome. Multiple small hyperintense white matter lesions are common and do not unequivocally imply brain infarction (Fig. 61.2). Occlusions usually occur in vessels below the resolution limits of angiography; hence angiography or magnetic resonance angiography is not indicated unless clinical findings suggest medium- or large-vessel disease. Echocardiography or cardiac MRI may show severe Libman-Sacks endocarditis and intracardiac thrombi.

Pathological Studies

Skin, renal, and other tissues show noninflammatory occlusion of all caliber arteries and veins, acute and chronic endothelial injury and its sequelae, and recanalization in late lesions. The finding of inflammatory necrotizing vasculitis suggests concomitant SLE or other connective tissue disease. There are no other diagnostic immunofluorescence or electron microscopical findings.

CLINICAL PEARLS

- The clinical manifestations of antiphospholipid antibodies (aPL) represent a spectrum (from asymptomatic to catastrophic antiphospholipid syndrome [APS]); thus patients should not be evaluated and managed as having a single disease manifestation.
- Stroke and transient ischemic attack are the most common presentation of arterial thrombosis; deep vein thrombosis, often accompanied by pulmonary embolism, is the most common venous manifestation of APS.
- Pregnancy losses in patients with aPL typically occur after 10 weeks' gestation (fetal loss), but early losses also occur (preembryonic or embryonic losses).
- Catastrophic APS is a rare, abrupt, life-threatening complication of APS, which consists of multiple thromboses of medium and small arteries occurring over a period of days.
- APS diagnosis should be made in the presence of characteristic clinical manifestations and *persistently* (at least 12 weeks apart) positive aPL.

TREATMENT

Treatment recommendations for persistently aPL-positive patients are determined by the specific clinical indication (Table 61.6).

Asymptomatic Individuals

The ideal strategy for primary thrombosis prevention in asymptomatic, persistently aPL-positive individuals requires a risk-stratified approach based on aPL profile, age, systemic autoimmune diseases, traditional cardiovascular disease, or risk factors for venous thrombosis. Elimination of reversible risk factors for thrombosis (smoking, oral contraceptives) and prophylaxis during high-risk periods (surgical interventions or prolonged immobilization) is crucial for primary thrombosis prophylaxis in persistently aPL-positive individuals. The effectiveness of aspirin is not supported by the literature; in a randomized, double-blind, placebo-controlled trial, low-dose aspirin (81 mg) appeared to be no better than placebo in preventing first thrombotic episodes in persistently asymptomatic aPL-positive patients.[15] The general-population cardiovascular disease (CVD) risk prediction tools and prevention guidelines formulated based on risk–benefit calculations should play the primary role in decision making for aspirin therapy. Estrogen and estrogen-containing oral contraceptives are considered unsafe for asymptomatic women serendipitously known to bear high-titer antibody. There is no reliable information regarding the safety of progestin-only contraception, "morning after" contraception, or raloxifene, bromocriptine, or leuprolide in APS patients. However, progestin-only contraception

TABLE 61.6 Treatment Recommendations in Persistently Antiphospholipid Antibody–Positive Patients

Clinical Circumstance	Recommendation
Asymptomatic	No treatment[a]
Venous or arterial thrombosis	Warfarin international normalized ratio (INR) 2–3.0 indefinitely
Recurrent thrombosis	Warfarin INR 3–3.5 indefinitely ± low-dose aspirin
First pregnancy	No treatment[a]
Single pregnancy loss, <10 weeks	No treatment[a]
Recurrent fetal loss or loss after 10 weeks; history of no thrombosis	Prophylactic-dose[b] heparin with low-dose aspirin throughout the pregnancy, discontinue heparin 6–12 weeks postpartum
Recurrent fetal loss or loss after 10 weeks; history of thrombosis	Therapeutic-dose heparin[c] with low-dose aspirin throughout pregnancy, warfarin postpartum
Catastrophic antiphospholipid syndrome (APS)	Anticoagulation + corticosteroids + intravenous immunoglobulin or plasma exchange
Livedo reticularis	No treatment
Valve nodules or deformity	No known effective treatment; full anticoagulation if emboli or intracardiac thrombi are demonstrated
Thrombocytopenia, >50 000/mm^3	No treatment
Thrombocytopenia, ≤50 000/mm^3	Prednisone and/or intravenous immunoglobulin

[a] Aspirin 81 mg/day may be given.
[b] Prophylactic dose such as enoxaparin 30–40 mg subcutaneously (SQ) once daily.
[c] Therapeutic dose such as enoxaparin 1 mg/kg SQ twice daily or 1.5 mg/kg SQ once daily.

is theoretically safer than estrogen-based contraception. A small retrospective review of women undergoing artificial reproductive technology procedures demonstrated no thrombotic events.

Venous and Arterial Thromboembolism

Anticoagulation with unfractionated heparin or low-molecular-weight heparin (LMWH) followed by warfarin is the treatment for APS patients with vascular events. Heparin inhibits complement, a fact that makes it theoretically a preferred but impractical agent in most patients. For patients with a positive LA test that prolongs the aPTT, monitoring heparin can be accomplished by measuring antifactor Xa levels.

Two prospective controlled studies concluded that recurrence of thromboses in APS patients can be prevented with warfarin to an international normalized ratio (INR) of 2.0–3.0.[16,17] Although these studies provide strong evidence for moderate-intensity anticoagulation after an aPL-related venous event, the intensity of the anticoagulation is still debatable for APS patients with arterial events, since such patients constituted a minority in these studies. Although some APS patients may require high-intensity anticoagulation, in the absence of risk-stratified studies the definition of high risk is based currently on clinical judgment.

Most aPL-positive patients receive warfarin after ischemic strokes; however, the Antiphospholipid Antibodies and Stroke Study (APASS) concluded that for selected aPL-positive patients who have neither atrial fibrillation nor high-grade arterial stenosis, aspirin (325 mg/day) and warfarin (target INR 1.4–2.8) are equivalent in efficacy and in association with major bleeding complications.[18] The generalizability of these results is limited, as the study group had an average age of 60 years (higher than the average for APS populations), the aPL determination was performed only once at study entry, and the cutoff for assigning a patient to the positive aCL group was very low. However, based on APASS results, aspirin is an option for older patients with a single low positive aCL test who present with stroke.

Recently, four direct oral anticoagulants (DOACs) that target thrombin (dabigatran) or factor Xa (rivaroxaban, apixaban, and edoxaban) have been approved for the treatment of patients with venous thromboembolism. Use in patients with APS is limited, and cases of therapeutic failure have been reported. A prospective study investigated the use of rivaroxaban compared with standard-intensity warfarin in patients with APS, using the change in endogenous thrombin potential (ETP), a laboratory parameter, as the primary outcome. Although the ETP did not reach the noninferiority threshold, there was no increase in thrombotic events in patients taking rivaroxaban.[19] The study was not powered for clinical outcomes, however, and other studies are in progress.

Venous thrombosis in aPL-positive patients typically has a high recurrence rate if anticoagulation is discontinued, and lifelong anticoagulation is usually recommended. A recent systematic review of the literature revealed that the available evidence in support of an association between the presence of aPL and risk of recurrence is of low quality, however, suggesting that indefinite anticoagulation may not be needed by all patients.[20] For example, it is unknown whether patients whose event was triggered by an acquired, reversible risk factor for thrombosis can discontinue anticoagulation or switch to aspirin when the trigger factor is eliminated. Normalization of the LA or aCL tests is *not* an indication to discontinue anticoagulation. For well-anticoagulated patients who continue to have thromboses, antiplatelet drugs, hydroxychloroquine, statins, intravenous immunoglobulin (IVIG), and plasmapheresis have theoretical bases for efficacy and have all been used. There are no systematic studies of treatment for CAPS. Detailed reviews conclude that the most effective treatment combines full-dose anticoagulation, high-dose corticosteroids, plasma exchange, and IVIG.

Pregnancy Morbidity

Pregnancy is a prothrombotic state; management strategies in persistently aPL-positive patients should focus on prevention of both pregnancy morbidity and maternal thrombotic complications.[9] In pregnancy, heparin and low-dose aspirin combination increases fetal survival rate from 50% to 80% among women who have had a fetal loss and tests positive for aPL. If patients fail this regimen, the next step is to add IVIG, an approach not supported by controlled studies. Most experts in the field use LMWH (*e.g.,* enoxaparin) due to lower risk of thrombocytopenia and osteoporosis—prophylactic doses for women who have had only pregnancy morbidity, or full anticoagulant doses for women who have had prior thromboses (Table 61.6). Treatment begins after confirmation of pregnancy, continues until 48 hours before anticipated delivery (to allow epidural anesthesia), and resumes for 8–12 weeks postpartum (if no prior thromboembolism), or else the patient is transitioned to warfarin for continued therapy. No studies unequivocally justify treatment of women with aPL during a first pregnancy, women with only very early losses, or women whose aPL titers are low or transient. Nonetheless, it is common to offer such patients low-dose aspirin.

Other Clinical Manifestations of APS

There is no consensus for the treatment of patients with noncriteria and/or nonthrombotic manifestations of aPL. Corticosteroids and/or IVIG are the first-line treatments for platelet counts less than 50 000/mm^3. An open-label phase IIa pilot study of aPL found that rituximab may be effective in controlling some but not all noncriteria manifestations of APS, although aPL profiles did not change with treatment.[21]

Perioperative Management

Serious perioperative complications may occur despite prophylaxis. Patients with APS are at additional risk for thrombosis when they undergo surgery. Thus perioperative strategies should be clearly identified before any surgical procedure or before pharmacological and physical antithrombosis interventions are vigorously employed; in addition, periods without anticoagulation must be kept to an absolute minimum, and any deviation from a normal course must be considered a potential disease-related event.[22]

Additional Therapeutic Considerations

There is experimental and clinical evidence in lupus patients that hydroxychloroquine (HCQ) might decrease the incidence of thrombosis, and recent *in vitro* studies have demonstrated that HCQ might protect endothelial cells and syncytialized trophoblast cell lines from the disruptive effect of antiphospholipid antibodies.[23] In patients with systemic autoimmune diseases (particularly lupus), HCQ is commonly employed for disease control and should be considered independent of patients' aPL status. However, further controlled studies are needed to determine the effectiveness of HCQ for primary prophylaxis in aPL-positive patients.

Although statins have been used in primary and secondary cardiovascular disease prevention, no data exist for their use in thrombosis prevention in aPL-positive patients. Experimental evidence in APS mouse models and the clinical trials demonstrating rosuvastatin's protective effect against a first major cardiovascular event and venous thrombosis in a general population without hyperlipidemia but with elevated high-sensitivity C-reactive protein levels[24] justifies clinical studies of statins in nonpregnant aPL-positive patients.

Antiplatelet agents other than aspirin or long-term LMWH treatment, empirically used, have not been formally tested in clinical trials. Ongoing prospective, randomized clinical trials investigating the role of the direct oral anticoagulants in the treatment of thromboembolism in patients with APS are eagerly awaited, and other strategies, including complement inhibitors and novel antithrombotic therapies, may be candidates for future clinical trials.

CONCLUSIONS AND TRANSLATIONAL RESEARCH

ON THE HORIZON

Translational Research in Antiphospholipid Syndrome

- A better understanding of the cellular mechanisms of antiphospholipid antibody (aPL) mediated clinical events that will help us design more specifically targeted treatments.
- Identification of patients who are at risk for future aPL-related events.
- Controlled studies of theoretically useful medications such as hydroxychloroquine, statins, complement inhibitors, or anti-B-cell therapies.

APS is a systemic autoimmune disease consisting of thromboses, pregnancy losses, and persistent high-titer aPL. Inflammation and complement activation are established mechanisms of aPL-related manifestations in murine models; however, definitive studies in humans do not exist. The disease is too variable clinically, and its mechanisms too diverse, to expect that a single mechanism defined in a single model will apply to all aspects of this disease. Given that the mechanisms of aPL-induced thrombosis are not well understood and that thrombosis is multifactorial and that controversies exist about the strength of association between aPL and clinical events, drug development specific for aPL-positive patients has been challenging. Anticoagulation is the primary treatment today, but a future therapeutic approach will likely include immunomodulatory agents.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Miyakis S, Lockshin MD, Atsumi T, et al. International consensus statement on an update of the classification criteria for definite antiphospholipid syndrome (APS). J Thromb Haemost 2006;4:295–306.
2. Giannakopoulos B, Krilis SA. The pathogenesis of the antiphospholipid syndrome. N Engl J Med 2013;368:1033–44.
3. The Antiphospholipid Antibodies in Stroke Study (APASS) Group. Anticardiolipin antibodies are an independent risk factor for first ischemic stroke. Neurology 1993;43:2069–73.
4. Stephenson MD. Frequency of factors associated with habitual abortion in 197 couples. Fertil Steril 1996;66:24–9.
5. Erkan D, Lockshin MD. What is antiphospholipid syndrome? Curr Rheumatol Rep 2004;6:451–7.
6. Betapudi V, Lominadze G, Hsi L, et al. Anti-β2GPI antibodies stimulate endothelial cell microparticle release via a nonmuscle myosin II motor protein-dependent pathway. Blood 2013;122:3808–17.
7. Girardi G, Berman J, Redecha P, et al. Complement C5a receptors and neutrophils mediate fetal injury in the antiphospholipid syndrome. J Clin Invest 2003;112:1644–54.
8. Roman MJ, Shanker BA, Davis A, et al. Prevalence and correlates of accelerated atherosclerosis in systemic lupus erythematosus. N Engl J Med 2003;349:2399–406.
9. Ruiz-Irastorza G, Crowther M, Branch W, et al. Antiphospholipid syndrome. Lancet 2010;376:1498–509.
10. Erkan D, Espinosa G, Cervera R. Catastrophic antiphospholipid syndrome: updated diagnostic algorithms. Autoimmun Rev 2010;10:74–9.
11. Pengo V, Tripodi A, Reber G, et al. Update of the guidelines for lupus anticoagulant detection. Subcommittee on Lupus Anticoagulant/Antiphospholipid Antibody of the Scientific and Standardisation Committee of the International Society on Thrombosis and Haemostasis. J Thromb Haemost 2009;7:1737–40.
12. Galli M, Luciani D, Bertolini G, et al. Lupus anticoagulants are stronger risk factors for thrombosis than anticardiolipin antibodies in the antiphospholipid syndrome: a systematic review of the literature. Blood 2003;101:1827–32.
13. Pengo V, Ruffatti A, Legnani C, et al. Incidence of a first thromboembolic event in asymptomatic carriers of high-risk antiphospholipid antibody profile: a multicenter prospective study. Blood 2011;118:4714–18.
14. Erkan D, Derksen WJ, Kaplan V, et al. Real world experience with antiphospholipid antibody tests: how stable are results over time? Ann Rheum Dis 2005;64:1321–5.
15. Erkan D, Harrison MJ, Levy R, et al. Aspirin for primary thrombosis prevention in the antiphospholipid syndrome: a randomized, double-blind, placebo-controlled trial in asymptomatic antiphospholipid antibody-positive individuals. Arthritis Rheum 2007;56:2382–91.
16. Crowther MA, Ginsberg JS, Julian J, et al. A comparison of two intensities of warfarin for the prevention of recurrent thrombosis in patients with the antiphospholipid antibody syndrome. N Engl J Med 2003;349:1133–8.
17. Finazzi G, Marchioli R, Brancaccio V, et al. A randomized clinical trial of high-intensity warfarin vs. conventional antithrombotic therapy for the prevention of recurrent thrombosis in patients with the antiphospholipid syndrome (WAPS). J Thromb Haemost 2005;3:848–53.
18. Levine SR, Brey RL, Tilley BC, et al. Antiphospholipid antibodies and subsequent thrombo-occlusive events in patients with ischemic stroke. JAMA 2004;291:576–84.
19. Cohen H, Hunt BJ, Efthymiou M, et al. Rivaroxaban versus warfarin to treat patients with thrombotic antiphospholipid syndrome, with or without systemic lupus erythematosus (RAPS): a randomised, controlled, open-label, phase 2/3, non-inferiority trial. Lancet Haematol 2016;3:e426–36.
20. Garcia D, Akl EA, Carr R, et al. Antiphospholipid antibodies and the risk of recurrence after a first episode of venous thromboembolism: a systematic review. Blood 2013;122:817–24.
21. Erkan D, Vega J, Ramon G, et al. A pilot open-label phase II trial of rituximab for non-criteria manifestations of antiphospholipid syndrome. Arthritis Rheum 2013;65:464–71.
22. Erkan D, Leibowitz E, Berman J, et al. Perioperative medical management of antiphospholipid syndrome: hospital for special surgery experience, review of literature, and recommendations. J Rheumatol 2002;29:843–9.
23. Rand JH, Wu XX, Quinn AS, et al. Hydroxychloroquine protects the annexin A5 anticoagulant shield from disruption by antiphospholipid antibodies: evidence for a novel effect for an old antimalarial drug. Blood 2010;115:2292–9.
24. Glynn RJ, Danielson E, Fonseca FA, et al. A randomized trial of rosuvastatin in the prevention of venous thromboembolism. N Engl J Med 2009;360:1851–61.

MULTIPLE-CHOICE QUESTIONS

1. Which of the following statements concerning treatment options for a patient with antiphospholipid antibodies is correct?
   - **A.** Aspirin is an effective thromboprophylactic strategy for asymptomatic individuals with antiphospholipid antibodies.
   - **B.** Patients with arterial thromboembolism and APS should be treated with aspirin and a direct oral anticoagulant
   - **C.** Patients with venous thromboembolism and APS should be treated with warfarin with a target INR of 2-3
   - **D.** A woman with prior pregnancy loss and APS should receive intravenous immunoglobulin and steroids during subsequent pregnancies

2. A lupus anticoagulant is identified by which of the following laboratory test results?
   - **A.** Elevated antibody levels to prothrombin
   - **B.** An antibody that interferes with phospholipid-dependent coagulation reactions
   - **C.** An antibody that induces an increased thrombin potential
   - **D.** The presence of a specific factor inhibitor

3. Which of the following patients should be tested for possible APS?
   - **A.** A 65 year old man with a stroke
   - **B.** A 22 year old woman taking oral contraceptives with a pulmonary embolism
   - **C.** A 34 year old woman who presents with fetal demise at 34 weeks gestation
   - **D.** A 27 year old man new-onset seizures

62

Immunohematological Disorders

Jay N. Lozier, Pierre Noel

Congenital (primary) or acquired (secondary) immunodeficiencies, medications, and lymphoproliferative and rheumatological disorders are frequently associated with immune-mediated cytopenias. These processes can affect erythrocytes, leukocytes, and platelets (individually or in combination). In this chapter, we address immune-mediated cytopenias of each hematological component, discussing pathophysiology, clinical presentation, differential diagnosis, and treatment options.

IMMUNE-MEDIATED HEMOLYTIC ANEMIA

Immune-mediated hemolysis can be autoimmune, alloimmune, idiopathic, or secondary to drugs or other diseases. Regardless of the underlying cause, immunoglobulin G (IgG) or IgM (rarely IgA) antibodies are directed against antigens on the red blood cell (RBC) membrane (Table 62.1).[1,2] These disorders can be categorized on the basis of the underlying cause and the type of antierythrocyte antibody that mediates the process.

Autoimmune Hemolysis Mediated by Warm Antibody

Although warm-antibody autoimmune hemolytic anemia is rare, it increases in prevalence above the age of 50 years and tends to be more common in women than in men, as is typical of most autoimmune diseases. In this condition, IgG autoantibodies and complement components are present on the RBC surface (see below). Although the autoantibody target is most commonly the Rh antigen, a variety of other targets are known. The specificity of the antibody, however, does not affect the clinical presentation or management of hemolysis.

Drug-Induced Immune Hemolysis

Over 125 drugs have been associated with immune hemolysis.[3] Cefotetan, ceftriaxone, and piperacillin currently account for over 80% of cases. The prognosis of drug-related immune hemolysis is much better than that of idiopathic hemolysis because the hemolysis stops once the offending drug has been removed. The drugs most commonly associated with fatal hemolytic anemia are cefotetan (8%) and ceftriaxone (6%).[3] Fludarabine and cladribine have been associated with immune hemolysis. Following multiple courses of alkylating agents, fludarabine causes autoimmune hemolytic anemia in 20% of patients with chronic lymphocytic leukemia (CLL). The combination of fludarabine and cyclophosphamide with or without rituximab (FC, FCR) may protect against the development of autoimmune hemolytic anemia.[3,4]

Although the biochemical mechanisms of drug-related immune hemolysis are not completely clear, several hypotheses are generally accepted. Most commonly, complexes of drug and IgG and/or IgM that adsorb to the RBC surface and fix complement. The resultant intravascular hemolysis is acute and often severe enough to cause renal failure from toxicities of hemoglobin to renal epithelium.

A second less common mechanism develops primarily in patients receiving very high doses of penicillin (rarely used) for at least 1 week. High-titer antipenicillin IgG develops and binds to penicillin that is covalently attached to the RBC membranes. The resultant hemolysis is less acute than that caused by immune complexes but can be life-threatening.

In a third mechanism, a drug stimulates the production of an antibody that reacts with the patient's RBCs independently from the drug. Serologically, this antibody is indistinguishable from an idiopathic autoantibody. This has become rare as the use of the primary causal agent, methyldopa (an antihypertensive medication), has declined. Although these autoantibodies commonly cause positive clinical antibody tests (see below), they rarely cause hemolysis *in vivo,* and when they do, it usually ceases within 2 weeks of discontinuing the drug.

Cold Agglutinin Diseases

This type of hemolytic anemia is mediated predominantly by anti-I or anti-i IgM that agglutinates red cells at temperatures well below 37°C.[5] These antibodies engage the complement pathway resulting in C3b-mediated RBC phagocytosis mainly by Kupffer cells, whereas membrane-associated complex is a minor mechanism at low IgM titers. The severity of the clinical illness depends on the concentration of the IgM and its "thermal amplitude." Thermal amplitude describes the temperature range over which it binds to RBCs: for example, antibodies that exclusively bind at 4°C are only active *in vitro,* whereas those that bind at >30°C can bind to RBCs as they circulate in the periphery and begin the process of complement fixation, which can persist even as the cells return to body core temperatures. The activity of the IgM is also determined by its relative affinity for the I- and i-antigens, which varies from one individual to the next.

Two general types of cold agglutinin disease are recognized: a chronic idiopathic disease presenting in patients over age 50 years, caused by monoclonal anti-I IgM, and a transient disease secondary to certain infections (*e.g.,* mycoplasma, Epstein-Barr virus [EBV], cytomegalovirus [CMV]), caused by polyclonal anti-i and anti-I (see Table 62.1). Avoidance of cold environments is important in both categories of cold agglutinin disease. In addition, cold agglutinin disease can be associated with B-cell lymphoproliferative disorders, and this typically is responsive to rituximab and/or rituximab combined with fludarabine.[6]

Paroxysmal Cold Hemoglobinuria

Paroxysmal cold hemoglobinuria (PCH) is caused by anti-P IgG that is very effective in fixing complement and producing intravascular hemolysis.[2] Although rare, it is most common in

TABLE 62.1 Classification of Immune Hemolytic Disorders

Autoimmune

Warm Antibody-Mediated

Idiopathic
Secondary
Drugs, lymphoid malignancies, infections
Other autoimmune diseases

Cold Antibody-Mediated

Cold agglutinin disease
Idiopathic
Secondary
Infection, lymphoid malignancies

Paroxysmal Cold Hemoglobinuria

Idiopathic
Secondary to infections

Alloimmune

Secondary to red cell transfusions (alloantibodies, isoantibodies)
Secondary to fetal-maternal hemorrhage
Secondary to transplanted lymphocytes

children following a viral illness and can be managed by avoidance of cold. In the past, it was more commonly associated with syphilis (see Table 62.1). There is also an autoimmune variety of PCH that may require immunosuppression with corticosteroids. Splenectomy is not helpful as a consequence of the fact that the hemolysis is intravascular.

Hemolytic Transfusion Reactions

Because individuals with group O RBCs have preformed iso-anti-A and -B, they must only be transfused with group O cells. Similarly, individuals with group A RBCs with preformed anti-B isoantibodies must only receive group A RBCs, and individuals with group B RBCs must only receive group B RBCs. Because of the absence of either group A or B antigens on the surface of group O RBCs, such cells can be used in transfusion of A, B, or AB individuals in emergency situations. Failure to abide by these rules results in acute intravascular hemolysis that can cause renal failure, disseminated intravascular coagulation, and death.

Other hemolytic transfusion reactions are caused by alloantibodies, predominantly IgG. Therefore a multiply transfused patient is at risk for hemolysis from alloantibodies if the patient receives incompatible blood. Multiparous women are at similar risk because of exposure to paternal fetal RBC antigens. Fortunately, for unclear reasons, most RBC antigens do not elicit an immune response although the Rh, Kell, Kidd, and Duffy antigens are clearly immunogenic. Exposure particularly to Kidd or Duffy antigens may stimulate alloantibody formation that can rise to sufficient titer to cause hemolysis with an onset typically delayed by approximately 1 week. Such delayed transfusion reactions may be subtle or cause an abrupt drop in hemoglobin with jaundice and hemoglobinuria.

Immune Hemolysis Associated With Transplantation

Any transplanted tissue may contain “passenger” lymphocytes from the donor that will survive and proliferate in the recipient if the recipient is sufficiently immunosuppressed.[7] When the RBCs of the recipient carry A or B antigen and the donor is ABO incompatible, the transplanted lymphocytes will respond to the recipient RBCs as foreign, and allogeneic anti-A or anti-B antibodies will be produced and can lead to significant hemolysis. If the transplantation involves hematopoietic stem cells (HSCs), this dilemma will resolve once the donor’s erythropoiesis prevails and the donor lymphocytes are no longer exposed to the recipient RBCs.

Immunopathogenesis

The antigens for antierythrocyte IgG are usually proteins, the most clinically important of which are the Rh-associated glycoproteins (RhAG), D, C, c, E, and e.[8] In contrast, antierythrocyte IgM is directed at polysaccharides, which include the ABO and I-antigens (I, i) found on the anion and glucose transporter proteins in the RBC membrane.[9,10]

Antibodies are distinguished by being “warm” and “cold” reactive, respectively, meaning that they bind antigens at core body temperature (warm) or they bind antigens preferentially at lower temperatures (cold) in the peripheral circulation or *ex vivo*. This distinction results from the different thermodynamics of binding to protein (hydrophobic) and polysaccharide (electrostatic) antigens.[10]

IgG antibodies typically bind at warm temperatures and IgM antibodies typically bind at cold temperatures, although there can be overlap and exceptions (*e.g.,* the Donath-Landsteiner IgG antibodies that are seen in paroxysmal cold hemoglobinuria). IgG and IgM also differ in their ability to fix complement, and this affects the resulting mechanism of hemolysis. To attach the first component of the classical complement pathway, two IgG molecules must bind in close proximity on the RBC. However, because of its pentameric structure, a single IgM molecule can initiate complement activation.

Erythrocyte-bound IgG becomes attached to the Fc receptors of splenic macrophages, which may engulf all or part of the cell or release lysosomal enzymes that digest its membrane (antibody-dependent cell-mediated cytotoxicity [ADCC]).[11] RBC fragments escaping from this encounter lose more membrane than cytoplasm and become spherical (spherocytes) as a consequence of this change in the surface-to-volume ratio. If IgG has initiated complement activation on the cell surface, binding of C3b to splenic macrophages will augment erythrocyte phagocytosis in the spleen.[12]

When IgM fixes complement, the process begins in the cooler peripheral circulation, where IgM binds to RBCs. If the amount of IgM bound is relatively high with at least some of it remaining on the cell at 37°C (*e.g.,* anti-A or anti-B isoantibodies), the cascade of complement activation goes to completion. Doughnut-shaped holes are formed in the cell membrane that allow the influx of water and sodium, inducing intravascular osmotic rupture of the cell.[11] However, if the IgM elutes from the RBC as it returns to body core temperature, the complement reaction attenuates. In this circumstance, the components remain on the cell but do not cause intravascular hemolysis. Instead, the cells are cleared by hepatic macrophages via complement-binding sites.[13]

This has important implications for clinical tests for the presence of antibody and complement on the RBC surface during the workup of immune-mediated hemolysis. Antibody-mediated hemolysis causes variable degrees of anemia and reticulocytosis. Intravascular hemolysis releases hemoglobin into the circulation, but this is too small an amount to cause measurable hemoglobinemia, although it will result in consumption of haptoglobin, which is rapidly depleted. In contrast, when hemolysis results from complement-mediated lysis, such as follows an ABO-incompatible blood transfusion, hemoglobinemia becomes massive, overcoming the scavenging capacity of plasma hemoglobin binders (haptoglobin, hemopexin, albumin), resulting in hemoglobinuria. Because hemoglobin is toxic to the renal tubular epithelium, renal function may become impaired. RBC membrane fragments released by massive intravascular hemolysis are a rich source of procoagulant phosphatidylserine and can precipitate disseminated intravascular coagulation.

In contrast, the consequences of extravascular hemolysis (*i.e.,* via phagocytosis in the reticuloendothelial system of the liver and the spleen) are much less severe. In macrophages, iron is removed from the hemoglobin and recycled to the circulation to support a compensatory reticulocytosis and the heme porphyrin is metabolized to bilirubin.

Patients with immune-mediated anemia have an increased incidence of venous thromboembolism, and detection of a lupus anticoagulant in these patients places them at a particularly high risk for this complication.[14]

Diagnosis

With few exceptions, if the mechanism of hemolysis is immune mediated, an anti-RBC antibody can be demonstrated, either on the RBC surface, in serum, or both.[1,2] With autoimmune hemolysis, IgG or IgM and/or complement components can be identified by a direct antibody test (DAT), originally known as a *direct Coombs test* (Fig. 62.1). For this assay, a patient's RBCs are washed and suspended in buffer. Surface-bound IgG is detected by adding anti-IgG antibody, which, being divalent, can bind to IgG on adjacent RBCs and agglutinate them into visible aggregates. Because of its pentameric structure, IgM on the cells can cause agglutination without the addition of a second antibody. Even when IgM has been previously eluted from the cell surface as a result of warming in the central circulation, its earlier presence *in vivo* can be detected by telltale remnants of complement that are fixed to the RBC. In this setting, detection requires the addition of anticomplement (*e.g.,* anti-C3dg) antibody.

Alloantibodies can also be detected by the DAT if allogeneic RBCs from a previous transfusion are still circulating. If these have been cleared, however, RBC antibodies can be identified in the patient's serum by adding the serum to a panel of RBCs carrying different antigens. Agglutination is detected as described above; this constitutes the indirect antibody test.

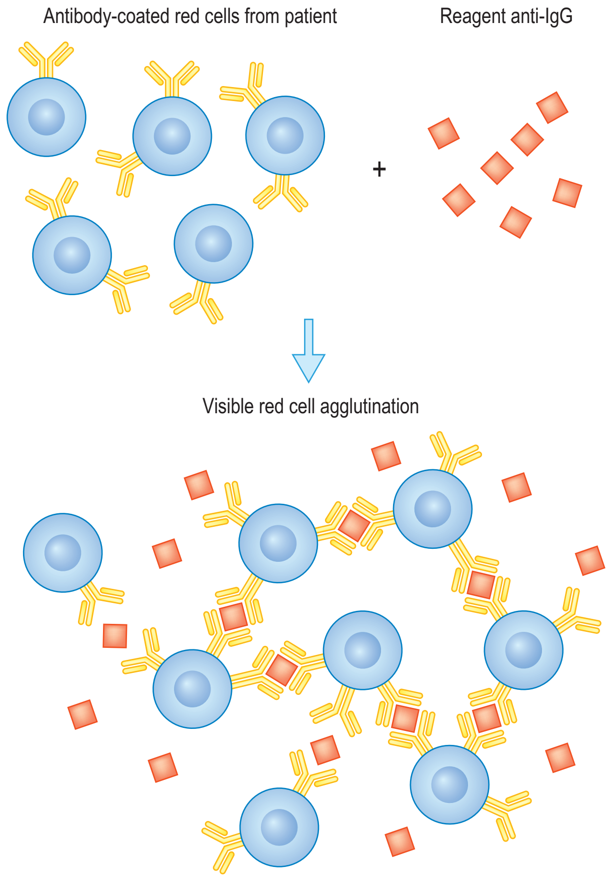

Antibody-coated red cells from patient
Reagent anti-IgG
+
Visible red cell agglutination

FIG 62.1 **The Direct Antibody Test (DAT).** The test is positive when immunoglobulin G (IgG: light blue triangles)–coated red blood cells are cross-linked by anti-IgG antibody (dark blue triangles) to form visible cell aggregates. Cell-bound complement and/or IgM can be detected by using anticomplement or anti-IgM reagent antibodies.

CLINICAL PEARLS

In the workup of hemolytic anemia, the following clinical laboratory studies provide important clues as to mechanism and may lead to a diagnosis of an immune hemolytic anemia.

Direct antibody test (DAT): The presence of antibodies on the surface of red blood cells (RBCs) suggests an immune-mediated hemolysis. The presence of antibody and complement on the surface of the RBC suggests drug-related hemolysis, whereas the presence of complement alone may suggest an immunoglobulin M (IgM) or cold antibody-related hemolysis.

Peripheral blood smear: Examination of the peripheral smear and various RBC indices (chiefly the mean corpuscular volume [MCV]) give mechanistic clues to the etiology of the hemolytic process. Immune-mediated hemolysis is characterized by spherocytosis, even microspherocytosis in severe cases, as antibody-coated RBCs traversing the reticuloendothelial system assume a spherical form, rather than that of a normal biconcave disc. The appearance of other pathological forms, such as schistocytes, sickle cells, targets, or tear drop forms (dacrocytes), suggest other causes of hemolysis, such as thrombotic thrombocytopenic purpura (TTP), or mechanical, shear-induced hemolysis (*e.g.,* aortic stenosis), sickle cell disease, thalassemia, or extramedullary hematopoiesis from bone marrow fibrosis, metastasis, or failure. Nucleated RBCs can be seen in any form of hemolytic anemia if it is severe enough.

Reticulocytes: Evaluation of the reticulocyte count indicates whether bone marrow is capable of making new erythrocytes in response to hemolysis.

Lactic acid dehydrogenase (LDH): The LDH is typically elevated with ongoing hemolysis, since LDH is an important housekeeping enzyme found in erythrocytes. LDH is found in cells from all tissues, each having a characteristic isoenzyme form of LDH. It is rarely necessary to distinguish LDH from RBCs from other tissue sources, but LDH isoenzyme 1 is the predominant form found in RBCs.

Bilirubin: As heme is released from RBCs, it is metabolized to bilirubin, which is glycosylated and then excreted via hepatic metabolism. Initially, as large amounts of heme are released and metabolized, the bilirubin is predominantly indirect bilirubin (unconjugated), and then it is converted to direct (conjugated) bilirubin. This can be altered by cholestasis, either from biliary obstruction or hepatic disease or Gilbert disease, or hepatic immaturity in the premature infant or newborn.

Haptoglobin: Haptoglobin is extremely sensitive to even small amounts of hemolysis. Its absence merely confirms that there is a significant hemolysis, but not its extent. The presence of normal haptoglobin effectively rules out significant hemolysis, and the return of measurable amounts of haptoglobin usually signals the end of hemolysis.

Therapy

The first line of therapy is corticosteroids, and 80% of patients achieve a partial or complete response to 1 mg/kg/day of prednisone (orally). Once a response is achieved, the prednisone dose is tapered slowly. Approximately 50% of patients require prednisone at a dose of 15 mg/day or less to maintain the hemoglobin level >10 g/dL, and it may take up to 3 weeks for patients to achieve a response. Patients who do not respond in 3 weeks should be started on second-line therapy. It is estimated that long-term complete responses not requiring prednisone can be achieved in 20% of patients.[15]

THERAPEUTIC PRINCIPLES

Autoimmune Anemia

- Hemolytic anemia induced by "warm" antibody
  - Acute: corticosteroids, transfusion if severe
  - Chronic: splenectomy, rituximab, immunosuppression
- Hemolytic anemia induced by "cold" antibody
  - Cold agglutinin disease: avoidance of cold; rituximab +/– fludarabine
  - Paroxysmal cold hemoglobinuria: avoidance of cold; corticosteroids
- Hemolytic disease of the newborn
  - Neonatal exposure to fluorescent light
  - Intrauterine transfusion or exchange transfusion

Splenectomy and anti-CD20 antibody (rituximab) are considered second-line therapy. Splenectomy is associated with short-term partial or complete responses in two-thirds of patients. The overall response rate to rituximab is approximately 80%, but rituximab is contraindicated in patients with untreated hepatitis B. The rare, but most severe, long-term complication of rituximab therapy is progressive multifocal leukoencephalopathy.[15]

Danazol, a synthetic anabolic steroid, has been used as a first-line agent in conjunction with prednisone; its effectiveness appears to be less in relapsed or refractory disease. The role of high-dose intravenous immunoglobulin (IVIG) remains controversial; its effectiveness remains to be determined in larger trials. Third-line therapy consists of immunosuppressive agents (*e.g.,* azathioprine, cyclophosphamide, alemtuzumab, mycophenolate mofetil, cyclosporine).[1,15]

IMMUNE-MEDIATED NEUTROPENIA

Immune neutropenia constitutes a heterogeneous group of acquired diseases in which the immune system responds to circulating neutrophils, selectively reducing their level to below 1500 cells/mm^3 (Table 62.2).

KEY CONCEPTS

Immune Neutropenia

- Immune neutropenia in children is caused by antineutrophil antibodies (ANAs).
- Immune neutropenia in adults has a more complex etiology.
- Immune complex-mediated neutrophil clearance and cell-mediated suppression of myelopoiesis often play a major role.

Isoimmune Neonatal Neutropenia

Transient neutropenia analogous to neonatal immune hemolysis or thrombocytopenia develops when IgG antineutrophil antibodies (ANAs) from an allosensitized or autoimmune mother cross the placenta and destroy fetal neutrophils.[16] Neutropenia is typically present at birth and resolves spontaneously, as maternal antibodies usually disappear within 12–15 weeks, but occasionally it can persist as long as 24 weeks after delivery.

TABLE 62.2 **Causes of Immune Neutropenia**

Primary
Isoimmune neonatal neutropenia
Autoimmune neutropenia of childhood
Adult autoimmune neutropenia
Secondary
Systemic autoimmune disease
Rheumatoid arthritis (*i.e.,* Felty syndrome)
Systemic lupus erythematosus
Sjögren syndrome
Lymphoproliferative malignancy
Large granular lymphocyte (LGL) leukemia
Lymphoma
Drug-Induced
Antiplatelets—ticlopidine
Inflammatory bowel disease drug—sulfasalazine
Antipsychotic—clozapine, phenothiazines
Antithyroid medications—propylthiouracil, methimazole
Retrovirals
Antibiotics—beta-lactams, cefepime, trimethoprim-sulfamethoxazole, vancomycin, rifampicin, quinine/quinidine
Diuretics—furosemide, spironolactone
Antiepileptic—lamotrigine
Rituximab, infliximab, etanercept

Primary Autoimmune Neutropenia

Primary autoimmune neutropenia is an antibody-mediated disease that presents commonly in early childhood.[17] Patients typically have normal blood counts at birth and develop neutropenia at 3–36 months of age. Children presenting at <2 years of age most commonly recover spontaneously within 2–3 years of diagnosis. Nevertheless, some children, particularly those presenting at older ages or manifesting other autoimmune findings, develop a chronic neutropenia disorder. Primary immune neutropenia (in the absence of other disease manifestations) is less common in adults. Rigorous incidence data are lacking, but a small retrospective study in Sheboygan County, Wisconsin, found an annual incidence of 5–10 cases of primary or secondary neutropenia per 100 000 people.

Neutropenia Associated With Systemic Autoimmune or Lymphoproliferative Diseases

Most patients with active systemic lupus erythematosus (SLE) develop neutropenia as part of a more global leukopenia (Chapter 51).[18] Separately, a smaller subset develops severe neutropenia, presumably mediated by ANAs. Sjögren syndrome (Chapter 54) and other systemic autoimmune diseases are also sometimes complicated by immune neutropenia. Lymphoproliferative disorders, such as chronic lymphocytic leukemia (Chapter 78), can occasionally be complicated by immune neutropenia as well.

Felty Syndrome

Immune neutropenia develops in about 1% of patients with rheumatoid arthritis, most commonly (but not exclusively) in association with splenomegaly, a combination designated as Felty syndrome.[19] Patients typically express human leukocyte antigen–DR4 (HLA-DR4) and have long-standing seropositive rheumatoid

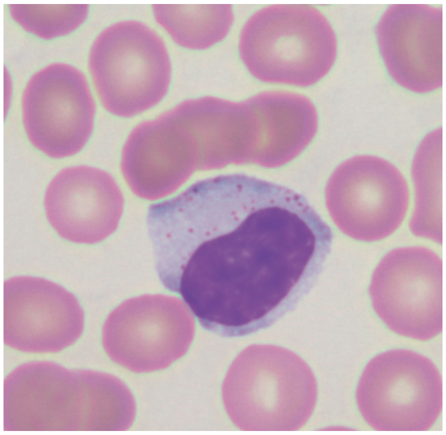


FIG 62.2 The Large Granular Lymphocyte (LGL) Is a Moderate-Sized Lymphocyte Containing Several Distinctive Azurophilic Granules. Additional immunophenotyping is required to distinguish the CD3$^+$, CD8$^+$, CD57$^+$ T-LGL associated with neutropenia from CD16$^+$, CD56$^+$ natural killer (NK)–LGL.

arthritis complicated by erosive joint disease, subcutaneous nodules, and/or leg ulcers. The natural history is often marked by repeated infection, with 5-year mortality rates of >30% reported in some studies. The complex pathophysiology of this disorder is discussed elsewhere in detail (Chapter 52).

T-Cell Large Granular Lymphocyte Leukemia

Large granular lymphocytes (LGL) are medium to large lymphocytes recognizable on light microscopy by their distinctive azurophilic granules (Fig. 62.2). These cells normally constitute <15% of circulating leukocytes and are composed of two major subsets. One is the natural killer (NK) LGLs that express CD2, CD16, and CD56 and is not linked to neutropenia. The other, T-cell (T) LGLs, expresses CD2, CD3, CD8, and CD57 with or without CD16, a phenotype typical of antigen-stimulated mature CD8 effector T cells. Polyclonal and transient monoclonal expansions of these cells sometimes appear in response to viral infection or other immune stimuli without adverse effect. However, some patients develop an indolent lymphoproliferative disease characterized by the accumulation of an autonomous T-LGL clone in blood and in other lymphoid organs, particularly bone marrow, the liver, and/or the spleen. Patients with this disease have a remarkably high incidence of immune neutropenia.[20,21] Even in the absence of gross marrow involvement, over 80% have a neutrophil count of <2000/mm^3 at presentation, and at some point 30–40% develop severe neutropenia with <500 neutrophils/mm^3.[20,21] The pathophysiology resembles Felty syndrome in many respects.

Clinical Overlap Between Felty Syndrome and T-LGL Leukemia

At the extremes these syndromes are easily separable. Patients with classical Felty syndrome have severe rheumatoid arthritis, usually requiring antiinflammatory therapy, incidentally complicated by late neutropenia. This is quite different from the pattern in patients with isolated T-LGL leukemia with neutropenia in the absence of clinical autoimmune disease. Nonetheless, confusing overlap can occur in two settings:

1. More than half the patients with a clinical Felty syndrome may have detectable T-LGL clones in their blood when studied with sensitive flow cytometric or molecular techniques.[20] There have been attempts in the past to classify this pattern as pseudo-Felty syndrome, but the clinical findings and course are often indistinguishable from “classic” Felty syndrome.
2. About half the patients with clinically apparent T-LGL leukemia have circulating rheumatoid factor and immune complexes in their blood; about a third, usually patients expressing HLA-DR4, develop clinically significant arthritis, sometimes requiring antiinflammatory agents.

Although the reason why clonal T-LGL disorders and autoimmunity often coexist is unclear, the tendency toward overlap is clear. Since the pathophysiology and therapy of both conditions are similar, these problems in classification usually have little impact on the initial management of neutropenia. When a patient with “Felty syndrome” develops aggressive T-LGL leukemia or a patient with T-LGL manifests severe rheumatological symptoms, the clinician must be prepared to alter therapy, as needed, to fit the clinical picture.

Although some clinicians have attempted to develop criteria for distinguishing Felty syndrome from T-LGL with pseudo-Felty syndrome, there is now substantial evidence that clonal T-LGL disorders are commonly found in rheumatology patients and that patients with clonal disorders seldom develop a progressive, neoplastic disorder. Conversely, although patients with T-LGL leukemia have a malignancy, it is typically quite indolent and in these cases, the clinical course is often dominated by rheumatological complication and/or neutropenia and not by progressive neoplastic disease.

Drug-Induced Immune Neutropenia

A wide variety of medications cause neutropenia.[22] In some cases, it may be antibody-mediated (see Table 62.2 for common examples); in others, it may be a direct toxic effect on marrow precursors. Often, the mechanism for neutropenia is unclear. A detailed discussion of this topic is beyond the scope of this chapter, but drug-induced neutropenia must always be considered in the differential diagnosis of acquired neutropenia. Rituximab (anti-CD20) is occasionally associated with late-onset neutropenia (LON).[23] LON is typically self-limiting and of no significant clinical consequence; the occurrence and severity of LON may be associated with the total dose of rituximab and the myelotoxicity of the accompanying chemotherapy administered. LON appears to coincide with post-rituximab B-cell recovery. In patients receiving DA-EPOCH (dose-adjusted etoposide/vincristine/doxorubicin/cyclophosphamide/prednisone regimen) chemotherapy for lymphoma, the incidence of LON was 8%, the median time of onset was 175 days (range of 77–204 days), and the duration was 11–16 days. A list of the most common agents associated with drug-induced neutropenia is included in Table 62.2.

Immunopathogenesis

Regulation of Antineutrophil Antibody Production

Isoimmune ANA production in pregnant women represents an “appropriate” immune response to foreign antigens. The production of autoimmune ANAs in other settings reflects immune dysregulation. Older studies attributed primary immune neutropenia of childhood to delayed maturation of T cells responsible for regulating B-cell responses. In this view, the spontaneous recovery usually observed reflects the eventual appearance of mature regulatory or inhibitory T cells. However, this hypothesis is not well supported with clinical laboratory data; the cause of antibody production in other autoimmune diseases remains unclear.

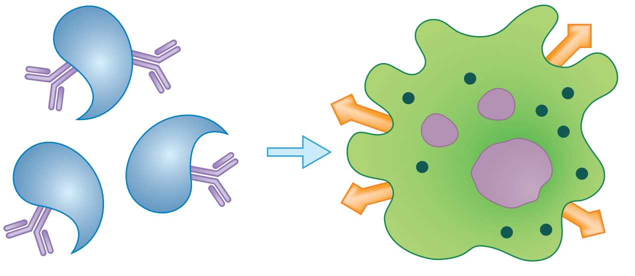


FIG 62.3 Antibody-Bound Platelets Are Cleared From the Circulation by Binding of the Fc Domain of Immunoglobulin G (IgG) to FcγIIA Receptors on Macrophages and Other Cells. The cross-linking of the macrophage receptors sets off a cascade of internal signaling that leads to increased expression of the inhibitory FcγIIB receptors (not shown).

Antibody Specificity

The most important target for antibody responses is the Fcγ receptor IIIb (FcγRIIIb), a low-avidity granulocyte-specific Fcγ receptor that binds IgG immune complexes (Fig. 62.3). This cell surface protein, a glycosyl phosphatidylinositol-linked variant of CD16 selectively expressed on neutrophils, contains several highly immunogenic polymorphisms.[24]

Human neutrophil antigen 1a (HNA-1a) and HNA-1b (previously designated NA1 and NA2) are two related polymorphic forms of FcγRIIIb. Antigenic differences are attributed to a cluster of five base substitutions, all found within exon 3. The two products can be distinguished because two of the base substitutions in HNA-1a disrupt sites of N-linked glycosylation, thereby lowering the size of the protein by ~30 000 Daltons (Da). The frequency of the gene for HNA-1a (*FCGR3B*) varies from 0.3 to 0.55 in different ethnic groups.

HNA-1c (previously designated SH) is another immunogenic polymorphism related to FcγRIIIb. This antigen, present in 5% of northern Europeans, is attributable to an amino acid polymorphism at position 266 in a (reduplicated) copy of the HNA-1b allele of FcγRIIIb.

FcγRIIIb-null phenotype, present in 0.1% of the northern European population, occurs in individuals homozygous for an extensive gene deletion involving all of FcγRIIIb. Because affected individuals have never been exposed to endogenous FcγRIIIb, they often develop a broad antibody response after neutrophil exposure through transfusion or pregnancy.

Other HNAs have also been identified.[25] Perhaps most clinically relevant is HNA-2a (previously NB1), a neutrophil glycoprotein designated CD177. Antibodies against this target have been reported in isoimmune and autoimmune neutropenia. HNA-3a (previously 5b) and HNA-4a (previously MART), an epitope on CD11b, are also immunogenic, but antibody formation against these is seen primarily in heavily transfused individuals and not in autoimmune neutropenia. As described below, individuals with chronic immune neutropenia may develop antibodies against other surface antigens, including CD11b/CDI8 and CD35.[26]

Using a sophisticated antigen discovery strategy, patients with Felty syndrome were found to have a high incidence of antibodies against an intracellular antigen eukaryotic elongation factor 1A-1 (eEF1A-1). Although the finding has been confirmed, this antibody probably is not specific for neutropenia, since it is also present in patients with autoimmune disease who do not have neutropenia.

Patterns of Autoantibody Specificity

Isoimmune neonatal neutropenia is associated with maternal IgG iso- or autoantibodies that can be generated in response to each of the polymorphic alloantigens noted above, particularly polymorphisms affecting FcγRIIIb.

Immune neutropenia in childhood is most commonly associated with IgG directed against the autoantigens HNA-1a and/or HNA-1b.[26] Sera from affected patients often also bind (albeit more weakly) to neutrophils expressing the alternative allele, and in some series, more than half the patients with this entity were shown to produce antibodies capable of binding to nonpolymorphic elements within FcγRIIIb. Using monoclonal antibody (mAb)–specific immobilization of granulocyte antigens (MAIGA) to search for antibodies directed against other neutrophil surface molecules, more complex patterns can be recognized. In one study, autoantibodies against FcγIIIb were seen in 27%, against CD11b/CDI8 (CR3) in 21%, against CD35 (CR1) in 14%, and FcγRII in 2% of patients' sera.[27] Prognostically, patients with antibody responses confined to HNA-1a and/or HNA-1b are more likely to have uncomplicated, self-limiting disease; patients with specificity for nonpolymorphic determinants on the FcγRIIIb molecule and other antigens more often have a generalized immune disorder and more persistent neutropenia.[26] Although antineutrophil assays are not highly quantitative, a recent retrospective study has suggested that patients with low levels of ANAs may have a more favorable long-term prognosis.[28]

In adults, it is often difficult clinically to distinguish immune neutropenia from nonimmune idiopathic neutropenia. Consequently, the sensitivity and specificity of antibody assays in this setting are uncertain. In general, antibodies against HNA-1a or HNA-1b are less common, and antibodies against surface receptors, such as CD11b/CD18 (CR3), are more common in older children and adults than in young children.

Sera from patients with Felty syndrome[20] and T-LGL leukemia[21] are often positive in ANA assays. Interpretation of these results is complicated by the high incidence of nonspecific immune complexes in these populations, which may bind nonspecifically through Fc and complement receptors to the neutrophil surface. Indeed, because it is difficult to distinguish these two types of binding, the incidence of "true" ANAs in these syndromes remains uncertain. Detectable antineutrophil antibodies are low in titer or absent in most patients carefully studied with either diagnosis.[20,21]

Impact of Antibodies and Immune Complexes on Neutrophil Survival

There is good experimental evidence that both ANAs and immune complexes can induce neutropenia *in vivo.* The relative importance of reversible sequestration and neutrophil destruction in inducing neutropenia varies with the experimental model, the character of the antibody/immune complex, spleen size, and presumably other factors as well.

The detection of antineutrophil antibodies in serum, however, does not automatically predict accelerated immune clearance of neutrophils *in vivo.* Some antibodies bind well to neutrophils under *in vitro* assay conditions, without provoking neutrophil destruction *in vivo.*[29] In part, this reflects the inability of these crude assays to distinguish effective from ineffective binding of immunoproteins.

Myelopoiesis in Immune Neutropenia

In primary immune neutropenia, bone marrow is typically normocellular or mildly hypercellular with an increased proportion of early myeloid forms (particularly myelocytes and promyelocytes) and decreased mature forms (neutrophils, bands, and metamyelocytes), a pattern designated maturation arrest. Although maturation arrest can also be seen in a number of other diseases, in this setting, it suggests an expansion in immature precursors with early release of mature components into blood. Rigorous kinetic studies in children with primary neutropenia are not available, but the available data suggest myelopoiesis in this setting is increased.

The findings are more complex in Felty syndrome and T-LGL leukemia. *In vivo* neutrophil kinetic studies and *in vitro* assays of marrow function often document reduced myelopoiesis in these settings.[20,21] This has been attributed to T-cell–mediated and cytokine-mediated suppression. T-LGL leukemia cells constitutively express Fas ligand on their surface and also release significant quantities of soluble Fas ligand into plasma *in vivo.* Reduced myelopoiesis in T-LGL leukemia and some patients with Felty syndrome may be linked to apoptosis instigated by the binding of Fas ligand expressed on the abnormal cells to Fas expressed on the surface of myeloid precursors.[21] Whatever the precise mechanism, reductions in myelopoiesis appear to be a common element in patients with these forms of secondary immune neutropenia.

Diagnosis

Clinical Presentation

Isoimmune neutropenia presents at birth and may persist for up to 6 months. Self-limiting primary autoimmune neutropenia typically presents in early childhood. In older children and adults, neutropenia is more commonly associated with other systemic autoimmune disease, especially rheumatoid arthritis and SLE or T-LGL leukemia. Drug-induced neutropenia must always be considered in patients taking medications.

Laboratory Findings

Blood counts typically demonstrate isolated neutropenia, sometimes with monocytosis. More generalized leukopenia, anemia, and/or thrombocytopenia suggest concurrent SLE or a primary bone marrow disorder, especially aplastic anemia or myelodysplasia.

Examination of the blood film for evidence of abnormalities in other cell lines or increased numbers of LGL is essential. The persistent presence of >2000 LGL/mm^3 for 6 months in itself is diagnostic of T-LGL leukemia; however, normal LGL counts in blood do not preclude this diagnosis. Perhaps a quarter of patients with T-LGL leukemia and immune neutropenia have fewer than 500 monoclonal LGL/mm^3 in blood.[21] The evaluation of patients with small T-LGL clones detected in blood on flow cytometric or molecular testing without clear tissue infiltration remains problematic. At least some of these patients probably have self-limiting “T-cell gammopathies of unknown origin” (Chapter 80) unassociated with overt lymphoproliferation or autoimmunity.

Bone marrow findings in immune neutropenias (as briefly reviewed above) can vary substantially. Perhaps the most important function of the bone marrow examination is to rule out hypoplasia/aplasia, myelokathexis, marked megaloblastic dysplastic changes, or abnormal infiltration with nonhematopoietic cells, which might suggest an alternative diagnosis. The marrow examination may also be helpful in confirming T-LGL leukemia.[30]

Detection of antineutrophil antibodies. Antibodies are assayed clinically using indirect assays (*i.e.,* by measuring the binding of antibodies from patient sera to fixed granulocytes from unrelated individuals). The granulocyte immunofluorescence test (GIFT), which exploits flow cytometry for detection, is most commonly used because of its high sensitivity. The granulocyte agglutination test (GAT) is less sensitive but is particularly valuable used in conjunction with GIFT to detect antibodies against HNA-3a or HNA-1b. Once the presence of an antibody has been confirmed, MAIGA is a valuable technique in identifying the target molecule recognized by the antibody, information that may be very helpful in identifying antibody specificity and in distinguishing granulocyte-specific antibodies from alloantibodies directed against HLA determinants. More precise epitope typing still requires a panel of granulocytes of varied phenotype. Unfortunately, to date, granulocyte panels are both difficult to prepare and impossible to store. Consequently, antibody typing remains a laborious, difficult task. At the second international granulocyte serology workshop, 12 centers independently tested a series of unknown sera. Many laboratories could detect strong HNA-1a antibodies, but the success rate was much lower in defining HNA-1b or HNA-2a antibodies, and individual laboratories varied greatly in their proficiency.[25]

Clinical use of antineutrophil antibody studies. In young children with neutropenia, a positive result is very helpful in distinguishing between immune-mediated and congenital causes. Isoimmune and congenital disease may be apparent from birth, and the former usually resolves within 2–6 months. Using GIFT or GAT assays, ANAs can be detected in more than 70% of children with primary immune neutropenia. When both are used in tandem the yield increases further. A strong positive result strongly supports the diagnosis of immune neutropenia. However, a negative result does not exclude the diagnosis.[28]

Primary autoimmune neutropenia in adults is difficult to distinguish from the ill-defined entity chronic idiopathic neutropenia.[1] Because there is no “gold standard” for distinguishing immune disease from nonimmune disease in this setting, the diagnostic sensitivity and specificity of the ANA assays are unclear. Assays are positive in perhaps a third of adults referred with chronic neutropenia, and a positive result in the absence of other systemic autoimmune disease certainly supports a diagnosis of immune neutropenia. Again, a negative result does not preclude an immune etiology, but it is more consistent with chronic idiopathic neutropenia.

In patients with systemic autoimmune disease or T-LGL leukemia, hyperglobulinemia and circulating immune complexes greatly complicate laboratory evaluation. ANA assays are frequently positive even in the absence of neutropenia. Since the specificity of a positive result is low, its diagnostic value is very limited, and the clinician must be vigilant for other possible causes, especially drug-induced neutropenia.

Therapy

Overview

All patients with neutrophil counts below 1000/mm^3 have some increased risk of infection, but some remain asymptomatic even with absolute neutrophil counts of 500/mm^3 or less. Growth factors can usually improve neutropenia and reduce the risk of infection, but given their expense, inconvenience, and possible side effects, they should be reserved for use in patients with a very low count, or a previous pattern of frequent infection. The indications for immunosuppressives, steroids, and splenectomy are more complex.

THERAPEUTIC PRINCIPLES

Immune Neutropenia

- Palliative treatment of neutropenia is reserved for patients with a neutrophil count below $500/\text{mm}^3$ or recurrent infection.
- Recombinant granulocyte–colony-stimulating factor (G-CSF) is the most effective single agent for palliating neutropenia.
- Immunosuppressive agents, steroids, and splenectomy are reserved for patients with persistent or refractory neutropenia or with other detrimental manifestations of systemic autoimmunity.

Colony-Stimulating Factors

Controlled trials are lacking in this disease setting, but granulocyte–colony-stimulating factor (G-CSF) or granulocyte macrophage–colony-stimulating factor (GM-CSF) usually enhance neutrophil counts in each of the clinical groups discussed. Because of their safety, speed, and efficacy, they have replaced steroids and splenectomy as first-line symptomatic therapy. They should be used at the lowest effective dose, with particular caution in patients with systemic autoimmune disease, as they are prone to leukocytoclastic vasculitis as a complication of therapy.[18]

Immunosuppressive Agents

Because disease is usually self-limiting and responsive to G-CSF, immunosuppressive agents are seldom used in children with primary immune neutropenia. Chronic low-dose methotrexate is considered first-line therapy for patients with Felty or LGL leukemia–associated neutropenia. Cyclophosphamide and cyclosporine are considered second-line therapy for LGL leukemia, and purine analogues, splenectomy, and alemtuzumab are considered third-line therapy for LGL leukemia.

Other Therapy

Each of the following treatments can be effective in reversing neutropenia, but their use has diminished considerably in recent years. IVIG can temporarily reverse neutropenia, particularly in children, probably by blocking Fc receptors responsible for triggering neutrophil destruction. However, G-CSF, which is more convenient to administer and at least as effective, has largely replaced use of IVIG.

Splenectomy and steroids can each reduce immune destruction by suppressing the body's capacity to clear IgG- and complement-coated cells. Over a longer time frame, these treatments can also suppress antibody production, in the first case by removing a major site of production and in the second by reducing antineutrophil antibody production and blocking T-cell–mediated myelosuppression. They can reverse neutropenia in many patients, but their long-term impact on outcome remains unclear. Given their risk and side effects, both modalities are generally reserved for patients resistant to CSFs and low-dose immunosuppressives.

Prophylactic Antibiotics

Where recurrent infection is a problem, oral trimethoprim–sulfamethoxazole (TMP-SMX) is commonly used for prophylaxis, particularly in children. This approach is very reasonable, given its success in other immunocompromised groups, but it has not been tested in a controlled trial. Immunization with pneumococcal vaccine is also recommended in situations where therapeutic splenectomy has been used or is being planned.

IMMUNE-MEDIATED THROMBOCYTOPENIAS

Immune Thrombocytopenia

KEY CONCEPTS

Immune Thrombocytopenic Purpura

- Antibody-mediated destruction or decreased production of platelets
- Both B and T cells important in etiology
- Clearance of platelets mediated by Fcγ receptors
- Diagnosis depends on:
  - Clinical presentation of low platelets
  - Exclusion of other causes of thrombocytopenia

Immune thrombocytopenia (ITP) is an autoimmune syndrome involving antibody and cell-mediated destruction of platelets.[31] ITP can occur in the absence of an identified predisposing factor (primary ITP) or in association to drug exposure or other identified cause (secondary ITP). Recent recommendations from an international working group recommend that a platelet count of $<100 \times 10^9$/L be required for diagnosis.[32]

ITP During Pregnancy and the Neonatal Period

Gestational thrombocytopenia is the most common cause of thrombocytopenia in pregnancy, affecting 5% of all pregnancies and accounting for 75% of cases of pregnancy-associated thrombocytopenia. Gestational thrombocytopenia develops in the late second or third trimester and is not associated with an increased incidence of pregnancy-related complications or the delivery of thrombocytopenic offspring. Preeclampsia and HELLP (hemolysis, elevated liver enzymes, and low platelet count) syndrome are also present in the third trimester and are associated with thrombocytopenia. Thrombotic thrombocytopenic purpura (TTP) and hemolytic uremic syndrome need to be considered in the differential diagnosis of thrombocytopenia of pregnancy. ITP accounts for 5% of pregnancy-associated thrombocytopenia.[33] Lack of hypertension, fever, hemolytic anemia, uremia, and liver function abnormalities serve to distinguish ITP from these other conditions.

Neonatal Alloimmune Thrombocytopenia and Posttransfusion Purpura

Neonatal alloimmune thrombocytopenia (NAT) caused by antihuman platelet antigen 1a (HPA-1a) antibodies occurs in 1 : 1250 pregnancies in the Caucasian population. Severe hemorrhage occurs in 1 : 12 500–1 : 25 000 pregnancies. NAT is caused by maternal antibodies against paternally derived antigens on fetal platelets, most commonly HPA-1a. These antibodies cross the placenta and sensitize fetal HPA-1a–positive platelets, which are then removed in the spleen. Two percent of Caucasian women carry the less frequent HPA-1b and can be immunized against HPA-1a during pregnancy (25%) or at the time of delivery (75%). The HPA-1a antigen is most efficiently presented by HLA-DRB3*0101; 35% of the women at risk (with HPA-1b and HLA- DRB3*01:01) are immunized.[34] In most circumstances, first-time cases of NAT are identified following the birth of a markedly thrombocytopenic neonate; antenatal management is, thus, only possible in subsequent pregnancies. Screening all pregnancies for NAT is under evaluation in several countries. Treatment of NAT consists of maternal administration of high doses of IVIG and corticosteroids.[35]

Platelet transfusions can induce antibodies to either HPA or antigens encoded by the major histocompatibility complex

(MHC). Eight percent of transfused patients mount detectable antibodies to HPA antigens and 45% to HLA antigens. Platelet antibodies cause rapid clearance of transfused platelets and lead patients to become refractory to platelet transfusions.[36] Post-transfusion purpura (PTP) can follow transfusion with platelets or RBCs (in which platelets are found in numbers sufficient to sensitize the recipient).

Drug-Induced Thrombocytopenia

Drug-induced thrombocytopenia (DITP) is an idiosyncratic immune-mediated reaction. The drug-dependent antibodies bind to specific epitopes on platelet surface glycoproteins only in the presence of the sensitizing drug. The drugs bind noncovalently and reversibly to platelets, commonly to GPIIb-IIIa and GPIb-V-IX, and also to the antibody. The Fab domains of the antibodies bind to the drug-platelet epitope. Drug-dependent antibodies inducing thrombocytopenia typically develop 1–2 weeks after exposure to a drug; exceptions to this rule include eptifibatide, tirofiban, and abciximab, as naturally occurring antibodies to these drugs can cause thrombocytopenia within a few hours of the first exposure. Thrombocytopenia with platelet counts frequently below 20×10^9/L develops acutely, recovery occurs 1–2 days after discontinuation of the drug and is usually complete after 1 week, but rarely thrombocytopenia persists for several weeks. Quinidine, quinine, rifampin, tegretol, TMP-SMX, vancomycin, danazol, acetaminophen, abciximab, eptifibatide, tirofiban, and gold salts are the most common culprits. Treatment consists of discontinuing the offending drug; platelet transfusions are sometimes necessary.[37,38]

Heparin-induced thrombocytopenia (HIT) is a special case that is caused by antibodies to platelet factor 4 (PF4)–heparin complexes. It can be associated with life-threatening thrombosis. The antibody–PF4–heparin complex activates platelets, resulting in a high risk of both arterial and venous thrombotic events.[39] Thrombocytopenia occurs as a result of clearance of platelet aggregates induced by the antibody and usually appears 5–7 days after treatment with heparin (or low-molecular-weight heparin) unless a patient has been previously exposed to heparin. In the event of prior exposure, especially within the last 100 days, thrombocytopenia can occur within 1 day of heparin administration. Even small doses of heparin given as "flushes" to maintain intravenous catheter patency can be sufficient to cause HIT with thrombosis. In cases of suspected HIT, all heparins should be stopped, and an alternative anticoagulant agent, such as the direct thrombin inhibitors argatroban and lepirudin, should be used. Fondaparinux is an anticoagulant synthesized from the pentasaccharide core of the heparin molecule and may be another option for the treatment of HIT, since it binds to HIT antibodies but does not activate platelets and cause thrombosis. In cases of confirmed or strongly suspected HIT, anticoagulation should be continued for 3 months because the risk of thrombosis persists in this group of patients. Coumadin is generally used for long-term anticoagulation; it should be started concomitant with a heparin alternative because of the increased risk of thrombosis during the initial depletion of anticoagulant factors (proteins C and S) by warfarin (Coumadin). The role of direct oral anticoagulants that target thrombin or activated factor X remains to be defined.

Pathogenesis

In the majority of patients the underlying defects leading to autoantibody production remain unclear. In some patients, ITP follows exposure to viral or bacterial antigens. Molecular mimicry appears to play a role in the development of self-reactive platelet antibodies following infections. Human immunodeficiency virus (HIV), hepatitis C virus (HCV), and *Helicobacter pylori* infections have been associated with ITP. *H. pylori* CagA antigen appears to cross-react with platelet antigens,[40] which may explain the association with ITP and *H. pylori* infection. Interestingly, empiric treatment of *Helicobacter* infection with amoxicillin, clarithromycin, and proton pump inhibitors in patients suspected of infection leads to 53% remission of ITP whether or not *Helicobacter* is eradicated, suggesting immunomodulatory effects of treatment that do not involve elimination of the bacterial antigen with cross-reactivity to platelet antigens.[41] Patients who are not infected with *H. pylori* but are treated empirically with antibiotics do not have a significant platelet response rate (6.5%), which suggests that empiric therapy is not useful in the absence of a documented infection.[41]

NAT is caused by maternal antibodies against the HPAs that the fetus carries but which the mother lacks (most commonly HPA-1a). It is more likely to cause intracerebral hemorrhage (10–20% of cases) than is maternal ITP, and it is very likely to recur if it has occurred in a previous pregnancy.[42] IVIG and platelet transfusions using maternal platelets (to ensure that the offending antigen is not present in the transfused product) are often necessary.

Antibodies against human platelet antigens are also responsible for PTP, in which the recipient has an acquired antibody directed against a platelet antigen on the donor platelets. It is not understood why the antibodies also destroy the patient's own "antigen-negative" platelets. Treatment consists of IVIG and corticosteroids, and sometimes plasma exchange.[34] Transplantation-mediated alloimmune thrombocytopenia (TMAT) may occur as the result of donor-origin antibodies produced by passenger B cells directed against the recipient platelet alloantigen (HPA-1a).[43]

The most common epitopes for platelet antibodies in ITP are the platelet GPIIb/IIa and GPIb-IX receptors.[44] The autoantibodies serve as opsonins resulting in the clearance of platelets by FcγR-bearing cells in the reticuloendothelial system (see Fig. 62.3). There is upregulation of genes involved in cell-mediated cytotoxicity via $CD3^+CD8^+$ T lymphocytes,[45] with T-helper 1 (Th1)–associated cytokines predominating.[46] Regulatory T cells (Tregs) are decreased,[47] and B-cell activation is increased.[48] There is evidence of suppression of megakaryopoiesis by both T lymphocytes[49] and ITP plasma/IgG.[50] In addition, megakaryocyte and platelet production are dependent on thrombopoietin signaling through binding to the Mpl receptor, and patients with ITP have reduced thrombopoietin levels despite the presence of low platelet counts.[51]

Laboratory Diagnosis[52]

Peripheral blood smear examination is important to evaluate for the presence of schistocytes, leukocyte adhesion bodies in MYH-9-related disease, and giant or large platelets in inherited thrombocytopenias and to exclude ethylenediaminetetraacetic acid (EDTA)–dependent platelet agglutination.

Bone marrow aspirate, biopsy, flow cytometry, and cytogenetics should be considered in patients older than 60 years of age and in patients with systemic symptoms. The detection of *H. pylori* with the urea breath test is recommended in adults, and testing in children is recommended in high prevalence areas. Routine serological evaluation for HIV and HCV is recommended in adults. Baseline Igs (IgG, IgA, IgM) should be measured in both adults and children to diagnose such conditions as common variable immunodeficiency and selective IgA deficiency (Chapter 34), in which ITP is a common complication.

Antiplatelet antibody testing to specific platelet glycoproteins is not routinely recommended; the test is neither sensitive nor specific for ITP. Routine testing for anticardiolipin antibodies is not recommended in the absence of symptoms of antiphospholipid syndrome. DAT should be obtained if hemolytic anemia is suspected and in patients in whom anti-D antiglobulin treatment is considered. Blood group Rh(D) typing should be obtained if anti-D antiglobulin treatment is considered.

Where a microangiopathic process is evident, especially with concomitant renal failure, fever, or cognitive impairment, measurement of ADAMTS13 (ADintegrin-like And Metalloprotease with ThromboSpondin type 13 motifs) is warranted to rule out TTP.

Therapy of ITP[53]

THERAPEUTIC PRINCIPLES

Immune Thrombocytopenic Purpura (ITP)

- Primary goal—prevent bleeding
- Treatment not usually necessary if platelets >30 000/µL
- Prednisone and/or intravenous immunoglobulin (IVIG) usual first medications
- Anti-D may avert/delay splenectomy
- Treat refractory ITP with immunosuppressives
  - Mycophenolate mofetil, cyclosporine, alemtuzumab, cyclophosphamide, rituximab (not approved by the US Food and Drug Administration [FDA] for this use)
  - Splenectomy for persistent disease (immunize patient with pneumococcal vaccine before surgery)

In most circumstances, the goals of therapy are to keep the platelet count above 30×10^9/L and to minimize toxicity. In the absence of hemostatic comorbidity, trauma, or surgery, intracranial hemorrhage is rare in patients with a platelet count above 20×10^9/L. Treatment of ITP should be individualized, but first-line therapy usually consists of corticosteroids supplemented with either IVIG or anti-Rh(D) in Rh(D)-positive, nonsplenectomized patients. Intravascular hemolysis, disseminated intravascular coagulation, and renal failure have been reported with the use of anti-Rh(D). Anti-Rh(D) should be avoided in patients with underlying hemolysis or a positive result of the DAT that is not the result of prior therapy. Several uncontrolled studies suggest that bolus oral therapy with oral dexamethasone for 1–4 cycles increases response rates and prolongs remission duration.

Second-line therapy should be initiated if there is absence of a robust response at 1 month or once significant steroid related toxicity supervenes. Two-thirds of patients obtain a durable long-term remission following splenectomy, and these patients should receive immunizations with the pneumococcal, *Haemophilus influenzae* type b, and the quadrivalent meningococcal vaccines before splenectomy. A single course of rituximab (anti-CD20) (375 mg/m^2 weekly for 4 weeks) is associated with 40% complete remission at 1 year and 15–20% complete remission at 5 years. Patients who relapse after an initial response usually respond to a second course. Rituximab is contraindicated in patients with active hepatitis B. Cases of multifocal leukoencephalopathy have been reported in HIV-negative patients treated with rituximab. The thrombopoietin receptor agonists (TRAs) romiplostim and eltrombopag have shown significant sustained activity in patients with ITP. Romiplostim is administered subcutaneously and eltrombopag orally. Increased bone marrow reticulin fibrosis has been described in some patients receiving TRAs. Initial dosing of eltrombopag should be reduced by 50% in patients of Southeast Asian origin. Danazol is an attenuated androgen administered orally and has been used, with some success, as a steroid-sparing agent. Third-line therapy is reserved for patients who are unresponsive to or ineligible for first- and second-line therapies. These agents include cyclophosphamide, azathioprine, mycophenolate mofetil, cyclosporine, alemtuzumab, and vinca alkaloids. One novel approach to ITP consists of blockade of the FcγIIIA receptors on macrophages in the spleen that mediate uptake of antibody-bound platelets. Prior attempts at FcγIIIA receptor blockade to ameliorate ITP using mAbs to FcγIIIA have improved platelet counts in half the patients refractory to other treatments but were abandoned because of nausea, vomiting, fever, and rarely anaphylaxis that were presumed to be caused by immune stimulation from cross-linking of FcγIIIA receptors by bivalent IgG. Preclinical studies suggest monovalent anti-FcγIIIA fused to albumin (to prolong the half-life) may block the FcγIIIA receptor without stimulating the immune system, and this approach may be tested clinically, again.[54]

Therapy of ITP During Pregnancy and the Neonatal Period

Patients with ITP with platelet counts higher than $20\text{–}30 \times 10^9$/L usually do not require treatment during the first and second trimesters of pregnancy; platelet counts should be monitored more closely as delivery approaches. Obstetrical anesthesiologists recommend a minimum platelet count of 75×10^9/L to allow administration of spinal or epidural anesthesia. A platelet count of at least 50×10^9/L is generally considered adequate to allow a Cesarean section.[52]

First-line therapy consists of prednisone and IVIG. Prednisone is initiated at a low dose (10–20 mg) and then adjusted to the minimum dose that produces a hemostatically effective platelet count. Corticosteroids may exacerbate hypertension, diabetes, and osteoporosis during pregnancy and are associated with weight gain and psychosis. Corticosteroid use in the first trimester has been associated with congenital anomalies, such as orofacial clefts, and low-dose corticosteroids have no impact on the fetal platelet count. IVIG is an effective means of increasing the platelet count rapidly and is preferred over prednisone in this setting. The American Society of Hematology (ASH) guidelines consider IVIG to be an appropriate first-line agent for severe thrombocytopenia or for thrombocytopenic bleeding in the third trimester. Anti-Rh(D) is safe and effective for the mother and fetus in nonsplenectomized Rh(D)-positive patients.[33,52]

Second-level therapy before delivery may require utilizing higher doses of corticosteroids combined with IVIG to achieve an adequate platelet count. Splenectomy is an option for patients who fail corticosteroids and IVIG. Remission of ITP is achieved in 75% of patients. Splenectomy is optimally performed during the second trimester. The risk of inducing premature labor is high during the first trimester, and the enlarged uterus renders the procedure difficult if not impossible during the third trimester. Vinca alkaloids, rituximab, danazol, thrombopoietin receptor agonists, and immunosuppressive drugs (other than azathioprine) should be avoided during pregnancy.[33,52]

Management of parturition is based on the finding that there is no correlation between platelet counts or ITP status of the mothers and the development of neonatal thrombocytopenia. The most reliable predictor of neonatal thrombocytopenia is a history of thrombocytopenia at delivery in a prior sibling.[33] In a large meta-analysis, fetal platelet counts below 50×10^9/L were observed in 10.1%, and fetal platelet counts below 20 × 10^9/L were observed in 4.2%.[42] The risk of intracranial hemorrhage is below 1%, and the risk of fetal platelet count determination by either fetal scalp sampling or umbilical blood sampling likely exceeds the risk of fetal intracranial hemorrhage, negating its clinical utility. The mode of delivery has no impact on the incidence of intracranial hemorrhage.

Platelet counts should be monitored for 1 week following delivery. In a study of 61 patients, 66% of neonates had a further drop in platelets following delivery; the nadir occurred at day 2, and the counts stabilized or began to rise at day 7.[34] ASH guidelines recommend that children with a platelet count below 20 × 10^9/L or those with hemorrhage be treated with IVIG. The concurrent use of corticosteroids is controversial because of the risk of neonatal sepsis. Neonates with platelet counts below 50 × 10^9/L should have brain imaging performed to detect occult intracranial hemorrhage.

ON THE HORIZON

- For many years immunosuppressive therapy for immune thrombocytopenic purpura (ITP) has consisted of corticosteroids and infusions of intravenous immunoglobulins (IVIGs), and more recently anti-Rh antibodies (in suitable candidates), and/or rituximab in patients whose disease is refractory to first-line therapy. Splenectomy may be indicated in those who fail medical therapy altogether.
- All ITP seems to be mediated by interaction of antibody-coated platelets with Fc receptors in the reticuloendothelial system of the spleen (and the liver). Many investigators have sought to interrupt the interaction between immunoglobulin (Ig)–coated platelets and Fcγ receptors by use of monoclonal antibodies to the Fcγ receptor itself. However, the use of traditional murine monoclonal antibodies (mAbs) to Fcγ has been associated with adverse events, perhaps as a result of cross-linking of Fcγ receptors on the surface of macrophages, resulting in undesirable immune stimulation and even anaphylaxis. One strategy to circumvent these problems would be to use monovalent Fab fragments that would bind but not cross-link the Fcγ receptors. Yu et al.[54] have reported on the design of a monovalent fusion protein composed of the single chain variable region (scFv) of the anti-FcγRIIIA antibody, fused to human serum albumin so as to increase its size and prolong its half-life. These investigators have shown in a mouse model of ITP that the novel construct can prevent IgG-platelet interaction with FcγRIIIA receptors without harmful immune stimulation.[54] If this can be replicated in larger animal models, it may be eventually tested for use in humans with ITP. Further, in theory this approach could be applied to other diseases, such as autoimmune hemolytic anemia.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Barros MM, Blajchman MA, Bordin JO. Warm autoimmune hemolytic anemia: recent progress in understanding the immunobiology and treatment. Transfus Med Rev 2010;24:195–210.
2. Petz LD, Garratty G. Classification and clinical characteristics of autoimmune hemolytic anemias. In: Petz LD, Garratty G, editors. Immune hemolytic anemias. 2nd ed. Philadelphia, PA: Churchill Livingstone; 2004. p. 61.
3. Garratty G. Immune hemolytic anemia associated with drug therapy. Blood Rev 2010;24:143–50.
4. Hoffman PC. Immune hemolytic anemia—selected topics. Hematology Am Soc Hematol Educ Program 2009;80–6.
5. McNicholl FP. Clinical syndromes associated with cold agglutinins. Transfus Sci 2000;22:125–33.
6. Berentsen S, Randen U, Vagan AM, et al. High response rate and durable remissions following fludarabine and rituximab combination therapy for chronic cold agglutinin disease. Blood 2010;116:3180–4.
7. Petz LD. Immune hemolysis associated with transplantation. Semin Hematol 2005;42:145–55.
8. Avent ND, Reid ME. The Rh blood system: a review. Blood 2000;95:375–87.
9. Mollison PL, Engelfriet CP, Contreras M. Immunology of red cells. In: Mollison PL, Engelfriet CP, Contreras M, editors. Blood transfusion in clinical medicine. 10th ed. London: Blackwell Science; 1997. p. 60.
10. Hughes-Jones NC. Red-cell antigens, antibodies and their interaction. Clin Haematol 1975;4:29–43.
11. Petz LD, Garratty G. Mechanisms of immune hemolysis. In: Petz LD, Garratty G, editors. Immune hemolytic anemias. 2nd ed. Philadelphia, PA: Churchill Livingstone; 2004. p. 133.
12. Kurlander RJ, Rosse WF, Logue GL. Quantitative influence of antibody and complement coating of red cells on monocyte-mediated cell lysis. J Clin Invest 1978;61:1309–19.
13. Frank MM, Schreiber AD, Atkinson JP, et al. Pathophysiology of immune hemolytic anemia. Ann Intern Med 1977;87:210–22.
14. Moise KJ. Red blood cell alloimmunization in pregnancy. Semin Hematol 2005;42:169–78.
15. Lechner K, Jager U. How I treat autoimmune hemolytic anemias in adults. Blood 2010;116:1831–8.
16. Maheshwari A, Christensen RD, Calhoun DA. Immune-mediated neutropenia in the neonate. Acta Paediatr Suppl 2002;91:98–103.
17. Capsoni F, Sarzi-Puttini P, Zanella A. Primary and secondary autoimmune neutropenia. Arthritis Res Ther 2005;7:208–14.
18. Starkebaum G. Chronic neutropenia associated with autoimmune disease. Semin Hematol 2002;39:121–7.
19. Balint GP, Balint PV. Felty's syndrome. Best Pract Res Clin Rheumatol 2004;18:631–45.
20. Mohan SJ, Maciejewski JP. Diagnosis and therapy of neutropenia in large granular lymphocyte leukemia. Curr Opin Hematol 2009;16:27–34.
21. Lamy T, Loughran TP. How I treat LGL leukemia. Blood 2011;117:2764–74.
22. Tesfa D, Keisu M, Palmblad J. Idiosyncratic drug-induced agranulocytosis: Possible mechanisms and management. Am J Hematol 2009;84:428–34.
23. Grant C, Wilson WH, Dunleavy K. Neutropenia associated with rituximab therapy. Curr Opin Hematol 2010;18:49–54.
24. Bux J. Molecular nature of granulocyte antigens. Transfus Clin Biol 2001;8:242–7.
25. Bux J, Chapman J. Report on the second international granulocyte serology workshop. Transfusion 1997;37:977–83.
26. Bruin MC, von dem Borne AE, Tamminga RY, et al. Neutrophil antibody specificity in different types of childhood autoimmune neutropenia. Blood 1999;94:1797–802.
27. Bux J, Behrens G, Jaeger G, et al. Diagnosis and clinical course of autoimmune neutropenia in infancy: analysis of 240 cases. Blood 1998;91:181–6.
28. Kobayashi M, Nakamura K, Kawaguchi H, et al. Significance of the detection of anti-neutrophil antibodies in children with chronic neutropenia. Blood 2002;99:3468–71.
29. McCullough J, Weibler BJ, Clay M, et al. Effect of leukocyte antibodies on the fate in vivo of indium-111-labelled granulocytes. Blood 1981;58:164–70.
30. Morice WG, Kurtin PJ, Tefferi A, et al. Distinct bone marrow findings in T-cell granular lymphocytic leukemia revealed by paraffin section immunoperoxidase stains for CD8. TIA-1, and granzyme B. Blood 2002;99:268–74.
31. Cines DB, Bussel JB, Liebman HA, et al. The ITP syndrome: pathogenic and clinical diversity. Blood 2009;113:6511–21.
32. Rodeghiero F, Stasi R, Gernsheimer T, et al. Standardization of terminology, definitions of outcome criteria in immune thrombocytopenic purport of adults and children: report from an international working group. Blood 2009;113:2386–93.
33. Stavrou E, McCrea KR. Immune thrombocytopenia in pregnancy. Hematol Oncol Clin North Am 2009;23:1299–316.

34. Husebekk A, Killie MK, Kjeldsen-Kragh J, et al. Is it time to implement HPA-1 screening in pregnancy? Curr Opin Hematol 2009;16:497–502.
35. Vinograd CA, Bussel JB. Antenatal treatment of fetal alloimmune thrombocytopenia: a current perspective. Haematologica 2010;95:1807–11.
36. Kiefel V, Konig C, Kroll H, et al. Platelet antibodies in transfused patients. Transfusion 2001;41:766–70.
37. George JN, Aster RH. Drug-induced thrombocytopenia: pathogenesis, evaluation, and management. Hematology Am Soc Hematol Educ Program 2009;153–8.
38. Aster RH, Curtis BR, McFarland JG, et al. Drug-induced immune thrombocytopenia: pathogenesis, diagnosis, and management. J Thromb Haemost 2009;7:911–18.
39. Otis SA, Zehnder JL. Heparin-induced thrombocytopenia: current status and diagnostic challenges. Am J Hematol 2010;85:700–6.
40. Takahashi T, Yujiri T, Shinohara K, et al. Molecular mimicry by *Helicobacter pylori* CagA protein may be involved in the pathogenesis of *H. pylori*-associated chronic idiopathic thrombocytopenic purpura. Br J Haematol 2004;124:91–6.
41. Arnold DM, Bernotas A, Nazi I, et al. Platelet count response to *H. pylori* treatment in patients with immune thrombocytopenic purpura with and without *H. pylori* infection: a systematic review. Haematologica 2009;94:850–6.
42. Burrows RF, Kelton JG. Pregnancy in patients with idiopathic thrombocytopenic purpura: assessing the risks for the infant at delivery. Obstet Gynecol Surv 1993;48:781–8.
43. West KA, Anderson DR, McAlister VC, et al. Alloimmune thrombocytopenia after organ transplantation. N Engl J Med 1999;341:1504–7.
44. McMillan R. Antiplatelet antibodies in chronic adult immune thrombocytopenic purpura: assays and epitopes. J Pediatr Hematol Oncol 2003;25:S57–61.
45. Olson B, Anderson PO, Jernas M, et al. T-cell mediated cytotoxicity toward platelets in chronic idiopathic thrombocytopenic purpura. Nat Med 2003;9:1123–4.
46. Panitsas FP, Theodoropoulou M, Kouraklis A, et al. Adult chronic idiopathic thrombocytopenic purpura (ITP) is the manifestation of a type-1 polarized immune response. Blood 2004;103:2645–7.
47. Yu J, Heck S, Patel V, et al. Defective circulating CD25 regulatory T cells in patients with chronic immune thrombocytopenic purpura. Blood 2008;112:1325–8.
48. Emmerich F, Bal G, Barakat A, et al. High-level serum B-cell activating factor and promoter polymorphisms in patients with idiopathic thrombocytopenic purport. Br J Haematol 2007;136:309–14.
49. Olson B, Ridell B, Carlsson L, et al. Recruitment of T cells in the bone marrow of ITP patients possibly due to elevated expression of VLA-4 and CX3CR1. Blood 2008;112:1078–84.
50. McMillan R, Wang L, Tomer A, et al. Suppression of in vitro megakaryocytic production by antiplatelet autoantibodies from adult patients with chronic ITP. Blood 2004;103:1364–9.
51. Aledort LM, Hayward CP, Chen MG, et al. Prospective screening of 205 patients with ITP, including diagnosis, serological markers, and the relationship between platelet counts, endogenous thrombopoietin, and circulating anti-thrombopoietin antibodies. Am J Hematol 2004;76:205–13.
52. Provan D, Stasi R, Newland AC, et al. International consensus report on the investigation and management of primary immune thrombocytopenia. Blood 2010;115:168–86.
53. Vesely SK, Perdue JJ, Rizvi MA, et al. Management of adult patients with persistent idiopathic thrombocytopenic purpura following splenectomy: a systematic review. Ann Intern Med 2004;140:112–20.
54. Yu X, Menard L, Prechl J, et al. Monovalent Fc receptor blockade by an anti–Fcγ receptor/albumin fusion protein ameliorates murine ITP with abrogated toxicity. Blood 2016;127:132–8.

MULTIPLE-CHOICE QUESTIONS

1. Which of the following treatments is LEAST likely to improve platelet counts in an patient with type A- blood and immune thrombocytopenic purpura (ITP)?
   - **a.** Rituximab 375 mg/m2 IV
   - **b.** Prednisone 1 mg/kg PO
   - **c.** WinRho 250 IU/kg (50 micrograms/kg) IV
   - **d.** IVIG, 2 g/kg continuous IV infusion over 48 hours

2. Which is true of hemolysis caused by a "cold" antibody elicited by exposure to drugs?
   - **a.** Usually caused by immunoglobulin G (IgG) antibody against Rh antigens
   - **b.** Usually accompanied by a direct antibody test (DAT) (Coombs test) that is positive for IgM and IgG and negative for complement
   - **c.** Usually accompanied by a DAT (Coombs test) that is negative for IgM and IgG and positive for complement
   - **d.** Usually caused by a Donath-Landsteiner IgG against Rh antigen on the surface of the red blood cell (RBC)

3. Hemolytic anemia of the newborn:
   - **a.** Usually occurs when the mother of the infant is Rh^+ and the father is Rh^-
   - **b.** Is more likely in a first pregnancy than subsequent pregnancies between the same two parents
   - **c.** May be prevented by administration of RhoGam at week 28 of pregnancy when the parents are Rh incompatible
   - **d.** Is more common if the parents are ABO incompatible

4. Neutropenia associated with administration of rituximab:
   - **a.** Is usually self-limiting
   - **b.** Precludes future use of rituximab
   - **c.** Typically occurs 7–10 days after administration of rituximab
   - **d.** Does not respond to administration of granulocyte–colony-stimulating factor (G-CSF)

63

Bullous Diseases of the Skin and Mucous Membranes

Adela Rambi G. Cardones, Russell P. Hall III

PEMPHIGUS

Pemphigus is a group of cutaneous autoimmune blistering diseases characterized by intraepidermal blistering of the skin and mucous membranes. Pemphigus comprises four distinct conditions: pemphigus vulgaris (PV), pemphigus foliaceus (PF), paraneoplastic pemphigus, and immunoglobulin A (IgA) pemphigus.

Pemphigus Vulgaris

The most common and most severe form of pemphigus, PV is still a rare disease, with an incidence estimated at 0.5–1.6 per 100 000.[1-3] PV presents most often in the fourth to sixth decades but can occur at any age. PV is characterized by the presence of flaccid blisters and erosions of the skin and mucous membranes (Fig. 63.1). Virtually all patients with PV develop oral erosions at some point during the course of their disease. Most present with oral erosions, and for many, these are their only disease manifestation for months to years (Fig. 63.2). These erosions are often persistent and shaggy; they can involve all areas of the oral mucosa and extend into the esophagus. In severe cases, the conjunctiva, nasal, anal, cervical, or urethral mucosae may be affected.

PV blisters are flaccid and easily ruptured; they can occur on inflamed or noninflamed skin. Once broken, these blisters leave large, nonhealing erosions. Blisters can occur on any area of skin, most often beginning on the head, neck, or trunk. Generalized blistering is not uncommon, especially if diagnosis and/or treatment have been delayed.

Pemphigus vegetans is an unusual variant of PV, in which lesions develop mainly in the axilla, groin, and other flexural areas. The characteristic lesions are plaques of hypertrophic granulation tissue with occasional pustules. These can occur *de novo* or after healing of PV lesions. The etiology of this unusual clinical presentation is unknown.

Before the advent of glucocorticoids, PV was almost uniformly fatal, with patients developing large areas of denuded skin and dying from overwhelming sepsis. Since the introduction of systemic glucocorticoids, the mortality rate has fallen to approximately 10%. The residual mortality is largely due to the side effects of the high doses of glucocorticoids required to treat the disease, particularly in debilitated or older patients.[4]

Histologically, PV is characterized by intraepithelial blisters, with acantholysis (breaking apart of the suprabasilar portion of the epidermis). Biopsy of early PV lesions shows that basal epidermal cells remain attached to the dermis, forming a “row of tombstones,” with loss of attachment of individual keratinocytes to each other, resulting in intraepidermal blisters.

In the early 1960s, investigators found IgG antibodies bound to the epithelial keratinocyte cell surface in the skin of patients with PV and showed that the serum of these patients contained IgG antibodies that bound in an identical pattern to normal human skin.[5] These findings demonstrated that PV is an autoimmune disease directed against a normal component of stratified squamous epithelium.

Pemphigus Foliaceus

PF is characterized by blisters within the epidermis, resulting in superficial erosions with scaling and crusting (Fig. 63.3). In contrast to patients with PV, patients with PF rarely develop mucous membrane lesions. Because the blisters are superficial, many patients present with erythematous, scaling, and crusted plaques, which occur principally on the head, neck, and trunk (Fig. 63.4). The prognosis for patients with PF is markedly better than that for patients with PV, although patients with severe PF do occasionally die as a result of infections and/or the side effects of therapy.

An unusual variant of PF is the endemic form of the disease, fogo selvagem (FS) (which is the Portuguese term for “wild fire”). Although clinically identical to the sporadic form of PF, FS only occurs in certain rural areas of South America and North Africa.[6] The epidemiology suggests an infectious etiology, and an arthropod vector may be involved in spreading the disease.

Pemphigus erythematosus, or Senear-Usher syndrome, is another PF variant. These patients have typical features of PF with additional characteristics suggestive of systemic lupus erythematosus (SLE) (Chapter 51), such as malar involvement and antinuclear antibodies. The clinical course tends to parallel that of PF, and most patients do not develop SLE. The pathogenesis of this form of PF is poorly understood.

PF can also be associated with drugs. The most common agents that induce pemphigus are D-penicillamine and angiotensin-converting enzyme inhibitors.[7] Although these drugs can induce a PV-like picture, the clinical manifestations are most often similar to PF. Patients with drug-induced pemphigus have autoantibodies directed against the keratinocyte cell surface; the mechanism(s) causing these antibodies are unknown, but reactive sulfhydryl or amide groups on the drugs are thought to be responsible.[7]

Histological examination of PF blisters reveals intraepidermal blisters with acantholysis of the most superficial portion of skin (the granular cell layer), resulting in blisters with an epithelial base rather than the basal cell layer as seen in PV.

The immunohistological features of PF are indistinguishable from those of PV (Fig. 63.5). Cell-surface IgG deposits are present,

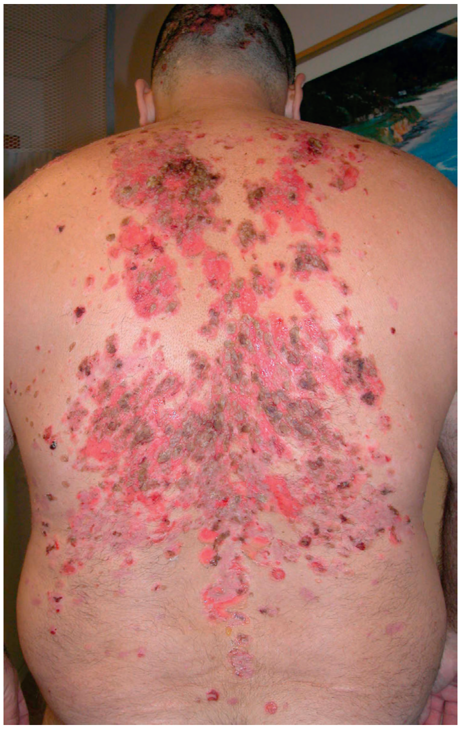


FIG 63.1 Patient with pemphigus vulgaris showing ulcerative plaques and crust involving the back.

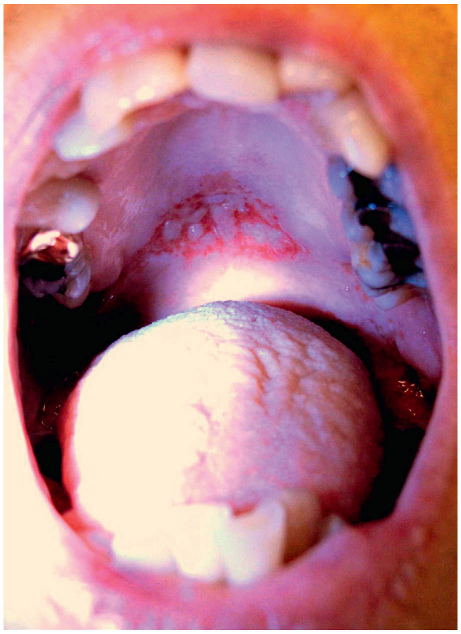


FIG 63.2 Patient with pemphigus vulgaris showing ulceration of the soft palate.

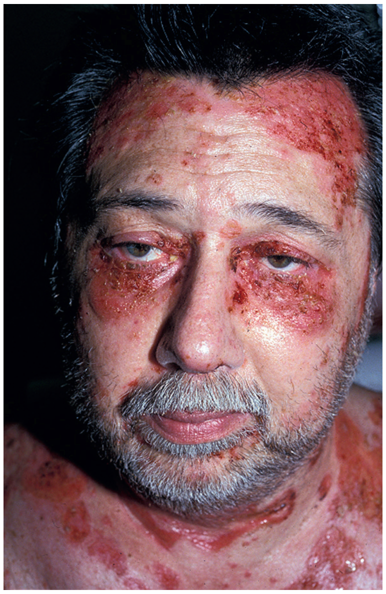


FIG 63.3 Patient with pemphigus foliaceus showing superficial erosions and crust formation on the head and neck.

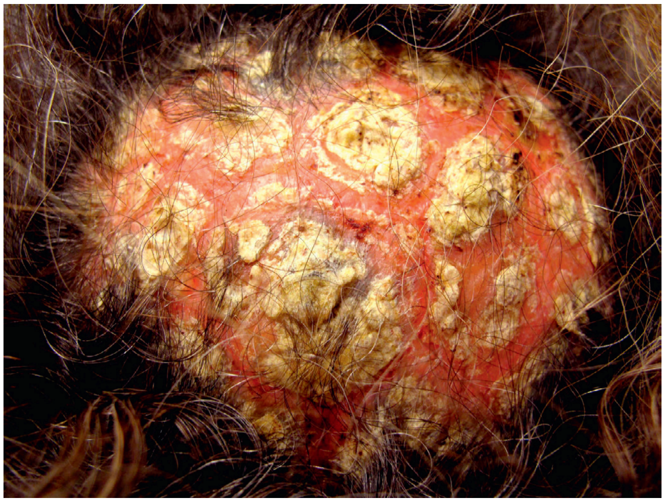


FIG 63.4 Patient with pemphigus foliaceus showing involvement of the scalp with erosions and adherent crust with alopecia. The patient's hair regrew fully after initiation of treatment.

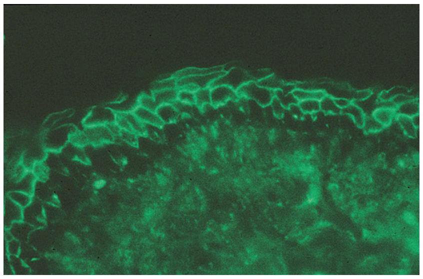


FIG 63.5 Direct immunofluorescence of perilesional normal-appearing skin from a patient with pemphigus foliaceus using antibodies directed against human immunoglobulin G (IgG). Cell-surface IgG deposits are seen on the epidermal keratinocytes. A similar pattern is seen in patients with pemphigus vulgaris.

and serum from patients with PF contains IgG antibodies, which bind to normal stratified squamous epithelium. Although the immunofluorescence patterns of PF and PV are virtually identical, the antigenic targets of these autoantibodies are different (see below).

Paraneoplastic Pemphigus

Anhalt described a series of patients with severe oral erosions and polymorphic skin lesions suggestive of coexistent PV and erythema multiforme.[8] All these patients had associated malignancies, which were mostly lymphoreticular, including lymphoma, chronic lymphocytic leukemia, and thymoma. The histological features of the skin lesions in these patients resembled PV, including suprabasilar blister formation, but with additional keratinocyte cell necrosis as found in erythema multiforme. Skin biopsies from these patients revealed IgG bound to keratinocyte cell surfaces, and granular–linear deposits of IgG and C3 at the basement membrane. Serum contained circulating IgG antibodies, which bound to normal keratinocyte cell surfaces of squamous epithelium. Although initial reports suggested that these antibodies bound to stratified and nonstratified squamous epithelium, this is not always the case. The antigenic targets of these antibodies were also distinct from both PV and PF (see below).

IgA Pemphigus

IgA pemphigus presents as an annular, vesiculopustular eruption characterized by intraepidermal blisters with numerous neutrophils. Intercellular deposits of IgA are present. This entity is also known as *intraepidermal neutrophilic IgA dermatosis, intercellular IgA vesiculopustular dermatosis,* and *intercellular IgA dermatosis.*[9]

IgA pemphigus is a distinct clinical entity with two subtypes that differ in histology and epidermal IgA deposition patterns: subcorneal pustular dermatosis (SPD) and intraepidermal neutrophilic IgA dermatosis (IEN). Biopsy specimens from SPD show subcorneal acantholysis and pustules, with intercellular IgA deposition mainly in the upper epidermis. The IEN subtype is characterized by pustules deeper in the epidermis, with intercellular IgA deposition throughout the epidermis. Circulating IgA antibodies to intraepidermal structures are present in ~50% of cases. Desmocollin 1 (a desmosomal cadherin) has been identified as a target autoantigen in most cases with the SPD phenotype. In some IEN-type cases, desmoglein-1 (Dsg1) and Dsg3 are autoantibody targets, but in most IEN cases the autoantigen remains unknown. Patients of both types commonly present with flat pustules, often on a slightly erythematous base, which tend to coalesce to form annular patterns. IgA pemphigus mainly affects the trunk, axillae, and groin but not the mucosa. Histology of lesional skin reveals subcorneal or intraepidermal blisters, with acantholytic cells and a neutrophilic infiltrate.[9]

Pathogenesis

The antigenic targets in pemphigus have been identified as components of desmosomes, key structures involved in keratinocyte adhesion (Table 63.1). IgG from patients with PF binds to a 160-kilodalton (kDa) glycoprotein, Dsg1. IgG from patients with PV with only mucosal lesions binds predominantly to Dsg3, an epithelial cadherin with significant homology to Dsg1. Patients with PV who have both skin and mucosal diseases have IgG antibodies directed against both Dsg1 and Dsg3.[10] In paraneoplastic pemphigus, IgG antibodies bind to members of the plakin family of desmosomal/hemidesmosomal proteins (desmoplakin I, desmoplakin II, envoplakin, periplakin, plectin, and BPAG1) as well as to Dsg1 and Dsg3.[8]

It has been clearly shown that these autoantibodies are pathogenic. Neonatal mice injected with IgG from patients with PV, PF, or paraneoplastic pemphigus develop cutaneous blisters, histological changes, and immunofluorescence findings consistent with pemphigus.[8,11]

The causes of the autoimmune response remain unknown. An increased frequency of human leukocyte antigen (HLA)-DR4

TABLE 63.1 Immunopathology of Autoimmune Blistering Diseases

Disease	Direct IF (Patient's Skin)	Indirect IF (Patient's Serum)	Antigen(s)
Pemphigus vulgaris	IgG keratinocyte cell surface	IgG binds keratinocyte cell surface (normal stratified squamous epithelium)	Dsg3, +/– Dsg1
Pemphigus foliaceus	IgG keratinocyte cell surface	IgG binds keratinocyte cell surface (normal stratified squamous epithelium)	Dsg1
Paraneoplastic pemphigus	IgG keratinocyte cell surface; C3, granular/linear at epidermal basement membrane (normal stratified and nonstratified squamous epithelium)	IgG binds keratinocyte cell surface	Plakin proteins, Dsg1, Dsg3
Bullous pemphigoid	Linear IgG, C3 at epidermal basement membrane	IgG binds in linear pattern normal stratified squamous epithelium basement membrane (epithelial side)	180-kDa, BPAG2 230-kDa, BPAG1
Herpes gestationis	Linear C3 at epidermal basement membrane (in 30–50% of patients linear IgG also seen)	IgG binds in linear pattern normal stratified squamous epithelium basement membrane (epithelial side)	180-kDa, BPAG2
Linear IgA bullous dermatosis	Linear IgA at epidermal basement membrane	IgA binds in linear pattern normal stratified squamous epithelium basement membrane (epithelial side)	97-kDa protein (portion of BPAG2)
Epidermolysis bullosa acquisita	Linear IgG epidermal basement membrane (IgA rarely)	IgG binds in linear pattern normal stratified squamous epithelium basement membrane (dermal side)	Type VII collagen (noncollagenous domain)
Dermatitis herpetiformis	Granular IgA dermal papillary tips	Negative	Epidermal TGase-3

Dsg, desmoglein; IF, immunofluorescence; Ig, immunoglobulin; kDa, kilodalton; TGase-3, transglutaminase-3.

and HLA-DR14 is found in patients with PV, although the individual susceptibility gene differs with ethnic origin: PV is associated with PV HLA-DRB1*04:02 in Ashkenazi Jews; DRB1*14:01/04 and DQB1*05:03 in non-Jewish patients of European or Asian descent. T cells from patients with PV have been observed to secrete high levels of interleukin-4 (IL-4) and IL-10 in response to specific Dsg3 peptides presented by DRB1*04:02. T cells reactive against Dsg3 are found in patients with PV and also in some normal healthy subjects with HLA DRB1*04:02. Which additional factors are required for HLA-susceptible individuals to develop clinical disease remains unknown.

Both the sporadic and endemic (FS) forms of PF are associated with HLA class II alleles. Susceptibility has been correlated with the presence of DR4, DR14, and DR1 alleles, but no single DR4 or DR14 allele has been associated with the disease. The HLA-DRB1 alleles DRB1*04:04, *14:02, *14:06, or *01:02 have been identified as risk factors for FS (relative risk >14).[12] These alleles all share a common sequence in the third hypervariable domain of the *DRB1* gene.

Environmental factors have been implicated in FS, the endemic form of PF. Healthy people living in endemic areas in South America have anti-Dsg1 IgG antibodies. The proportion of normal individuals with anti-Dsg1 IgG decreases with distance from the endemic focus of FS.[13] In addition, normal controls and patients with FS from endemic areas have IgM against the nonpathogenic extracellular domain of Dsg1. This is uncommon in other pemphigus phenotypes or in patients with FS who have migrated to urban centers, suggesting exposure to an environmental antigen that may share epitopes with Dsg1.[14] An IgG4 response seems important: In endemic areas, both normal subjects and those with preclinical disease or in remission had anti-Dsg1 antibodies that were predominantly IgG1[15]; in those with active disease, anti-Dsg1antibodies were predominantly IgG4.[15]

FS has been linked to bites from bloodsucking black flies found in the endemic areas of Brazil. Pathogenic IgG4 antiDsg1 antibodies from patients with FS cross-react with the LJM11 sand fly salivary gland antigen.[12,16] This suggests that FS may be triggered in HLA-susceptible individuals by an immune response that is cross-reactive with Dsg1.

Although autoantibodies are clearly important in pathogenesis, the mechanism by which they cause acantholysis is unknown. Complement components (C3, C1q) are often present in lesional skin, but autoantibodies can induce acantholysis *in vitro* and in neonatal mice without complement activation. Overall, it seems that complement may augment acantholysis, but pemphigus IgG on its own can bind keratinocytes and induce loss of cell–cell adhesion. It has been suggested that urokinase plasminogen activator is necessary for skin lesions to develop, but pemphigus antibodies remain pathogenic in plasminogen-activator knock-out mice.[17] Acantholysis may be triggered by intracellular signaling after antibodies bind to cell-surface desmogleins: p38MAPK inhibition prevents disease in the mouse model.[18]

Therapy

Before the introduction of glucocorticoid therapy, pemphigus was uniformly fatal. The introduction of high-dose systemic glucocorticoids and improvements in wound care have markedly reduced mortality caused by PV, but it remains a serious condition. Most blistering can be controlled with prednisone (1–2 mg/kg/day, usually in divided doses). The dose can be slowly tapered toward alternate-day therapy. Although this will control disease in many patients, in others, PV may flare up during corticosteroid therapy taper and require additional immunosuppression.

Azathioprine, mycophenolate mofetil, cyclophosphamide, methotrexate, and cyclosporine are the additional immunosuppressive agents most often utilized. Azathioprine and cyclosporine have steroid-sparing effects in long-term treatment, but adding these agents to steroids does not induce long-term remission any better than steroids alone.[19] In small trials, mycophenolate mofetil, used alongside prednisone, induced disease control more quickly compared with prednisone alone. However, a recent larger study of this failed to show statistical significance.[20] Because of its relatively low toxicity, azathioprine (1–3 mg/kg/day) is frequently chosen for adjuvant immunosuppression. Other possible adjunctive therapies include parenteral gold, plasmapheresis, and extracorporeal photochemotherapy. Treatment for PF is similar to that for PV, but therapy can be less aggressive, as PF is usually less severe than PV and rarely fatal. Dapsone is used for intraepidermal IgA neutrophilic dermatosis.[9] Occasionally, systemic glucocorticoid therapy is needed, either alone or with dapsone. Therapy for paraneoplastic pemphigus has been disappointing and usually unsuccessful.

Various biological agents, including etanercept, infliximab, high-dose intravenous immunoglobulin (IVIG), and rituximab, have been tried in pemphigus, with some success. Trials with etanercept showed mixed results with heterogeneous responses. In a trial comparing infliximab plus prednisone versus prednisone alone, neither group achieved their primary endpoints,[21] but at 26 weeks, three of 10 patients given infliximab showed positive responses, compared with none who received prednisone alone. Median anti-Dsg1 and anti-Dsg3 levels were lower in the infliximab group at 18 and 26 weeks. A single course of IVIG was effective in a double-blind study of patients with steroid-refractory PV.[22]

There are numerous case reports of successful use of rituximab in severe or recalcitrant pemphigus. In most cases, rituximab was added to systemic corticosteroid therapy. Three prospective studies of rituximab in pemphigus, alongside corticosteroids or other immunomodulatory agents, have reported dramatic improvement or total clearance in most patients.[23,24] The mean time to achieve disease control, with re-epithelialization of mucosal and cutaneous erosions, was 3 months.[24] The response to rituximab appears durable, with 58% of patients remaining in complete remission 6 years after therapy.[24]

KEY CONCEPTS

Pemphigus

- Clinical presentations correlate with the antigenic target of the autoantibodies.
- Pemphigus foliaceus presents with superficial intraepidermal blisters and antibodies directed against Dsg1 alone.
- Pemphigus vulgaris with only oral mucosal lesions most often has antibodies against Dsg3.
- Pemphigus vulgaris with skin and mucosal lesions most often has antibodies against both Dsg1 and Dsg3

BULLOUS PEMPHIGOID

Clinical Features

Bullous pemphigoid (BP) is characterized clinically by extremely pruritic, tense blisters on inflamed or noninflamed skin. These lesions range from small (1–5 mm) vesicles to large bullae several

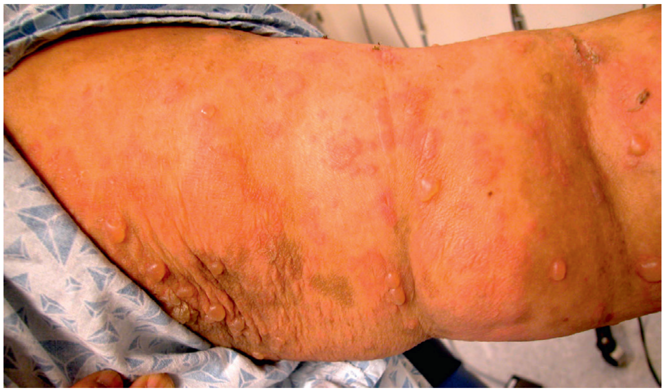


FIG 63.6 Patient with bullous pemphigoid showing erythematous plaques with tense subepidermal bullae.

centimeters in diameter (Fig. 63. 6). Patients with BP can also present with other skin lesions, including urticarial or eczematous dermatitis. Blisters can occur anywhere on skin, but often on skin in the extremities, groin, and axillae.

Oral and ocular mucosal lesions are infrequent in BP, in contrast to patients with mucous membrane pemphigoid (see below), in which mucosal lesions predominate.

Most patients with BP are over 60 years of age, although it can develop at any age. BP has been reported in association with other diseases, including neurological disease, diabetes mellitus, psoriasis, autoimmune diseases, and malignancy. Recent studies have suggested that only neurological disease is actually more prevalent, being over two-fold more common in patients with BP.[25] Drug-related BP is recognized, with furosemide and phenacetin being the most frequent offenders. BP has also been linked to ultraviolet light and therapeutic radiation. In general, the prognosis for BP is good, with a 1-year survival rate of 90% in the United States. Higher mortality rates are reported in Europe, but the reasons for this difference are not currently understood.

Biopsy of an early lesion of BP can show several different patterns consistent with the polymorphic appearance of the eruption. The classic finding is a subepidermal blister with a dermal inflammatory infiltrate, comprised predominantly of eosinophils with some lymphocytes, histiocytes, and neutrophils. The epidermis over this blister is often intact with minimal abnormality, whereas the blister cavity is filled with inflammatory cells. Sometimes neutrophils predominate, or there may be cell-poor lesions with very few inflammatory cells.

Occasionally, there may only be mild epidermal edema, with eosinophil infiltration but no blister. The polymorphic histology of BP means the diagnosis is not based solely on the histological and clinical findings.

The diagnosis of BP can be confirmed by direct immunofluorescence of perilesional skin. Over 90% of patients with BP have linear C3 deposits at the dermal–epidermal junction (DEJ), whereas 70–90% have linear IgG deposits at the DEJ (Fig. 63.7). In some patients, only C3 is seen in the skin. In 70–90% of patients with BP, circulating IgG is present and binds in a linear pattern at the DEJ of normal human skin. Anti-BP180 (BPAG2) antibody titers indicate disease activity, especially at disease onset. Anti-BP230 (BPAG1) antibodies are a fairly sensitive and specific diagnostic marker but do not correlate strongly with disease activity.

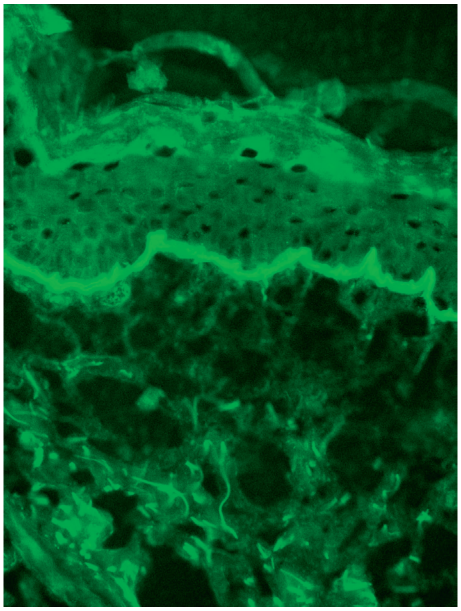


FIG 63.7 Direct immunofluorescence of perilesional normal-appearing skin from a patient with bullous pemphigoid using antibodies directed against human immunoglobulin G (IgG). A linear band of IgG is present at the basement membrane.

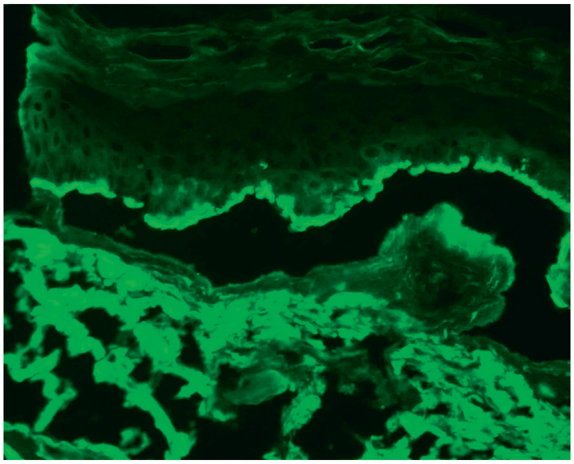


FIG 63.8 Direct immunofluorescence sample from a patient with bullous pemphigoid after incubation with 1 mol/L sodium chloride (NaCl), showing localization of immunoglobulin G (IgG) immunoreactants to the roof (epidermal side) of the blister cavity.

Although the immunofluorescence findings in patients with BP are characteristic, they are not diagnostic. Other blistering diseases, such as bullous lesions in SLE, pemphigoid gestationis (PG), and epidermolysis bullosa acquisita (EBA), can have similar patterns of immunoreactivity. Saline treatment splits skin within the lamina lucida: IgG from patients with BP binds predominantly to the epidermal side, whereas antibodies from patients with EBA bind only to the dermal side. Saline-treated skin from patients with BP also shows that IgG is deposited *in vivo* on the epidermal side of the split skin (Fig. 63.8).

Pathogenesis

Immunofluorescence microscopy demonstrates IgG in the lamina lucida of the basement membrane. The target antigens are components of hemidesmosomes. These are the major attachment sites of epidermal basal cells and where cytoskeletal proteins link through the plasma membrane to the dermis. Two predominant antigens have been identified: 230-kDa and 180-kDa proteins. These have been cloned and sequenced, confirming that they are distinct proteins, associated as a complex within the hemidesmosome. The 230-kDa BP antigen (BPAG1) is an intracellular protein with sequence homology to desmoplakin I, a member of an adhesion junction plaque protein family. The 180-kDa BP antigen (BPAG2) contains both intracellular and extracellular domains joined by a transmembrane region. The extracellular portion of BPAG2 has alternating collagen and noncollagen domains and has been termed *type XVII collagen.* Epitope mapping studies of BPAG2 have identified an immunodominant epitope on an extracellular noncollagen linker domain that is also targeted by IgG antibodies in PG.

Although patients with BP have antibodies against BPAG1 and BPAG2, the pathogenic role of these autoantibodies remains uncertain. Targeted disruption of the gene encoding BPAG1 did not result in blister formation or epidermal–dermal adhesion abnormalities. This suggests that anti-BPAG1 antibodies may not directly inhibit the function of BPAG1 but may maintain or enhance the inflammatory response in BP. Interestingly, targeted knockout of BPAG1 in mice causes progressive deterioration in motor function and sensory neurodegeneration. Additionally, certain portions of BP230 (BPAG1n1 and BPAG1n3) seem to play important roles in organizing cytoskeletal networks *in vivo.*[26] These findings complement prior studies implicating BP230 (BPAG1) in the function of the central nervous system of humans and may explain the overlap between BP and neurological diseases.[27]

Patients with BP (and PG) develop antibodies against a noncollagenous, extracellular portion of BPAG2.[28] Serum levels of BPAG2 autoantibodies correlate with disease activity in patients with BP. Interestingly, patients with an inherited form of skin blistering (generalized atrophic benign epidermolysis bullosa) have a mutation in BPAG2 resulting in a dysfunctional or missing protein.[29]

In an animal model of BP, rabbits were immunized to produce antibodies directed against mouse BPAG2. Passive transfer of this rabbit IgG to newborn mice resulted in blisters, an inflammatory infiltrate, and deposition of immunoreactants. In this model, complement activation, mast cell degranulation, and neutrophil infiltration are important in blister formation.

Complement activation and other extracellular mediators are also implicated in the pathogenesis of BP skin lesions. Complement and the terminal components of the complement cascade are present in the skin of patients with BP. Early lesions in these patients contain eosinophil granule proteins and eosinophil-derived gelatinase, suggesting an early role for eosinophils in the lesion development. Taken together, these findings suggest that IgG autoantibody binds, most likely, to BPAG2 and activates complement and local cytokine production. This attracts eosinophils and neutrophils that release enzymes leading to blister formation. IgE antibodies against BPAG2 have also been found in sera from patients with BP and are implicated in the early lesions of BP, the development of eosinophilic infiltration, and degradation of the basement membrane zone.[30]

The factors that induce development of autoantibodies in patients with BP are unknown. T cells from these patients respond *in vitro* to the extracellular region of BPAG2; this reactivity is restricted by HLA-DQB1*03:01. Subjects with normal HLA-DQB1*03:01 showed similar T-cell reactivity but produced only T-helper-1 (Th1-type) cytokines, whereas T cells from patients with BP produced both Th1-type and Th2-type cytokines. These results suggest that the ability to mount Th2-pattern responses to BP antigen(s) may be critically important.

Therapy

The mainstay of therapy for BP is treatment with systemic glucocorticoids: most patients respond rapidly to oral prednisone (0.5–2.0 mg/kg/day), but starting prednisone at doses >0.75 mg/kg/day confers no advantage. Mild or localized disease can sometimes be managed with potent topical corticosteroids (confirmed in controlled trials).[31]

When new blister formation has stopped and healing has begun, systemic steroids can be tapered slowly. The speed of tapering is dictated by the severity of the patient's initial disease and any flare-ups that occur during tapering. Most patients can stop systemic steroids completely within 6–18 months, but recurrence of disease activity is not uncommon. Flare-ups should be managed with the lowest dose of systemic steroids possible; occasionally topical corticosteroids are enough. Usually, BP is self-limiting, lasts between 1.5 and 5 years, and responds promptly to systemic glucocorticoids.[31]

A minority of patients requires prolonged high-dose systemic glucocorticoids. In these individuals, adding azathioprine, cyclophosphamide, or methotrexate will often allow tapering or discontinuation of systemic steroids, but data on which is best for steroid sparing are limited.[31] Other unproven adjunctive therapies that may help in some patients include dapsone, cyclosporine, and rituximab. Our preferred options for additional therapy are azathioprine (1–2 mg/kg/day) or mycophenolate mofetil (1000–2000 mg/day).[32]

A recent review of IVIG for BP concluded that 70% of patients experienced some improvement but there was no clinical benefit in the remainder. IVIG 2 g/kg over 5 days at monthly intervals for 3 months has been commonly used; typically more than one cycle is needed to prevent recurrence. In some patients, IVIG appears to be steroid sparing. IVIG was ineffective if patients received low-dose IVIG or only a single infusion.

Rituximab eliminates $CD20^+$ B cells through complement-dependent and antibody-dependent cell-mediated cytotoxicity, as well as by inducing structural changes and apoptosis. The targeted B cells remain absent from the circulation for 6–12 months. Several groups have reported clinical success with the use of rituximab for pemphigus, but it has only had limited success in the treatment of BP.[33]

An open-label trial of anti-IgE (omalizumab) demonstrated benefit in 5 of 6 patients with BP,[34] albeit with varying degrees of success.

KEY CONCEPTS

Bullous Pemphigoid

- Most common autoimmune blistering disease that presents in older adults
- Severely pruritic with clinical presentations, including urticarial lesions, small vesicles, and/or large tense vesicles
- Linear immunoglobulin G (IgG) and C3 present at the epidermal basement membrane, binding to the epidermal side of "1 M NaCl split" skin
- Disease activity correlates best with IgG anti-BP 180 antibodies

EPIDERMOLYSIS BULLOSA ACQUISITA

EBA is a chronic subepidermal blistering disease that typically presents in the fourth to sixth decades. There are two forms of EBA—a noninflammatory type with blisters on distal extremities and an inflammatory type that closely resembles BP. Patients with non-inflammatory EBA have peripherally distributed blisters that heal with scarring and milia formation. Their skin is extremely fragile, often resulting in numerous erosions in areas of mechanical trauma, such as hands, feet, elbows, and knees. Lesions are often seen on oral mucous membranes, sometimes including the esophagus. Patients with classic EBA may also develop ocular, vaginal, urethral, and rectal mucosal lesions. Ocular changes are common, clinically resembling mucous membrane pemphigoid (MMP). Other cutaneous manifestations include scarring alopecia and variable degrees of nail dystrophy. Patients with nonclassic (inflammatory) EBA often present similarly to those with BP, with widespread tense bullae on an erythematous base, which heal without scarring (Fig. 63.9).[35]

EBA has been associated with several diseases, particularly inflammatory bowel disease (IBD) and bullous SLE. These associations may partly result from the association of EBA with HLA-DR2.[36]

Skin biopsies of early lesions from patients with EBA show subepidermal blisters with variable degrees of inflammation. In patients with classic EBA, lesional skin biopsies often have minimal inflammatory cell infiltrate. In contrast, patients with inflammatory EBA may have substantial collections of mononuclear cells, neutrophils, and eosinophils in the superficial dermis.

Direct immunofluorescence of perilesional skin biopsies from patients with EBA shows linear deposits of IgG at the DEJ, as in BP. Linear deposits of C3, IgM, IgA, and fibrinogen have also been reported.[35] However, in EBA, these deposits are localized exclusively below the lamina lucida. Direct immunofluorescence of saline-split perilesional skin shows IgG in the blister floor in EBA (Fig. 63.10); in BP, IgG is in the blister roof.

Indirect immunofluorescence using normal human skin shows circulating anti–basement membrane zone antibodies in 30–50% of patients with EBA. However, IgG antibody can be detected in 85% of patients when using saline-split skin, which is the more sensitive and specific substrate. As expected from the *in vivo* deposition pattern, IgG from patients with EBA localizes to the blister floor. Sera from these patients recognize the 300-kDa protein, type VII collagen, primarily targeting its immunodominant NC1 domain.

Pathogenesis

Immunoelectron microscopy of EBA biopsy specimens has shown immunoglobulin deposits localized to the lamina densa zone of the basement membrane, below the lamina lucida. Transmission electron microscopy of lesional EBA skin shows decreased or absent anchoring fibrils. Anchoring fibrils are implicated in epidermal–dermal adherence via linkage of the hemidesmosome through the basement membrane, and their absence may explain the observed skin fragility. The lack of inflammatory infiltrate in many patients with EBA suggests that autoantibodies may disrupt the interaction between anchoring fibrils and dermal matrix proteins.

The target antigen for the IgG autoantibodies present in the sera of patients with EBA is type VII collagen, a 300-kDa glycoprotein composed of a 145-kDa noncollagenous domain (NC1) at the amino-terminal end, an 18-kDa noncollagenous domain (NC2) at the carboxy terminus, and a central collagenous domain. IgG antibodies from patients with EBA appear specific for epitopes within the NC1 noncollagenous domain. Type VII collagen is the major structural component of anchoring fibrils and is produced by both epithelial keratinocytes and dermal fibroblasts. Passive transfer experiments in mice and active immune models

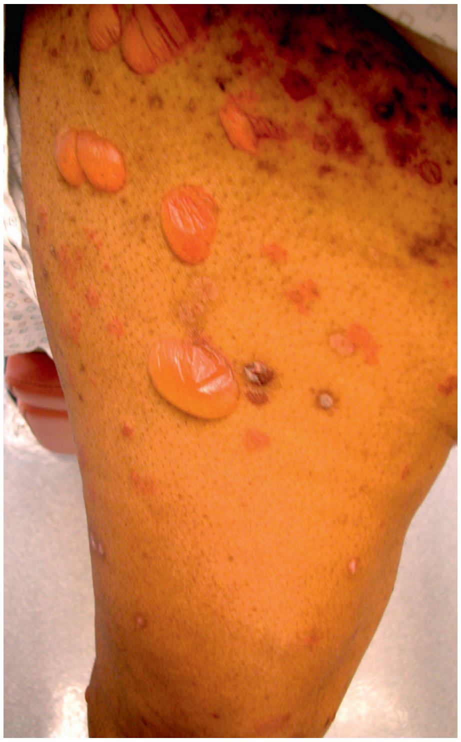


FIG 63.9 Patient with epidermolysis bullosa acquisita showing extremity involvement with tense bullae. Note the similarity to lesions of bullous pemphigoid, with somewhat less inflammation surrounding the base of bullae.

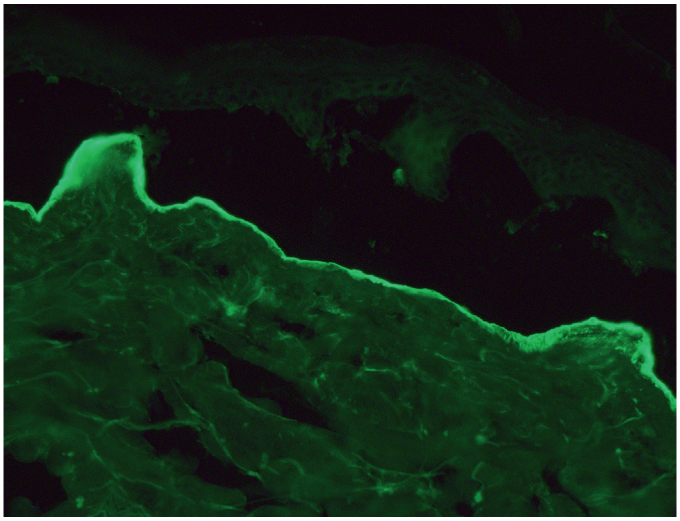


FIG 63.10 Direct immunofluorescence sample from a patient with epidermolysis bullosa acquisita (EBA) after incubation with 1 mol/L sodium chloride (NaCl), showing localization of immunoglobulin G (IgG) immunoreactants to the floor (dermal side) of the blister cavity.

have provided additional support for the critical role of anti-type VII collagen antibodies in the pathogenesis of EBA.[35] In experimental mouse-model EBA, induction of specific types of antibodies is linked to the MHC haplotype suggesting future work may identify non-MHC EBA susceptibility genes.[37] Utilizing a four-way, autoimmune-prone, advanced mouse intercross line immunized with a COL7 fragment to induce EBA, investigators identified quantitative trait loci (QTLs) on chromosomes 9, 12, 14, and 19 associated with disease development and QTLs on chromosomes 1, 15, and 19 associated with maximum disease severity.[4] In this model, gene–microbiota interactions appear to promote disease development.[38] Finally, identification of other mediators in inflammatory EBA, such as heat shock protein-90 (HSP-90), granulocyte macrophage–colony-stimulating factor (GM-CSF), chemokine ligand-1 (CXCL1), CXCL2, or IL-1, suggests other possible candidates for therapeutic targeting.[38]

Treatment

Spontaneous resolution is infrequent in EBA, and management is difficult. The main goals of therapy are to minimize blistering and scar formation, with particular attention to ocular and oral mucosal lesions. Systemic glucocorticoids are the mainstay of therapy, especially for patients with inflammatory EBA. Unfortunately, even high-dose systemic steroids do not usually improve skin fragility and trauma-induced blister formation. Mucosal lesions often respond to systemic steroids (prednisone 0.5–1.5 mg/kg/day) but may recur with tapering and/or discontinuation of steroids. Several adjunctive agents, including azathioprine, cyclophosphamide, colchicine, dapsone, hydroxychloroquine, and plasmapheresis, have been proposed, but none has been consistently effective. Cyclosporine has been used to treat patients with EBA with some success, although toxicity can limit therapy. Photopheresis, IVIG, and rituximab have also been reported to be effective in some patients with refractory EBA.[35] Mucous membrane lesions can prove particularly difficult to manage and may require systemic therapy. Patients with ocular lesions need ophthalmologist review and may require systemic glucocorticoids to prevent conjunctival scarring. Oral mucous membrane lesions can sometimes be managed with frequent application of potent topical steroid ointments or gels (0.05% clobetasol propionate, 0.05% fluocinonide). If this fails and the degree of oral erosions inhibits appropriate nutrition, systemic glucocorticoid therapy may be required. Clinicians should also be aware of possible involvement of esophageal mucosa and/or tracheal mucosa and involve appropriate specialists to monitor and treat these potentially severe complications.

As well as therapies targeting B cells, novel therapeutic targets addressing autoreactive T cells are being studied in the treatment of EBA. These include inhibitors of the stress-inducible HSP-90.[39] Strategies targeting other features of EBA, such as neutrophil recruitment and complement activation, are also being explored.

The management of chronic skin wounds is vital in EBA. Protection of skin from trauma and early use of topical and systemic antibiotics are critical to improving the rate of healing. The development of new biological dressings for chronic ulcers has also proven helpful in the management of these wounds.

PEMPHIGOID GESTATIONIS

PG is a rare, itchy, blistering disease of pregnancy and the puerperium characterized by linear deposits of IgG and C3 at the DEJ. Formerly known as "herpes gestationis," PG has no relationship to herpes virus infections. The term *pemphigoid gestationis* eliminates this confusion and emphasizes the pathophysiological similarity of PG to BP.

PG is rare, occurring in <1 in 50 000 pregnancies in North America. It usually presents in the third trimester or in the immediate postpartum period. Disease flare-ups beyond the immediate postpartum period do occur, and patients have been reported to develop blisters when menses return or when oral contraceptives are used. Although PG may recur in subsequent pregnancies, this is not absolute.

PG presents as tense blisters, often on an urticarial base, ranging in size from small (3–5 mm) vesicles to large (1–2 cm) bullae. The blisters often begin on the abdomen, although the entire body can be involved. Pruritus is common and can be severe.

PG is associated with maternal and fetal morbidity. In the mother, morbidity mainly relates to skin disease, with extensive itching and blister formation. Recent data indicate a better prognosis for the fetus than previously reported, with a 20% risk for premature delivery. Small-for-gestational-age birth weight and spontaneous abortion are only marginally increased in frequency. Similar effects have been found in pregnant women receiving systemic corticosteroid treatment for allergy, asthma, inflammatory bowel disease, and so on. Low birth weights and prematurity are significantly associated with early onset of PG (in first or second trimester), rather than exposure to corticosteroids during pregnancy.[40] Up to 10% of children born to mothers with PG develop skin lesions similar to those in the mother, likely as a result of transplacental crossing of maternal autoantibodies. The typical presentation is an erythematous urticarial or vesicular rash; yellowish plaques on erythematous base and frank bullae have also been reported. The disease is generally mild and resolves spontaneously over days to weeks as maternal antibodies are cleared.

The histopathology of PG lesions is not diagnostic. Subepidermal blisters are seen with necrotic basal keratinocytes and a perivascular infiltrate containing eosinophils, neutrophils, lymphocytes, and monocytes.

Direct immunofluorescence studies of the skin of patients with PG shows linear deposition of C3 at the DEJ. In addition, 30–50% of patients have linear IgG deposits in a similar pattern. Occasionally IgA, IgM, C1q, C4, properdin, factor B, and C5 may be present in the lamina lucida. Indirect immunofluorescence reveals circulating IgG antibodies directed against the epidermal basement membrane in only ~30% of patients. In contrast, 50–75% of patients have IgG antibody directed against epithelial basement membrane as shown by complement fixation techniques. This IgG anti–basement membrane antibody fixes complement *in vitro* but is not detected by routine indirect immunofluorescence because of its low concentration. These autoantibodies can cross the placenta: Infants born to mothers with PG often have C3 deposits at the DEJ, but without any clinical skin disease. These IgG antibodies bind the epithelial side of saline-split skin, as seen in BP.

Pathogenesis

Sera from most patients with PG bind to BPAG2 but do not react with BPAG1. The role of pregnancy and other hormonal factors in the development of PG is not understood.

Patients with PG have an increased frequency of HLA-DR3; the greatest relative risk is when HLA-DR3 and HLA-DR4 are both present. Patients with PG have an increased frequency of anti-HLA antibodies compared with multiparous control subjects. Some studies have suggested that paternal haplotype is important in PG pathogenesis, but further prospective studies are needed. Patients with PG also have increased frequency of the C4 null allele, although this may just secondary to the increased frequency of HLA-DR3.[41]

Treatment

Therapy of PG is directed at relieving pruritus and preventing the development of new blisters. If the extent of disease is limited, patients can be managed with topical glucocorticoids and antihistamines. However, most patients with PG have extensive disease and need systemic glucocorticoids. Prednisone 20–60 mg/day is usually sufficient to control both pruritus and new blister formation; once control has been achieved, the dosage should be reduced to the lowest possible effective dose. Occasionally, it may be possible to stop prednisone before delivery. However, PG often flares up in the immediate postpartum period. Only rarely do patients require therapy beyond the initial postpartum period. In patients who do require prolonged therapy after delivery, immunosuppressive agents should be considered, as in the case of the other blistering diseases. Patients being treated with systemic steroids or immunosuppressive drugs should be counseled regarding breastfeeding restrictions.

MUCOUS MEMBRANE PEMPHIGOID

MMP, formerly termed *cicatricial pemphigoid,* is a rare, chronic, scarring blistering disease that predominantly affects the oral and conjunctival mucosae. Conjunctival involvement ranges from mild conjunctivitis to severe inflammation leading to symblepharon formation, entropion, trichiasis, corneal scarring, and, in some cases, to blindness. Oral lesions include gingivitis, often with erosions and desquamation, as well as erosions and ulcerations on the buccal and palatal mucosa. Esophageal and laryngeal involvement is rare but can cause scarring and stenosis. MMP can also affect the genitalia, rectum, and nasopharynx. Cutaneous involvement occurs in 10–25% of patients, although it is often limited in extent and predominantly on the scalp, back, or face.[42]

Mucosal biopsy reveals subepidermal blisters with a cellular infiltrate, usually consisting of lymphocytes and plasma cells. Increased numbers of neutrophils and/or eosinophils have also been reported. Skin biopsy specimens of lesions in patients with MMP are indistinguishable from those of BP.

Direct immunofluorescence of perilesional skin or mucosal samples reveals linear deposition of immunoreactants at the basement membrane zone. In most patients with MMP, these are IgG and C3, although IgA, IgM, and fibrin deposits have also been reported. In patients who have ocular or oral mucosal MMP but no skin lesions, undirected skin biopsy for immunofluorescence is unhelpful. Similarly, conjunctival biopsy can yield negative immunofluorescence in patients with otherwise typical MMP. This may be attributed to conjunctival inflammation. Careful evaluation for mucosal lesions and biopsy of perilesional mucosae is often diagnostic. On immunoelectron microscopy, immunoreactants have been found in the lamina lucida and the lamina densa.

Circulating anti–basement membrane zone antibodies may be present in low titers but are usually absent in patients with MMP. The diagnostic yield of indirect immunofluorescence is not improved by using different tissue substrates, including oral and conjunctival mucosae, normal human skin, and saline-split normal human skin. A BP180 C-terminal domain enzyme-linked immunosorbent assay (ELISA) has been evaluated as a method of characterizing MMP, although direct immunofluorescence remains the most sensitive test for diagnosing BP-180 MMP.[43]

Pathogenesis

Understanding the pathogenesis of MMP has been hampered by the heterogeneous clinical presentation and the relative lack of antibodies in serum. The presence of different immunoglobulin classes at the basement membrane on immunofluorescence and the variable localization of these immunoglobulins on immunoelectron microscopy support the hypothesis that MMP is a heterogeneous disease, in which a variety of basement membrane zone proteins are targeted. In a group of patients with immune-mediated subepithelial blistering disorders of the mucous membranes, immunoreactants were found on direct immunofluorescence in about 90% with both skin and mucous membrane lesions or with mucous membrane lesions alone.[44] In contrast, only 40% of patients with isolated ocular disease had C3 and/or IgG deposits. This may reflect the difficulty encountered in biopsy of inflamed conjunctiva. On indirect immunofluorescence, 11 of 14 (81%) patients with skin and mucous membrane disease had circulating IgG anti–basement membrane zone antibodies, of which nine were identified as having BPAG1 and BPAG2 by immunoblot and/or immunoprecipitation. In contrast, only 3 of 24 patients with only oral mucosal lesions or ocular lesions had positive results on indirect immunofluorescence, and one of the sera reacted with the BP antigens.

Sera from some patients with MMP react with BPAG2, whereas others have antibodies against type VII collagen. Sera from patients with ocular MMP react with a 205-kDa protein that has homology with the cytoplasmic domain of β_4 integrin and a 45-kDa protein. Some patients have antibodies against epiligrin (laminin 5), a ligand for major keratinocyte integrins ($\alpha_3\beta_1$ and $\alpha_6\beta_4$). Patients with MMP who have anti-epiligrin antibodies need a careful workup, as they have an increased risk of malignancy.[45] Additionally, >80% of patients with anti-epiligrin antibodies show laryngeal involvement, something seen in <10% of all patients with MMP.[46] These data suggest that the clinical presentation of MMP is associated with diverse autoantibodies that react with multiple antigens of the skin basement membrane zone. The different clinical presentations may therefore result not only from different target antigens but also perhaps from different antibody-binding sites on these antigens.

MMP has been linked to many different HLA class II alleles. It has been suggested that HLA-DR4 substantially increases the risk of ocular disease. An increased prevalence of HLA-DQB1*03:01 was described in patients with isolated ocular MMP,[46a] but this allele was later found to be associated with all clinical sites of involvement and possibly to anti–basement membrane IgG production as well as overall disease severity.

Therapy

Treatment of MMP is difficult and directed at minimizing the scarring and resultant ocular complications. Dapsone (100–200 mg/day) is effective at controlling the early inflammatory response. Tetracyclines can also control MMP and are often combined with nicotinamide. If these are unsuccessful, treatment with systemic glucocorticoids (1–2 mg/kg/day) may be needed, often with adjunctive immunosuppressive therapy, such as with azathioprine, mycophenolate mofetil, or cyclophosphamide. IVIG may be useful: It starts to work rapidly and is especially helpful in patients with progressive ocular MMP whose vision is threatened. In some patients, IVIG has arrested the progression of eye disease, maintaining vision and preventing blindness. However, no controlled studies have yet been performed. Case reports suggest some benefit with plasmapheresis, etanercept, and infliximab. Growing numbers of case series have suggested that rituximab may effective, with most patients showing complete response after a single course, but, again, no controlled trials have been published.[47] A retrospective study of MMP patients given rituximab showed that all patients achieved disease control, compared with 40% of those who received conventional therapy.[48]

The clinical course of MMP is difficult to predict, and the response to treatment is often suboptimal. Some patients achieve long-lasting remissions, but MMP frequently recurs, and patients in remission should be monitored closely.

Ocular mucosal lesions are often severe in MMP and can cause significant morbidity. Consultation with an ophthalmologist is important in managing the potentially severe ocular complications. Physicians should also be aware of the risk of esophageal and tracheal involvement; patients with severe disease often need consultation with relevant specialists. Oral mucosal lesions can often be managed with frequent application of potent topical corticosteroids (0.05% clobetasol propionate or 0.05% fluocinonide). In severe cases, dapsone or systemic glucocorticoids may also be needed. Patients with limited skin disease can often be managed with topical therapy alone. If this fails, treatment as for BP is usually successful.

LINEAR IGA BULLOUS DISEASE

Linear IgA bullous disease (LABD) is a clinically heterogeneous blistering disease in which direct immunofluorescence of perilesional skin biopsies reveals linear deposition of IgA at the basement membrane zone. The majority of patients with LABD presents with pruritic vesicles and papules, localized primarily to the extensor surfaces, similar to the clinical pattern seen in patients with dermatitis herpetiformis (DH). Patients with LABD do not, however, have associated gluten-sensitive enteropathy, their disease is not improved on a gluten-free diet, and they do not have the characteristic HLA associations of patients with DH.

Other patients with linear IgA deposits present a clinical picture more suggestive of BP, with pruritic tense blisters on an erythematous base. Other clinical presentations have been described, with clinical, histological, and immunopathological overlap among patients with LABD, MMP, and EBA. Children can develop a subepidermal blistering disease with linear IgA basement membrane deposits, a condition previously termed *chronic bullous disease of childhood.* In these juvenile cases, blisters occur mainly in flexural surfaces, around the genitalia, and on the face, especially the perioral region. A drug-induced form of LABD can occur, mainly in association with vancomycin, but also with other medications.[49]

Histopathology of lesional skin of patients with LABD reflects the clinical heterogeneity seen in these patients. Biopsy usually reveals collections of neutrophils in the dermal papillary tips in a pattern virtually identical to DH. However, subepidermal blisters with eosinophils may also be seen, as in BP.

Direct immunofluorescence of perilesional skin in patients with LABD reveals a linear band of IgA, almost exclusively IgA1, at the basement membrane zone. These IgA skin deposits contain both κ and λ light chains, indicating that the antibodies are polyclonal. J chain is not present, suggesting that the IgA is monomeric and not of gut origin. Immunoelectron microscopy most often reveals IgA deposits in the lamina lucida; but IgA deposits have also been seen below the lamina densa. Circulating IgA antibodies against the basement membrane zone of stratified squamous epithelium are detected in only 10–30% of both adult and pediatric patients with linear IgA deposits. Other immunoreactants, especially IgG and C3, have been found in 30–40% of patients.

Pathogenesis

The paucity and heterogeneity of circulating antibody has made it difficult to determine the antigenic target in LABD. Sera from adults and children with linear IgA deposits within the lamina lucida reacted with a 97-kDa protein that is identical to part of BPAG2.[50] In patients with LABD or linear IgA disease of childhood with IgA deposits below the lamina densa, antibodies have been identified reactive to type VII collagen and an unknown 285-kDa dermal protein.[50] These findings support the hypothesis that there are heterogeneous target antigens in LABD and linear IgA disease of childhood. Regardless of the location of IgA deposits or the specific antigenic target, the mechanism of lesion formation in patients with LABD remains unknown.

Therapy

Dapsone (25–200 mg/day) is the mainstay of therapy of patients with all forms of LABD. Although dapsone is well tolerated by most patients, it has significant pharmacological and idiosyncratic adverse effects. Occasionally, patients do not respond to dapsone alone. In these cases, adding low doses of prednisone (10–20 mg/day) may result in significant improvement.

THERAPEUTIC PRINCIPLES

Treatment of Autoimmune Blistering Diseases

Systemic corticosteroids are the initial mainstay of treatment.

Dapsone is the mainstay of therapy for linear immunoglobulin A (IgA) dermatosis and dermatitis herpetiformis.

Steroid-sparing agents, including azathioprine, mycophenolate mofetil, methotrexate, and intravenous immunoglobulin (IVIG), are often needed for treatment.

Rituximab and other anti-CD20 biologics have been proven to be very effective as steroid-sparing agents in clinical studies.

DERMATITIS HERPETIFORMIS

Clinical Features

DH is an intensely pruritic blistering disease that typically presents in the second or third decade with erythematous papules and/or vesicles over extensor surfaces (Fig. 63.11). Patients often present with a broad spectrum of lesions. The severe pruritus leads to scratching, resulting in erosions that may be the presenting feature. Patients can also present with urticarial plaques or, less commonly, frank bullae. Lesions are usually symmetrically distributed over the extensor surfaces, especially the elbows, knees, buttocks, back, and posterior hairline. Symptomatic mucous membrane lesions are rarely present. Extreme pruritus is a clinical hallmark of DH. Patients characteristically report severe burning or stinging 12–24 hours before the appearance of lesions, which persists until vesicles break and crusts form. DH is generally considered a life-long dermatosis, although clinical remission can occur in as many as 12% of patients.[51] The wide variety of clinical presentations of DH often suggests a long differential

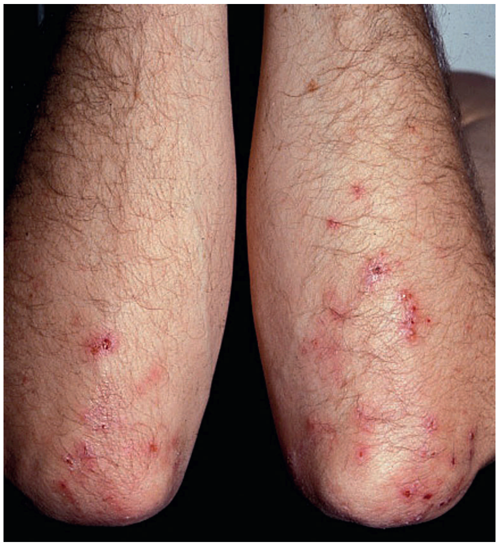


FIG 63.11 Patient with dermatitis herpetiformis showing erythematous papules with crusts on the elbows. Rare intact vesicles are present.

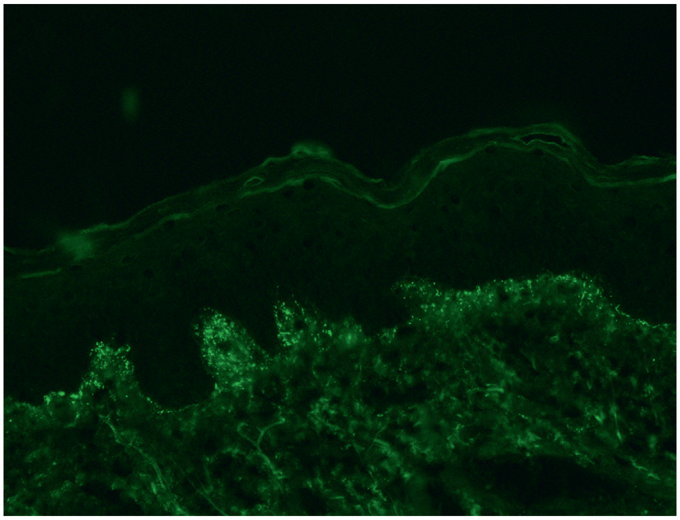


FIG 63.12 Direct immunofluorescence of normal-appearing perilesional skin of a patient with dermatitis herpetiformis using antibodies against human immunoglobulin A (IgA). Granular deposits of IgA are present at the dermal–epidermal junction.

diagnosis, including erythema multiforme, PG, BP, transient acantholytic dermatitis, papular urticaria, scabies, bug bites, and neurotic excoriations.

The frequency of DH varies in different ethnic groups. It is between 10 and 39 per 100 000 in Anglo-Saxon and Scandinavian populations, but up to 75 per 100 000 in Finland, although its incidence in the UK appears to be decreasing.[52,52a] DH occurs much less frequently in other populations (*e.g.,* those of Asian or African descent). This relates, in part, to differing frequencies of the DH-associated HLA antigens in different populations, as detailed below.

Biopsy of early DH lesions reveals neutrophilic infiltration of dermal papillae with fibrin, neutrophilic fragments, edema, and variable numbers of eosinophils. This histological pattern is not specific and has been reported in BP, linear IgA disease, bullous eruption in SLE, and leukocytoclastic vasculitis.

Granular IgA deposits can be found at the DEJ on direct immunofluorescence of DH skin biopsy specimens (Fig. 63.12). These granular deposits of IgA at the DEJ are specific for DH and have not been found in gluten-sensitive enteropathy (GSE, also known as celiac disease; Chapter 75) without DH or in asymptomatic relatives of patients with DH.[53] Although some patients with DH have circulating antibodies against tissue transglutaminase, identification of granular IgA deposits at the DEJ remains the gold standard for diagnosing DH. To maximize the diagnostic yield, biopsies for direct immunofluorescence studies should be taken from normal-appearing perilesional skin.

A gastrointestinal (GI) abnormality similar to that seen in isolated GSE was identified in 60–70% of DH patients.[54] This abnormality was reversible by avoiding dietary gluten, confirming that DH was a gluten-sensitive disease. Small bowel histological changes in DH include flattening of the normal villous architecture of the jejunal epithelium, with elongation of intestinal crypts and mononuclear cell infiltrates within the lamina propria and epithelium. These findings are often patchy and less severe than in isolated GSE. In effect, all patients with clinical DH and granular IgA deposits have associated GSE.

Most patients with DH have no GI symptoms. Only 10% of DH patients have typical symptoms of isolated GSE, such as bloating, diarrhea and malabsorption, while 20–30% have mild steatorrhea. However, many patients with DH have asymptomatic abnormal intestinal function, as shown by abnormal absorption of D-xylose, iron, folate, glucose, water, and bicarbonate.

GSE clearly plays a critical role in the pathogenesis of DH. Patients with DH who adhere to gluten-free diets can control their skin disease, normalize the morphological changes of the small intestine and, after years of gluten avoidance, lose their cutaneous IgA deposits. However, the exact relationship between the skin disease, cutaneous IgA deposits and GSE remains unknown.

HLA-DR3 is expressed in 90–95% of patients with DH, compared with ~23% controls. In addition, 95–100% of patients with DH express HLA-DQ2 (which is in linkage disequilibrium with HLA-B8 and HLA-DR3), compared with 40% of controls. The HLA-DR and HLA-DQ alleles involved in the pathogenesis of DH are DQB1*02:01, DQA1*0501, and DRB1*03:01, which are essentially identical to those associated with isolated GSE.

Patients with DH also have an increased frequency of gastric atrophy and gastric hypochlorhydria. Thyroid abnormalities, including hypothyroidism, hyperthyroidism, thyroid nodules, and thyroid cancer, also occur more frequently in these patients. Various other autoimmune diseases, including SLE, dermatomyositis, myasthenia gravis, Sjögren syndrome, and rheumatoid arthritis, have also been reported in patients with DH. It is thought that many of these associations relate to the high frequency of the HLA-B8, HLA-DR3, and HLA-DQ2 haplotypes in these patients. A five- to sixfold increased incidence of non-Hodgkin lymphoma has been reported among patients with DH, without any increased mortality.[55,56]

Pathogenesis

The pathogenesis of DH rests on three distinct associations: *(i)* granular deposits of IgA in the skin at the DEJ; *(ii)* gluten-sensitive enteropathy (albeit often asymptomatic); and *(iii)* increased frequency of the HLA-DR3/DQ2 haplotype. Attempts to understand DH must integrate all these factors.

The mechanism of IgA binding to DH skin is uncertain. Circulating antibodies that bind to normal human skin *in vitro* have not been detected in DH. Also, the IgA deposits do not appear to be caused by IgA-containing circulating immune complexes. Another mechanism by which IgA might bind to DH skin is that IgA produced in the gut, in response to wheat or other dietary antigens, may reach the circulation, where it could bind to wheat protein deposited in skin or cross-react with normal skin structures or molecules. DH sera contain IgA antibodies against endomysium, a connective tissue element surrounding smooth muscle, and are directed against tissue transglutaminase.[57] Patients with DH also have IgA antibodies against epidermal transglutaminase, with higher avidity compared with those found in patients with isolated GSE. The IgA deposits in DH skin colocalize with epidermal transglutaminase, suggesting epidermal transglutaminase-3 (eTG3) is the IgA target in the skin of patients with DH. Development of IgA anti-eTG3 antibodies is associated with prolonged periods of gluten exposure, suggesting that these antibodies result from epitope spreading. In a mouse model in which normal human skin is grafted onto SCID mice, human IgA anti-eTG3 antibody binds in a pattern identical to that seen in patients with DH. These observations suggest that IgA anti-eTG binds and slowly accumulates in the dermis, resulting in the pathognomonic IgA deposits seen in patients with DH.[58]

The presence of an ongoing mucosal immune response in patients with DH has been shown to result in systemic signs of inflammation, including elevated serum levels of soluble IL-2 receptor and IL-8, increased neutrophil expression of CD11b, and increased expression of endothelial cell E-selectin in skin.[59] These findings suggest that the mucosal immune response results in a proinflammatory environment in skin, which, when coupled with cutaneous IgA deposits, can result in development of typical DH skin lesions.

Therapy

DH can be treated either with dapsone or a gluten-free diet; the choice of therapy should be individualized to each patient. The treatment of choice for most patients with DH is dapsone or sulfapyridine because of the lack of GI symptoms in most cases and difficulty adhering to a strict gluten-free diet. Dapsone (100–200 mg/day) is sufficient to control cutaneous eruptions in most cases, with minimal side effects. Dapsone therapy results in almost immediate cessation of DH skin symptoms, but it does not affect the GI mucosal defect or symptoms. Patients with DH who are on dapsone therefore have ongoing (albeit low-grade) GSE. Dapsone is potentially toxic, and if the disease flares up, patients should not increase the dose without consulting their physician.

Strict adherence to a gluten-free diet will also control DH skin lesions. Nearly 80% of patients with DH who follow such a diet can stop or substantially reduce the dose of dapsone required to control their skin disease.[59] Patients may need to be on a gluten-free diet for at least 5 months before being able to reduce their dose of dapsone; being on this diet for 8–48 months is necessary before discontinuation of dapsone. Gluten-reduced diets, in which gluten content is not totally eliminated, are less effective in controlling DH skin lesions but may provide some relief. Patients on gluten-reduced diets are able to decrease their dapsone dosage by approximately 50% compared with ≥75% reduction in patients on gluten-free diets. Strict adherence to a gluten-free diet controls the cutaneous manifestations of DH in most patients, reverses the morphological changes in the small intestine and, after many years, may result in disappearance of the cutaneous IgA deposits. Cutaneous IgA has been shown to return after rechallenge with dietary gluten.[59]

Rarely, patients with DH fail to respond to a gluten-free diet despite what appears to be total dietary compliance. This has also been reported in patients with isolated GSE and other dietary factors have been postulated to play a role. In support of this concept, the skin lesions and GI changes of DH may also respond to an elemental diet, often within only 2–3 weeks of starting the diet. These results imply that other proteins besides gluten play a role in the DH pathogenesis. Although a gluten-free diet is an attractive alternative to medication for many patients, it is difficult to maintain. Gluten is present in most common grains (wheat, rye, barley) but not in rice and maize. Eating oats appears to be safe for patients with DH, but oat products may be contaminated with wheat protein. Patients should be informed that the success or failure of the diet cannot generally be assessed until they have been on the diet for at least a year.

TRANSLATIONAL RESEARCH

ON THE HORIZON

Precise characterization of the antigenic targets of autoantibodies and correlation with clinical presentations and outcomes

Identification of markers or disease activity and response to therapy

Specific immunoabsorption therapy with faster responses and less treatment-associated morbidity

Increased use of targeted biologic therapy (*e.g.*, anti–B-cell therapy) to provide rapid, long-lasting remission while minimizing morbidity and mortality.

The next 5–10 years should see further advances in the treatment of autoimmune blistering skin diseases. Anti–B-cell therapies, such as rituximab, offer a more targeted, steroid-sparing approach, and we can expect to see the development of more efficient ways of delivering these therapies. Translational studies of immunological changes among patients successfully treated with anti–B-cell therapies will yield insights into the role of different B-cell populations in the pathogenesis of autoimmune blistering disease. We can expect to see more precise determination of the antigenic targets of the autoantibodies and correlation with clinical presentations, response to therapy, and clinical outcomes.

An improved understanding of the clinical response to biologics and the mechanisms of those responses will lead to discovery of new combinations and earlier initiation of biological therapy. Early or immediate use of specific biological therapies, such as rituximab, may allow shorter courses of systemic corticosteroid therapy, long-lasting remissions, and potentially lead to "cures" of these severe autoimmune diseases.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Langan SM, Smeeth L, Hubbard R, et al. Bullous pemphigoid and pemphigus vulgaris—incidence and mortality in the UK: population based cohort study. BMJ 2008;337:160–3.

2. Joly P, Litrowski N. Pemphigus group (vulgaris, vegetans, foliaceus, herpetiformis, brasiliensis). Clin Dermatol 2011;29:432–6.
3. Bolognia JL, Jean L, Jorizzo JL, et al. Pemphigus. In: Bologina JL, Jean L, Jorizzo JL, et al, editors. Dermatology. Cambridge, UK: Elsevier Health Sciences; 2012. 29 p. 461–74.
4. Risser J, Lewis K, Weinstock MA. Mortality of bullous skin disorders from 1979 through 2002 in the United States. Arch Dermatol 2009;145:1005–8.
5. Beutner EH, Jordon RE. Demonstration of skin antibodies in the sera of pemphigus vulgaris patients by indirect immunofluorescent staining. Proc Soc Exp Biol Med 1964;117:505–10.
6. Aoki V, Sousa JX Jr, Diaz LA, et al. Pathogenesis of endemic pemphigus foliaceus. Dermatol Clin 2011;29:413–18, viii.
7. Maruani A, Machet MC, Carlotti A, et al. Immunostaining with antibodies to desmoglein provides the diagnosis of drug-induced pemphigus and allows prediction of outcome. Am J Clin Pathol 2008;130:369–74.
8. Anhalt GJ. Paraneoplastic pemphigus. J Investig Dermatol Symp Proc 2004;9:29–33.
9. Düker I, Schaller J, Rose C, et al. Subcorneal pustular dermatosis-type IgA pemphigus with autoantibodies to desmocollins 1, 2, and 3. Arch Dermatol 2009;145:1159–62.
10. Ishii K, Amagai M, Hall RP, et al. Characterization in pemphigus using antigen-specific enzyme-linked immunosorbent assays with baculovirus-expressed recombinant desmogleins. J Immunol 1997;159:2010–17.
11. Anhalt GJ, Labib RS, Voorhees JJ, et al. Induction of pemphigus in neonatal mice by passive transfer of IgG from patients with the disease. N Engl J Med 1982;306:1189–96.
12. Aoki V, Rivitti EA, Diaz LA, et al. Update on fogo selvagem, an endemic form of pemphigus foliaceus. J Dermatol 2015;42:18–26.
13. Warren SJ, Lin MS, Giudice GJ, et al. The prevalence of antibodies against desmoglein 1 in endemic pemphigus foliaceus in Brazil. Cooperative Group on Fogo Selvagem Research. N Engl J Med 2000;343:23–30.
14. Diaz LA, Prisayanh PS, Dasher DA, et al. The IgM anti-desmoglein 1 response distinguishes Brazilian pemphigus foliaceus (fogo selvagem) from other forms of pemphigus. J Invest Dermatol 2008;128:667–75.
15. Warren SJ, Arteaga LA, Rivitti EA, et al. The role of subclass switching in the pathogenesis of endemic pemphigus foliaceus. J Invest Dermatol 2003;120:104–8.
16. Qian Y, Jeong JS, Ye J, et al. Overlapping IgG4 responses to self- and environmental antigens in endemic pemphigus foliaceus. J Immunol 2016;196:2041–50.
17. Malhoney MG, Wang ZH, Stanley JR. Pemphigus vulgaris and pemphigus foliaceus antibodies are pathogenic in plasminogen activator knockout mice. J Invest Dermatol 1999;113:22–5.
18. Berkowitz P, Hu P, Warren S, et al. p38MAPK inhibition prevents disease in pemphigus vulgaris mice. Proc Natl Acad Sci USA 2006;103:12855–60.
19. Chams-Davatchi C, Esmaili N, Daneshpazhooh M, et al. Randomized controlled open-label trial of four treatment regimens for pemphigus vulgaris. J Am Acad Dermatol 2007;57:622–8.
20. Beissert S, Mimouni D, Kanwar AJ, et al. Treating pemphigus vulgaris with prednisone and mycophenolate mofetil: a multicenter, randomized, placebo-controlled trial. J Invest Dermatol 2010;130:2041–8.
21. Hall RP 3rd, Fairley J, Woodley D, et al. A multicentre randomized trial of the treatment of patients with pemphigus vulgaris with infliximab and prednisone compared with prednisone alone. Br J Dermatol 2015;172:760–8.
22. Amagai M, Ikeda S, Shimizu H, et al. A randomized double-blind trial of intravenous immunoglobulin for pemphigus. J Am Acad Dermatol 2009;60:595–603.
23. Joly P, Mouquet H, Roujeau JC, et al. A single cycle of rituximab for the treatment of severe pemphigus. N Engl J Med 2007;357:545–52.
24. Colliou N, Picard D, Caillot F, et al. Long-term remissions of severe pemphigus after rituximab therapy are associated with prolonged failure of desmoglein B cell response. Sci Transl Med 2013;5:175ra130.
25. Jedlickova H, Hlubinka M, Pavlik T, et al. Bullous pemphigoid and internal diseases—a case-control study. Eur J Dermatol 2010;20: 96–101.
26. Liu JJ, Ding J, Kowal AS, et al. BPAG1n4 is essential for retrograde axonal transport in sensory neurons. J Cell Biol 2003;163:223–9.
27. Chen YJ, Wu CY, Lin MW, et al. Comorbidity profiles among patients with bullous pemphigoid: a nationwide population-based study. Br J Dermatol 2011;165:593–9.
28. Giudice GJ, Emery DJ, Zelickson BD, et al. Bullous pemphigoid and herpes gestationis autoantibodies recognize a common non-collagenous site on the BP180 ectodomain. J Immunol 1993;151:5742–50.
29. McGrath JA, Darling T, Gatalica B, et al. A homozygous deletion mutation in the gene encoding the 180-kDa bullous pemphigoid antigen (BPAG2) in a family with generalized atrophic benign epidermolysis bullosa. J Invest Dermatol 1996;106:771–4.
30. Messingham KA, Holahan HM, Fairley JA. Unraveling the significance of IgE autoantibodies in organ-specific autoimmunity: lessons learned from bullous pemphigoid. Immunol Res 2014;59:273–8.
31. Kirtschig G, Middleton P, Bennett C, et al. Interventions for bullous pemphigoid. Cochrane Database Syst Rev 2010;(10):CD002292.
32. Beissert S, Werfel T, Frieling U, et al. A comparison of oral methylprednisolone plus azathioprine or mycophenolate mofetil for the treatment of bullous pemphigoid. Arch Dermatol 2007;143:1536–42.
33. Peterson JD, Chan LS. Effectiveness and side effects of anti-CD20 therapy for autoantibody-mediated blistering skin diseases: a comprehensive survey of 71 consecutive patients from the initial use to 2007. Ther Clin Risk Manag 2009;5:1–7.
34. Yu KK, Crew AB, Messingham KA, et al. Omalizumab therapy for bullous pemphigoid. J Am Acad Dermatol 2014;71:468–74.
35. Ishii N, Hamada T, Dainichi T, et al. Epidermolysis bullosa acquisita: what's new? J Dermatol 2010;37:220–30.
36. Gammon WR, Heise ER, Burke WA, et al. Increased frequency of HLA-DR2 in patients with autoantibodies to epidermolysis bullosa acquisita antigen: evidence that the expression of autoimmunity to type VII collagen is HLA class II allele associated. J Invest Dermatol 1988;91:228–32.
37. Ludwig RJ, Recke A, Bieber K, et al. Generation of antibodies of distinct subclasses and specificity is linked to H2s in an active mouse model of epidermolysis bullosa acquisita. J Invest Dermatol 2011;131:167–76.
38. Srinivas G, Moller S, Wang J, et al. Genome-wide mapping of gene-microbiota interactions in susceptibility to autoimmune skin blistering. Nat Commun 2013;4:2462.
39. Tukaj S, Zillikens D, Kasperkiewicz M. Inhibitory effects of heat shock protein 90 blockade on proinflammatory human Th1 and Th17 cell subpopulations. J Inflamm (Lond) 2014;11:10.
40. Chi CC, Wang SH, Charles-Holmes R, et al. Pemphigoid gestationis: early onset and blister formation are associated with adverse pregnancy outcomes. Br J Dermatol 2009;160:1222–8.
41. Shornick JK, Artlett CM, Jenkins RE, et al. Complement polymorphism in herpes gestationis: association with C4 null allele. J Am Acad Dermatol 1993;29:545–9.
42. Chan LS, Ahmed AR, Anhalt GJ, et al. The first international consensus on mucous membrane pemphigoid: definition, diagnostic criteria, pathogenic factors, medical treatment, and prognostic indicators. Arch Dermatol 2002;138:370–9.
43. Yasukochi A, Teye K, Ishii N, et al. Clinical and immunological study of 332 Japanese patients tentatively diagnosed with anti-BP180-type mucous membrane pemphigoid: a novel BP180 C-terminal domain enzyme-linked immunosorbent assay. Acta Derm Venereol 2016;96: 762–7.
44. Chan LS, Yancey KB, Hammerberg C, et al. Immune-mediated subepithelial blistering diseases of mucous membranes: pure ocular cicatricial pemphigoid is a unique clinical and immunopathological entity distinct from bullous pemphigoid and other subsets identified by antigenic specificity of autoantibodies. Arch Dermatol 1993;129:448–55.
45. Egan CA, Lazarova Z, Darling TN, et al. Anti-epiligrin cicatricial pemphigoid: clinical findings, immunopathogenesis, and significant associations. Medicine (Baltimore) 2003;82:177–86.
46. Shipman AR, Ali I, Murrell DF, et al. Mucous membrane pemphigoid: are laminin 5 antibodies a risk factor for laryngeal involvement? J Eur Acad Dermatol Venereol 2009;23:169–70.

46a. Setterfield J, Theron J, Vaughan RW, et al. Mucous membrane pemphigoid: HLADQB1* 0301 is associated with all clinical sites of involvement and may be linked to antibasement membrane IgG production. Br J Dermatol 2001;145(3):406–14.

47. Le Roux-Villet C, Prost-Squarcioni C, Alexandre M, et al. Rituximab for patients with refractory mucous membrane pemphigoid. Arch Dermatol 2011;147:843–9.
48. Maley A, Warren M, Haberman I, et al. Rituximab combined with conventional therapy versus conventional therapy alone for the treatment of mucous membrane pemphigoid (MMP). J Am Acad Dermatol 2016;74:835–40.
49. Fortuna G, Salas-Alanis JC, Guidetti E, et al. A critical reappraisal of the current data on drug-induced linear immunoglobulin A bullous dermatosis: a real and separate nosological entity? J Am Acad Dermatol 2012;66:988–94.
50. Zone JJ, Taylor TB, Meyer LJ. The 97 kDa linear IgA bullous disease antigen is identical to a portion of the extracellular domain of the 180 kDa bullous pemphigoid antigen, BPAg2. J Invest Dermatol 1998;110:207–10.
51. Paek SY, Steinberg SM, Katz SI. Remission in dermatitis herpetiformis: a cohort study. Arch Dermatol 2011;147:301–5.
52. West J, Fleming KM, Tata LJ, et al. Incidence and prevalence of celiac disease and dermatitis herpetiformis in the UK over two decades: population-based study. Am J Gastroenterol 2014;109:757–68.

52a. Salmi TT, Hervonen K, Kautiainen H, et al. Prevalence and incidence of dermatitis herpetiformis: a 40-year prospective study from Finland. Br J Dermatol 2011;165(2):354–9.

53. Lawley TJ, Strober W, Yaoita H, et al. Small intestinal biopsies and HLA types in dermatitis herpetiformis patients with granular and linear IgA skin deposits. J Invest Dermatol 1980;74:9–12.
54. Marks J, Shuster S, Watson AJ. Small bowel changes in dermatitis herpetiformis. Lancet 1966;2:1280–2.
55. Grainge MJ, West J, Solaymani-Dodaran M, et al. The long-term risk of malignancy following a diagnosis of coeliac disease or dermatitis herpetiformis: a cohort study. Aliment Pharmacol Ther 2012;35:730–9.
56. Viljamaa M, Kaukinen K, Pukkala E, et al. Malignancies and mortality in patients with coeliac disease and dermatitis herpetiformis: 30-year population-based study. Dig Liver Dis 2006;38:374–80.
57. Dieterich W, Laag E, Bruckner-Tuderman L, et al. Antibodies to tissue transglutaminase as serologic markers in patients with dermatitis herpetiformis. J Invest Dermatol 1999;113:133–6.
58. Zone JJ, Schmidt LA, Taylor TB, et al. Dermatitis herpetiformis sera or goat anti-transglutaminase-3 transferred to human skin-grafted mice mimics dermatitis herpetiformis immunopathology. J Immunol 2011;186:4474–80.
59. Hall RP, Takeuchi F, Benbenisty K, et al. Cutaneous endothelial cell activation in normal skin of patients with dermatitis herpetiformis associated with increased serum levels of IL-8, sE-Selectin and TNF-α. J Invest Dermatol 2006;126:1331–7.

MULTIPLE-CHOICE QUESTIONS

1. In the evaluation of a patient with suspected autoimmune blistering skin disease diagnosis requires which of the following (may have more than one correct answer):
   A. Direct immunofluorescence testing of lesional skin
   B. Routine histology of lesional skin
   C. Direct immunofluorescence testing of normal appearing perilesional skin
   D. A and B
   E. B and C

2. Which of the following statements regarding bullous pemphigoid are correct? May have more than one correct answer.
   A. Most common autoimmune blistering disease and most often presents in patients 20–40 years of age
   B. Can present as a vesicular, bullous, or urticarial eruption
   C. Patients have few if any clinical symptoms of their skin eruption
   D. Characterized by linear immunoglobulin G (IgG) and/or C3 at the basement membrane of stratified squamous epithelium by direct immunofluorescence testing of patient's skin
   E. Associated with gluten sensitive enteropathy

3. Patients with pemphigus vulgaris are characterized by (may have more than one correct answer):
   A. Subepidermal blister with eosinophils by routine histology
   B. Oral mucosal ulcerations are common either alone or in combination with skin lesions
   C. Immunoglobulin G (IgG) antibodies can be found against desmoglein-1 (Dsg1) and/or Dsg3
   D. Occurs in older adults more frequently and may be associated with neurological disease
   E. Recent evidence suggests that therapy directed at CD20 expressing cells may be very effective

64

Immunology of Psoriasis

Cristina Albanesi

CLINICAL AND HISTOLOGICAL FEATURES OF PSORIASIS

Psoriasis is a common chronic, relapsing immune-mediated disease involving skin and small joints of genetically predisposed individuals. It affects approximately 2% of the general population, with >50% of patients presenting in the first three decades of life. There is a wide spectrum of cutaneous manifestations of psoriasis. Individual lesions vary from pinpoint to large plaques, or even generalized erythroderma. More specifically, the clinical spectrum of psoriasis includes the plaque, guttate, small plaque, inverse, erythrodermic, and pustular variants.[1] The most common and well-recognized morphological presentation of psoriasis is that of the plaque type (Fig. 64.1). The disease is characterized by the formation of demarked erythematous plaques with large scaling. The scales are a result of a hyperproliferative epidermis with premature maturation of keratinocytes and incomplete cornification, with retention of nuclei in the stratum corneum (parakeratosis). The mitotic rate of the basal keratinocytes is increased compared with that of normal skin. As a result, the epidermis is thickened (acanthosis), with elongated rete ridges that form finger-like protrusions into the dermis. The granular layer of the epidermis (the starting site of terminal keratinocyte differentiation) is strongly reduced or missing. The epidermis is infiltrated by neutrophils and activated CD8 T lymphocytes, whereas within the dermis there is an inflammatory infiltrate composed mainly of $CD3^+$ T cells, dendritic cells (DCs), macrophages, mast cells, and neutrophils. Elongated and dilated blood vessels in the dermal papillae represent a further histological hallmark of psoriatic skin lesions (Fig. 64.2).[1]

IMMUNE-RELATED GENETIC FACTORS PREDISPOSING TO PSORIASIS

The genetic basis of psoriasis has long been recognized, since family members of patients with psoriasis are at greater risk of developing the disease. The concordance rate of psoriasis is approximately 70% in monozygotic twins and 20% in dizygotic twins, depending on the study and population.[2] The mode of inheritance is complex. It is thought that there is no single disease gene, but, rather, a complex set of gene variants, resulting in an aberrant response to environmental factors. At least nine chromosomal loci with statistically significant linkage to psoriasis have been identified, termed *psoriasis susceptibility loci 1* through *9* (*PSORS1* through *PSORS9*).[2] *PSORS1* is the major genetic determinant of psoriasis and is located within the human leukocyte antigen (HLA) complex on chromosome 6p (Chapter 5). An association with HLA-B13 was first identified and later with other class I molecules HLA-B17, HLA-B37, HLA-B57, HLA-Cw6, and HLA-Cw7, and class II molecules HLA-DR4 and HLA-DR7. Among these, the highest and most consistently reported relative risk is for HLA-Cw6 haplotype. A significantly higher frequency of HLA-Cw6 is associated with early-onset (type I) psoriasis compared with late-onset psoriasis (type II). Current data suggest that HLA-Cw6 is the susceptibility allele within *PSORS1,* but no disease-specific mutations have been identified, and variants in regulatory sequences potentially affecting several downstream genes cannot be ruled out. HLA-C might be involved in immune responses at the levels of both antigen presentation and natural killer (NK)–cell regulation. Recently, compelling evidence has emerged for an interaction between the HLA-C and *ERAP1* (involved in major histocompatibility complex (MHC) class I peptide processing) loci in psoriasis.[3] 36763*ERAP1* variants only influenced psoriasis susceptibility in individuals carrying the HLA-C risk allele.

Other predisposing polygenes were found in the *PSORS2* region, on chromosome 17q25. Two distinct regions harboring susceptibility loci have been identified: the first contains the genes *SLC9A3R1* and *NAT9* and the second the gene *RAPTOR.*[4] The *SLC9A3R1* gene encodes a PDZ domain-containing phosphoprotein, implicated in several biological processes occurring in T cells. *NAT9* encodes an N-acetyltransferase involved in MHC class I antigen-presentation and in immunological processes related to autoimmune diseases. Additionally, between *SLC9A3R1* and *NAT9* is a polymorphism for the binding site of a transcription factor RUNX1 that may affect regulation of the immune synapse.[4] Associations with alleles of interleukin (IL)-12, IL-23 receptor (IL-23R), IL-19/20, and IRF2 have also been described. Interestingly, IL-12 and IL-23R single nucleotide polymorphisms do not have interactions with HLA-Cw6.[4] Recently, polymorphisms have been found in two genes, *IL-36RN* and *CARD14,* and to be independently associated with psoriasis.[4,5] For *CARD14* (caspase recruitment domain-containing protein 14) gene, many missense mutations, leading to elevation of CARD14 mRNA in patients, were initially found in pediatric patients with a severe clinical presentation of psoriasis. CARD14 gene mutation protein, in association with an inflammatory stimulus, may induce aberrant activation of nuclear factor (NF)-κb, a transcription factor that controls the expression of many genes, including key chemokines upregulated in psoriasis, such as chemokine (C-X-C motif) ligand (CXCL)8 and CC chemokine ligand (CCL)20. Mutations in *IL-36RN* gene were described in patients with severe pustular psoriasis. This gene encodes the antiinflammatory protein IL-36Ra, antagonist of IL-36γ, a cytokine highly produced in psoriatic lesions and with important proinflammatory function.

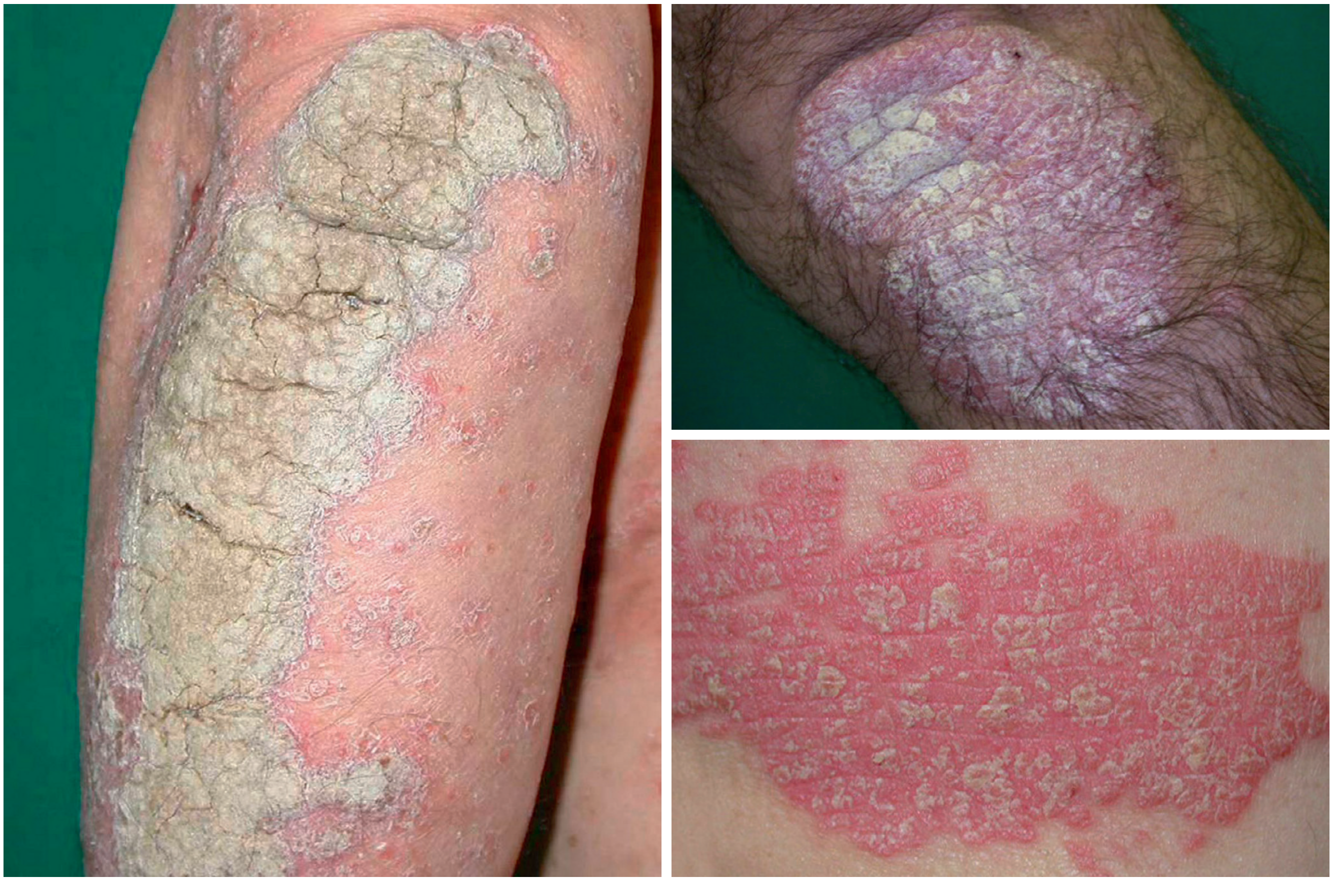


FIG 64.1 **Clinical Features of Plaque Psoriasis.** Scaly, erythematous, sharply demarcated plaques in different sizes and shapes are hallmarks of psoriasis.

KEY CONCEPTS

Psoriasis is a common chronic-relapsing immune-mediated skin disease affecting approximately 2% of the general population.

There is strong evidence that psoriasis is determined by genetic predisposition. A complex set of gene variants, rather than a single gene, are responsible for an aberrant response to environmental factors.

Psoriasis is a disease caused by the infiltration of effector immune cells in both the epidermis and dermis, which determines hyperproliferation of the epidermis with premature maturation of keratinocytes and incomplete cornification. As a result, the epidermis is thickened, with elongated rete ridges forming protrusions into the dermis.

Primary effector cells are dermal dendritic cells (DCs), in particular plasmacytoid DCs (pDCs), whose activation can depend on DNA-LL-37 or RNA-LL-37 complexes released by injured keratinocytes and leads to a massive production of interferon (IFN)-α.

pDC-released IFN-α or RNA-LL-37 complexes released by keratinocytes activate myeloid DCs (mDCs), which in turn induces type-1 and -17 T-cell responses. T-helper 22 response is also pathogenetically induced.

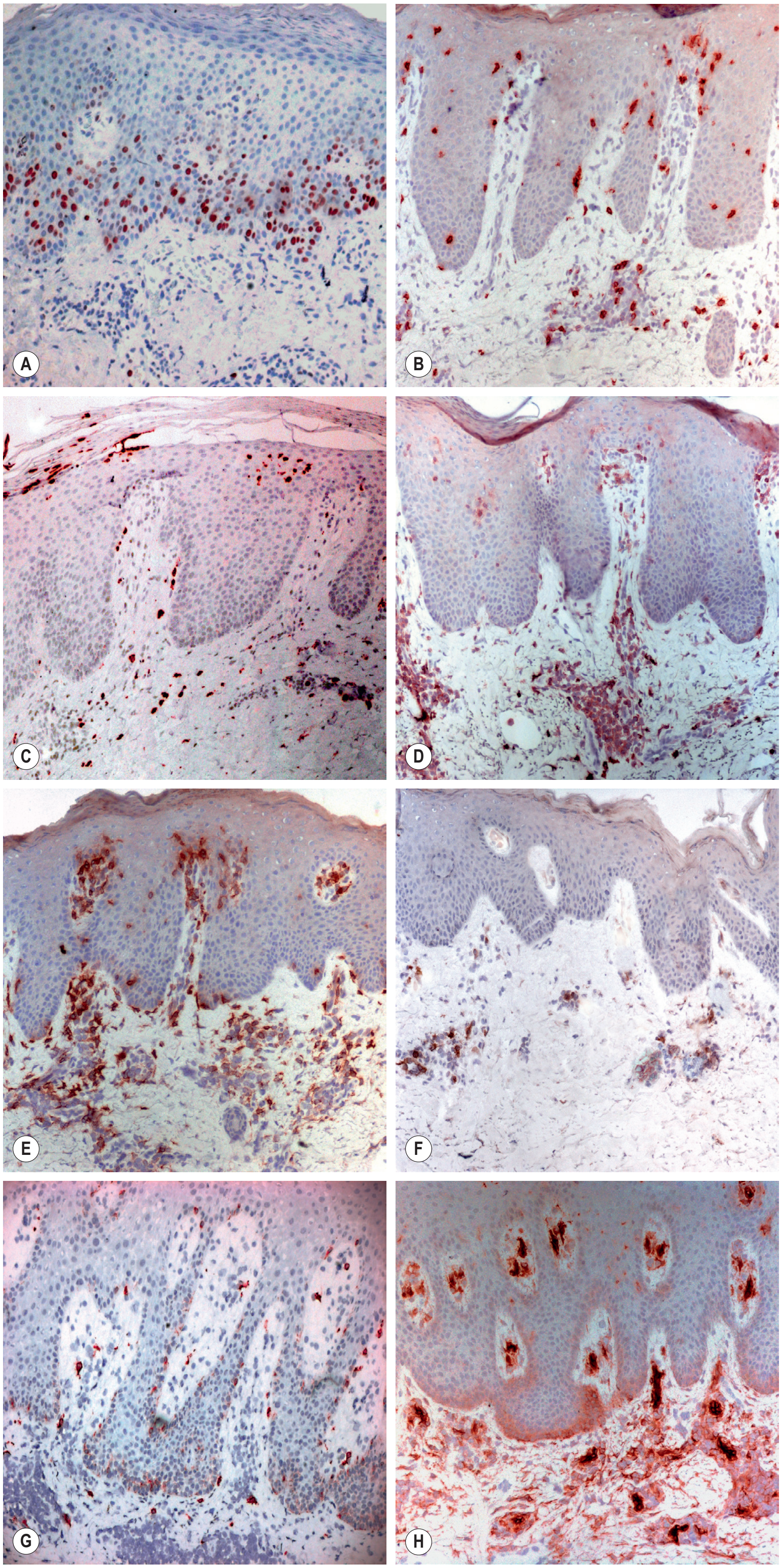
Pathological cytokines include T-cell-derived lymphokines, such as IFN-γ, tumor necrosis factor (TNF)-α, IL-17, IL-22, IL-21, and antigen-presenting cell–derived cytokines, such as IL-12 and IL-23.

Intrinsic alterations of keratinocytes in the activation of signal transduction pathways (*i.e.,* STAT3, IKK-2, AP-1, etc.) are fundamental for the amplification of psoriatic processes.

EFFECTOR CELLS AND IMMUNE MECHANISMS OPERATING IN PSORIASIS

Even though successful treatment regimens for the therapy of psoriasis are long established, the primary pathogenetic mechanism and the cell types involved in the onset of the disease are still under debate. Psoriasis is classically responsive to trigger factors that can induce psoriasis *de novo* or exacerbate skin lesions.[6] Trigger factors range from nonspecific stimuli, such as skin trauma (termed the *Koebner effect*) to more specific triggers, such as pathogens (*i.e.,* streptococci) or drugs (*i.e.,* lithium, interferon (IFN)-α). All of these factors generate a pathogenic cascade culminating in the expansion of lesional and/or circulating T cells in the psoriatic skin (Fig. 64.3). Much effort has been devoted to understanding the link between the trigger stimuli and the pathogenic T-cell cascade that leads to psoriasis. Recent evidence suggests that type-1 IFN may represent the missing link. The prototypical type I IFN, IFN-α, is abundantly produced by plasmacytoid dendritic cells (pDCs) during the early phase of psoriasis development (see Fig. 64.3A).[6,7] In turn, IFN-α indirectly stimulates the pathogenic T-cell cascade by promoting the activation and maturation of myeloid DCs (mDCs) or by direct stimulation of IFN-α–sensitive pathogenic T cells (see Fig. 64.3B). The T-cell infiltrate present in active psoriatic skin establishes a cytokine milieu that dictates specific and pathogenic gene signatures in resident skin cells. Thus cytokine-activated keratinocytes overexpress a number of inflammatory mediators that aberrantly amplify and sustain the psoriasiform tissue reactions (see Fig. 64.3C). Intrinsic defects and/or alterations of keratinocytes in their immune response to proinflammatory cytokines are fundamental to the induction of psoriatic processes, as demonstrated in genetically manipulated mouse systems.

FIG 64.2 Histological Components of a Mature Psoriatic Plaque. Psoriatic skin lesions are characterized by a hyperproliferative epidermis showing an increased mitotic rate of the basal keratinocytes (A, Ki67 immunostaining). As a consequence, the epidermis is thickened, with elongated rete ridges that form typical finger-like protrusions into the dermis. The epidermis becomes infiltrated by activated CD8 T lymphocytes and neutrophils (B and C, immunostaining for CD8 and CD15, respectively). Within the drmis, an inflammatory infiltrate mainly composed of $CD3^+$ T cells (D), $CD11c^+$ dendritic cells (E), BDCA-2^+ plasmacytoid dendritic cells (F), c-kit^+ mast cells (G), and neutrophils (C) is observed. Elongated and dilated ICAM-1^+ (H) blood vessels in the dermis represent another histological hallmark of psoriatic skin.



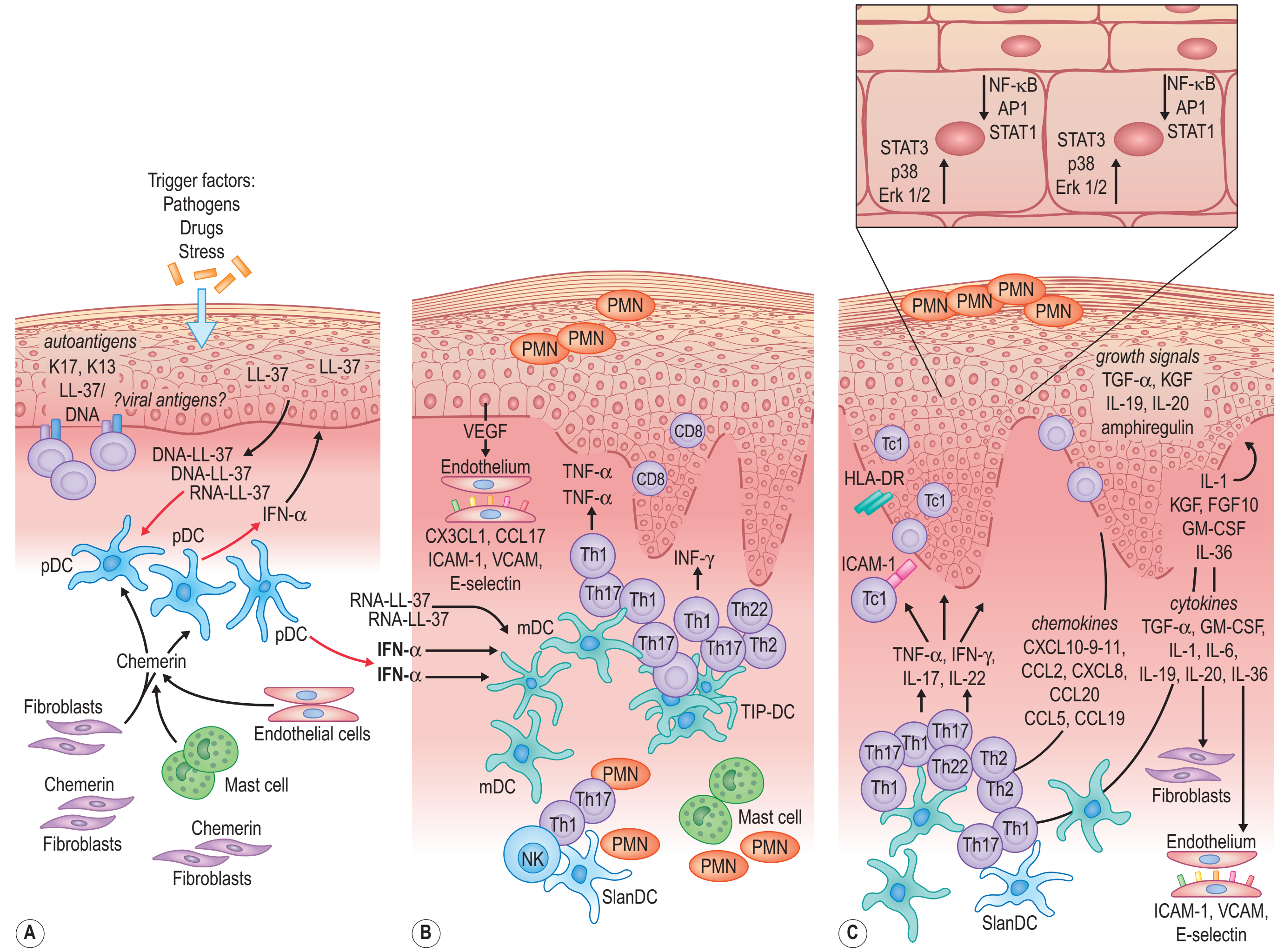

Trigger factors:
Pathogens
Drugs
Stress
autoantigens
K17, K13
LL-37/
DNA
?viral antigens?
LL-37
LL-37
DNA-LL-37
DNA-LL-37
RNA-LL-37
IFN-α
pDC
pDC
pDC
Chemerin
Fibroblasts
Endothelial cells
Mast cell
Chemerin
Fibroblasts
Chemerin
Fibroblasts
A
PMN
PMN
PMN
VEGF
Endothelium
CX3CL1, CCL17
ICAM-1, VCAM,
E-selectin
TNF-α
TNF-α
CD8
CD8
INF-γ
Th1
Th17
Th1
Th1
Th22
Th17
Th17
Th2
RNA-LL-37
RNA-LL-37
IFN-α
IFN-α
mDC
TIP-DC
mDC
PMN
Th17
Th1
NK
PMN
SlanDC
Mast cell
PMN
PMN
B
NF-κB
AP1
STAT1
STAT3
p38
Erk 1/2
NF-κB
AP1
STAT1
STAT3
p38
Erk 1/2
PMN
PMN
PMN
PMN
growth signals
TGF-α, KGF
IL-19, IL-20
amphiregulin
Tc1
HLA-DR
Tc1
ICAM-1
Tc1
IL-1
KGF, FGF10
GM-CSF
IL-36
cytokines
TGF-α, GM-CSF,
IL-1, IL-6,
IL-19, IL-20, IL-36
chemokines
CXCL10-9-11,
CCL2, CXCL8,
CCL20
CCL5, CCL19
TNF-α, IFN-γ,
IL-17, IL-22
Th17
Th1
Th17
Th22
Th2
Th1
Th2
Th1
Th17
Fibroblasts
SlanDC
Endothelium
ICAM-1, VCAM,
E-selectin
C

FIG 64.3 Scheme of Pathogenic Mechanisms Operating in Psoriasis. The psoriatic lesion starts to evolve after keratinocytes are injured, for instance by physical trauma or bacterial products. Thereafter, a cascade of events including the formation of complexes formed by keratinocyte-derived DNA and the cathelicidin LL-37 leads to the activation of plasmacytoid dendritic cells (pDCs), which routinely patrol psoriatic skin (A). Other pDCs are recruited in the early phase psoriasis development by the chemokine chemerin, derived primarily from dermal fibroblasts and, to a lesser extent, from mast cells and endothelial cells, and induced to release high amounts of IFN-α. IFN-α locally activates keratinocytes and participates in the activation processes affecting myeloid DCs. In turn, DCs migrate into draining lymph nodes and induce the differentiation of naïve T cells into effector cells, such as type 17 T-helper (Th) 17 or type 17 T cytotoxic (Tc) cells and type 1 Th1 or Tc1 cells. Effector cells recirculate and proliferate into psoriatic skin and produce massive amounts of proinflammatory cytokines, such as interferon (IFN)-γ and tumor necrosis factor (TNF)-α (B). The latter cytokine is also abundantly released by dermal DCs, mainly represented by TNF-α and inducible nitric oxide synthase-producing-DCs (TIP-DC). SlanDCs also reinforce immunity in psoriasis by interacting with and potentiating the activity of both neutrophils and natural killer (NK) cells, as well as inducing Th1 and Th17 responses. IFN-γ and TNF-α are responsible for the activation of resident skin cells, in particular keratinocytes, which respond to cytokines with a stereotypical set of genomic responses leading to synthesis of inflammatory mediators (C). Keratinocytes are also targets of T cell–derived IL-22, which induces proliferation and de-differentiation of psoriatic keratinocytes in a signal transducer and activator of transcription 3 (STAT3)-dependent manner. Keratinocyte-derived chemokines, cytokines, and membrane molecules have a major role in maintaining the recruitment of leukocytes into inflammatory sites. Because of their intrinsic defects, psoriatic keratinocytes aberrantly respond to cytokines and show altered intracellular signaling pathways, including STAT3 cascade (C). The uncontrolled hyperproliferation and differentiation observed in psoriatic skin could also derive from dysregulated production of tissue growth factors and regulators, such as transforming growth factor (TGF)-α, keratinocyte growth factor (KGF), amphiregulin, granulocyte macrophage–colony-stimulating factor (GM-CSF), fibroblast growth factor-10 (FGF-10), interleukin (IL)-19, IL-20, IL-36 produced by keratinocytes and fibroblasts. Psoriatic keratinocytes produce autoantigens (*i.e.,* keratin (K)17, K13, nucleic acid/LL-37 complexes) capable of inducing clonal T-cell responses. Finally, the inflammatory cytokine milieu also influences the immune functions of fibroblasts and endothelium, with the latter being critical for leukocyte trafficking and extravasation.

Plasmacytoid DCs as Inducers of Primary Immune Responses in Psoriasis

pDCs are characterized by a plasma cell morphology and a distinctive surface phenotype (CD4$^+$, CD45RA$^+$, CD123$^+$, BDCA-2$^+$, BDCA-4$^+$, CD62L$^+$, cutaneous lymphocyte-associated antigen (CLA)$^+$, and CD11c$^-$). They are considered as key effector cells in antiviral defense because of their ability to produce large amounts of type I IFN.[8] Upon viral stimulation, pDCs differentiate into a unique type of mature DC and induce an IFN-α–dependent activation of bystander mDCs with the ability to induce Th1-responses (Chapter 16), thus providing a necessary link between innate and adaptive immunity.[8] Several studies have demonstrated that pDCs infiltrate psoriatic skin, and that pDC-derived IFN-α initiates the expansion of autoimmune T cells, leading eventually to the skin lesions of psoriasis.[6,7,9,10] Blocking of type I IFN signaling with neutralizing antibodies to IFN-α/β receptors inhibited the development of psoriasis in symptomless pre-psoriatic skin engrafted onto immunodeficient AGR129 mice.[10] Moreover, inhibition of IFN-α release by pDCs through anti–BDCA-2 antibody prevented the activation and expansion of pathogenetic T cells and the development of a psoriatic phenotype. The detailed mechanisms responsible for the IFN-α-induced expansion of T cells in psoriasis are currently unknown, but it appears that IFN-α favors cross-presentation of sequestered tissue-specific autoantigens by mDCs. pDC-derived IFN-α may also enhance the survival of autoreactive T cells through the induction of IL-15 or by promoting a Th1 cell bias through the induction of T-bet and IL-12Rβ_2 expression. The pathogenetic role of IFN-α is also suggested by the observations that its signaling signature is present in resident skin cells of psoriatic plaques[11] and that psoriasis is exacerbated if patients with psoriasis are treated with recombinant IFN-α for unrelated conditions (*i.e.,* viral infections or tumors), or with imiquimod, a Toll-like receptor (TLR) agonist that induces production of IFN-α by pDCs.[12] In addition, excessive activation of type I IFN signaling in mice deficient for IFN regulatory factor (IRF)-2, a transcriptional repressor of IFN signaling, causes an inflammatory skin disease resembling psoriasis. Recently, a genome-wide analysis conducted on paired lesional and nonlesional psoriatic skin and on the skin of healthy donors revealed a significant overexpression of many components of the IFN-α pathway in patients with psoriasis, including the receptor subunits for type I IFN, IFN-AR1, and IFN-AR2, the transcriptional activators of IFN-α-inducible genes, signal transducer and activator of transcription (STAT) 1, IRF1, and IRF7.[11] All of these molecules are master regulators of IFN-α-mediated immune responses.

The molecular mechanisms leading pDCs to produce type I IFN involve the activation of TLR7 and TLR9, intracellular receptors that recognize viral/microbial nucleic acids within endosomal compartments. pDCs do not normally respond to self-DNA, but this restriction breaks down in some human autoimmune diseases. In psoriatic skin, pDCs can be activated to produce massive amounts of type I IFN in response to extracellular self-DNA fragments.[13] However, this process requires the coupling of self-DNA to the endogenous antimicrobial peptide LL-37, known to be overexpressed in psoriatic skin. LL-37 breaks innate tolerance to self-DNA by forming aggregated and condensed structures that can trigger a robust IFN-α induction via TLR9 activation.[13] LL-37 can also form complexes with RNA and activate pDCs through TLR7. In parallel, LL-37/RNA can alert myeloid DCs through their TLR8, driving T-cell activation and production of cytokines found in psoriasis.[14] This finding suggests a fundamental role for LL-37 in alerting resident skin pDCs of tissue damage associated with cell death and the release of self-DNA. pDCs are typically absent in unperturbed skin and peripheral tissues under homeostatic conditions but can enter secondary lymphoid organs through the expression of CD62L and chemokine receptors (CXCR4, CXCR3, CCR5, and ChemR23). pDCs can also infiltrate inflamed tissue of immune-mediated skin diseases, and, in particular, accumulate in the skin of patients with psoriasis early during disease development.[9,10] Many studies have shown an evident IFN-α signature (*e.g.,* increased expression of IRF7 and the presence of MxA, markers for IFN-α activity) in primary psoriatic plaques in the absence of detectable levels of the IFN-α cytokine. Indeed, IFN-α expression was detected early and transiently during the development of the psoriatic phenotype, and its effect persisted until lesions became chronic. Paralleling IFN-α expression, BDCA-2$^+$ pDCs were detected only during the early developmental stages of psoriasis.[10] In fact, pDC infiltration in psoriatic skin correlates with the expression of markers typical of early phases of psoriasis (CD15$^+$ neutrophils and c-kit$^+$ mast cells localized in the mid- and papillary dermis, and few CD8 T lymphocytes or ICAM-1$^+$ keratinocytes), whereas they are almost absent in long-lasting lesions.[9] Importantly, pDC recruitment in psoriatic skin is strictly associated with the expression of the chemokine chemerin, which is temporally produced by dermal fibroblasts and active during psoriatic plaque development.[9] pDC migration toward fibroblast-derived chemerin is completely dependent on the expression of ChemR23 receptor on pDC. Compared with other chemokines potentially active on pDC (CXCL10 and CXCL12), chemerin is the main, if not the only, protein responsible for the pDC chemotactic activity released by fibroblasts in psoriatic skin.[9]

DC Driving of T-Cell Responses in Psoriatic Skin

Although pDCs are responsible for triggering psoriasis, mDCs are the main amplifiers of local inflammation. Dermal mDCs are dramatically increased in psoriasis, and targeted immunotherapy reduces their quantity in patients with psoriasis, supporting the concept that mDCs have a key pathogenetic role.[15] Dermal mDCs are found at the dermal–epidermal junction as well as throughout the whole dermis. mDCs are able to capture extracellular antigens for presentation to T cells and also intracellular antigens from adjacent cell types via cross-presentation (Chapter 6). In addition, mDCs within psoriatic lesions are intrinsically stronger stimulators of T-cell proliferation compared with DCs derived from peripheral blood or from the skin of healthy patients. mDCs uniformly express CD11c, and they can be further subdivided on the basis of expression of CD1c (BDCA-1). Steady-state skin has a predominance of CD11c$^+$ CD1c$^+$ resident DCs, whereas CD11c$^+$ CD1c$^-$ mDCs predominate in psoriatic inflammation.[15] A small fraction of CD11c$^+$ CD1c$^+$ DC bears “maturation” markers, such as DC-LAMP, CD83, and endocytic receptor DEC-205/CD205, suggesting that they could function as conventional DCs and present antigens to T cells to trigger acquired immune responses.[15] These rare, phenotypically mature cells, often aggregating in dermal clusters, could be required for rapid antigen presentation to local T cells or for ongoing “micro-” immune responses. During psoriasis development, dermal CD11c$^+$ DC mature and acquire a CD1c$^-$ HLA-DR$^+$ CD45$^+$ CD14$^-$ DC-specific ICAM-3–grabbing nonintegrin (DC-SIGN)$^+$ phenotype.[15] These inflammatory mDCs are CCR7$^+$ and respond to the chemokine CCL19, suggesting that they may migrate to draining lymph nodes for antigen presentation. CD11c$^+$ CD1c$^-$ inflammatory DC express very high levels of tumor necrosis factor (TNF)-α and the enzyme inducible nitric oxide synthase (iNOS) and can be considered the human equivalent of TIP-DC (TNF-α and iNOS-producing DCs), which have been shown in mice to have effector functions in clearing some bacterial infections.[15] The proinflammatory nature of TNF-α in psoriasis and other inflammatory diseases is well established, and the receptors TNFR1 and TNFRII are expressed on a wide range of cells in psoriatic skin, including keratinocytes and endothelial cells. TNF-α induces expression of ICAM-1 on keratinocytes, facilitating the adhesion of circulating leukocytes. Moreover, TNF-α can stimulate keratinocytes and dermal fibroblasts to produce the potent neutrophil chemoattractant CXCL8, as well as the proinflammatory cytokines IL-6 and IL-1, which help generate and maintain Th17 cells (Chapters 10 and 16). The role of TNF-α in psoriasis is underlined by the therapeutic success of anti–TNF-α therapies in psoriasis.[15] In addition, polymorphisms of the TNF-α promoter region have been associated with psoriasis. However, iNOS production by inflammatory DCs leads to nitric oxide (NO) release, inducing vasodilation, inflammation, and antimicrobial effects in psoriatic skin. Interestingly, NO inhibitors (*e.g.,* statins) can have a beneficial effect on psoriasis, although this has not been studied in a randomized prospective manner. Inflammatory DCs also produce other cytokines (Chapter 9) (*e.g.,* IL-23 and -12), which are linked to psoriasis. IL-12 mainly induces IFN-γ production, whereas IL-23 also stimulates IL-17 and IL-22 release by T cells, as shown in mice.[16] Other evidence for a role of IL-23 in psoriasis includes the clinical efficacy of an anti-p40 monoclonal antibody (mAb) against psoriasis and the association of a single nucleotide polymorphism in the *IL-23R* gene in psoriasis patients. mDCs of psoriatic lesions also release the proinflammatory cytokine IL-15, which induces T-cell proliferation and monocyte activation, as well as skin hyperplasia by protecting keratinocytes from apoptosis.

Another population of inflammatory mDCs, defined by the selective expression of the 6-sulfo LacNAc residue on the P-selectin glycoprotein ligand 1 membrane molecule, has been identified in psoriatic skin.[17] 6-Sulfo LacNAc$^+$ DCs (slanDCs) have a well-defined phenotype (CD1c$^-$, CD11c$^+$, CD16$^+$, CD14$^-$) that clearly distinguishes them from classic CD1c$^+$ blood DC (CD1c$^+$, CD11c$^+$) or pDC (BDCA-2, BDCA-4). SlanDCs produce more TNF-α, IL-23, IL-12, IL-1β, and IL-6 and can thus induce Th1/Th17 cells. SlanDCs also reinforce innate immunity in psoriasis by interacting with and potentiating the activity of neutrophils and NK cells.[17]

The function of another DC subset, the Langerhans cell, is still controversial. Langerhans cells are only found in the epidermal compartment where they are similar in number and phenotype in lesional and nonlesional skin, but they fail to migrate in response to proinflammatory stimuli[18] (Chapter 19). These findings and data from Langerhans cell ablation models suggest that Langerhans cells may help sustain immune tolerance in psoriasis. In psoriatic skin, the epidermis contains an additional DC subset, known as *inflammatory dendritic epidermal cells* and are distinguishable from Langerhans cells by the expression of the macrophage mannose receptor CD206.

Activation of T Lymphocytes and Establishment of the Cytokine Milieu Influencing Keratinocyte Proliferation and Immune Functions

Psoriasis lesional skin shows many inflammatory T cells in both the papillary dermis and the epidermis. Immunophenotyping of T cells shows that these are mainly activated memory T cells expressing HLA-DR, CD25, CD27, and cutaneous lymphocyte antigen (CLA). Although there is differential T-cell receptor usage in T cells from psoriatic skin, the causative antigens responsible for T-cell activation in psoriasis remain unknown. Exposure of altered autoantigens from keratinocytes could be responsible for the activation and expansion of distinct T-cell subpopulations in psoriatic skin. These autoantigens may include keratin 17 (patients with active psoriasis have an increased frequency of circulating Th1 cells reacting to peptides from keratin 17) and corneodesmosin, an attractive candidate for psoriasis susceptibility based on its putative biological function in keratinocyte adhesion.[19] Keratinocytes could also be responsible for the activation of pathogenetic T cells by viral or bacterial products. For example, human papillomavirus 5 (HPV-5) DNA and antibodies to HPV-5 virus-like particles have been found in psoriasis. Streptococcal infections are also frequently associated with psoriasis, and streptococcal superantigens could be presented to T cells by binding to MHC class II molecules expressed by lesional keratinocytes. Putative psoriatic antigens are assumed to be keratinocyte proteins that might share structural homology with streptococcal proteins and might thereby induce cross-reactive T-cell responses against skin components.[20] Recently, it has been found that two-thirds of patients with moderate-to-severe plaque psoriasis harbor CD4 and/or CD8 T cells specific for LL-37.[21] LL37-specific T cells produce IFN-γ, and CD4 T cells also produce Th17 cytokines. LL37-specific T cells can infiltrate lesional skin and can be tracked in patients' blood. The presence of circulating LL37-specific T cells correlates significantly with disease activity, suggesting a contribution to disease pathogenesis.

T-cell migration from the dermis into the epidermis is a key event in psoriasis. It is controlled by the interaction of $\alpha_1\beta_1$ integrin (very late antigen 1) on T cells with collagen IV in the basement membrane of the epidermis. Blockade of this interaction inhibits the development of psoriasis in clinically relevant models. Based on the analysis of infiltrating cell types, their secreted products, and genetic signatures present in lesional skin, psoriasis has been considered for many years as a type-1 (*i.e.,* Th1)–mediated reaction, with IFN-γ playing a prominent role.[22] Consistent with a Th1 pattern of response, CCR10 is preferentially expressed by skin homing CLA$^+$ memory T cells, which secrete TNF-α and IFN-γ, but minimal amounts of IL-10 and IL-4, upon activation. However, other cytokines and T-cell subsets have also been identified during inflammatory responses in psoriasis. These include Th17 and Th22 cells, which produce large amounts of IL-17 and IL-22, cytokines that have relevant effects on epithelial cells.[23] Keratinocytes are strongly influenced by IL-17 and upregulate chemokines and immunomodulatory molecules in response to this cytokine.[24] A functional role of Th17 cells, and also Tc17 cells, in psoriasis is suggested by their reduction during successful anti–TNF-α treatment or via blockade with neutralizing mAb in a clinically relevant xenotransplantation mouse model.[25] IL-22 also acts pathogenetically in psoriatic skin by inducing proliferation and de-differentiation of keratinocytes, as well as promoting their production of antimicrobial peptides and chemokines, including CXCL8 and CXCL1.[26] Binding of IL-22 to its receptor, whose expression in the skin is confined to keratinocytes, mediates epidermal acanthosis through the activation of STAT3. These observations may explain the increased STAT3 expression in the epidermal compartment and the pathogenicity of STAT3 overexpression in the epidermis of transgenic mice (see below).[27] The demonstration that T cells are the primary regulators of keratinocyte proliferation and differentiation in psoriasis came from studies showing that supernatants from lesional skin-derived T cells transformed β-integrin$^+$ keratin 1/keratin 10$^-$ PCNA$^-$ stem cells of patients with psoriasis, but not stem cells of healthy subjects, into PCNA$^+$ active cycling cells. T cell–derived supernatants contained high levels of granulocyte macrophage–colony-stimulating factor (GM-CSF) and IFN-γ, and low levels of IL-3 and TNF-α.[28] However, among these cytokines, only IFN-γ affects the proliferation of psoriatic stem cells. *In vivo,* IFN-γ injection into prelesional psoriatic skin triggers keratinocyte proliferation and plaque development. Considering that IFN-γ is an antiproliferative cytokine and an inducer of squamous differentiation, these latter findings are paradoxical. This discrepancy may reflect an intrinsic defect in the response of psoriatic keratinocytes to IFN-γ, and/or altered localization and expression of the IFN-γ receptor complex in the epidermis of psoriatic skin.

Another T cell–derived cytokine that has been shown to regulate keratinocyte proliferation in psoriatic skin is IL-21, which is mainly released by CD4 T cells and natural killer T cells (NKT cells). IL-21 contributes to epidermal hyperplasia, and neutralization of IL-21 reduces both skin thickening and expression of inflammatory molecules. Interestingly, IFN-γ is necessary for IL-21–induced epidermal hyperplasia, while abrogation of IL-21 signals reduces IFN-γ expression in psoriatic T cells.[22,24]

Finally, the importance of T regulatory (Treg) lymphocytes (Chapter 18) in psoriasis has been examined in the peripheral blood and the inflamed skin of patients. The number of Treg cells (defined by expression of the transcription factor FOXP3) is increased in the peripheral blood of individuals with psoriasis, and this increase is positively correlated with the disease activity index. CD4$^+$CD25$^+$FOXP3$^+$ Treg cells are also present in psoriatic lesions, with more in lesional skin biopsy specimens than those from control or uninvolved skin. However, these Treg cells may be less functional in that Treg cells from both the peripheral blood and psoriatic skin lesions showed reduced ability to suppress effector T cells. This impairment may be dependent on IL-6 as blockade of IL-6 reversed the impairment in suppression observed in cocultures of Treg cells and effector T cells from patients with psoriasis.

Intrinsic Defects of Keratinocytes Are Fundamental for the Amplification of Psoriatic Processes

Endogenous defects in keratinocytes may be pathogenically relevant for psoriasis, as shown in mice with engineered epidermal phenotypes. Transgenic animals that overexpress the transcription factor STAT3 or that lack the inhibitor of NF-κB kinase-2 (IKK-2) in their epidermis develop skin lesions that closely resemble human psoriasis.[27] Similarly, the abrogation of JunB in keratinocytes triggers a skin phenotype with the histological features of psoriasis, including marked hyperplasia of the epidermis and dense dermal inflammatory cell infiltrates. The hyperplasia observed in these models may, in part, depend on overexpression of S100A8 and S100A9, two antimicrobial peptides with chemotactic activity and a recognized role in keratinocyte maturation and proliferation. The development of psoriatic lesions in mice with an epidermal deletion of STAT3 depends on the presence of activated T cells, whereas the inflammatory responses occurring in the skin of IKK2-transgenic mice are mediated by TNF-α. This implicates an intrinsically dysregulated interrelation between keratinocytes and cells of both the innate and acquired immune response in the pathogenesis of psoriasis. Within psoriatic lesions, alterations are observed in the levels of expression of several growth factors, such as insulin-like growth factor, keratinocyte growth factor (KGF), transforming growth factor (TGF)-α, and amphiregulin, all of which stimulate basal cell proliferation in an autocrine fashion.[22] Other cytokines aberrantly elevated in psoriasis include members of the inhibitory TGF-β family, and cytokines that stimulate keratinocyte proliferation and inflammation, such as IL-19 and IL-20.[22,24] An important pathogenic role has also been uncovered for IL-36γ, a cytokine abundantly induced by IL-17 in keratinocytes, whose overexpression in mouse skin leads to a disease quite similar to human plaque psoriasis, whereas inhibition in human psoriatic skin ameliorates the inflammation.[29]

Finally, two anti-inflammatory molecules, suppressors of cytokine signaling (SOCS)1 and SOCS3, are dysregulated in psoriatic keratinocytes.[30] These efficiently suppress the IFN-γ– and TNF-α–dependent molecular cascades in keratinocytes, and it has been hypothesized that strengthening of the action of SOCS1 in keratinocytes could be a valid therapeutic approach for treatment of IFN-γ- and TNF-α-dependent skin diseases, including psoriasis.

ON THE HORIZON

- There have been significant advances in therapies for psoriasis during the past 15 years that strongly reduce both symptoms and relapse rates. New therapeutic approaches include the targeting of interleukin (IL)-17A.
- Personalized therapies based on the employment of antagonists aligned with elements of genetic risk of the patients will be likely more successful.
- Newer strategies will also need to focus on treatments that do not require continuous, long-term immune suppression (*i.e.,* strategies to restore immune regulation or tolerance).
- Future work to identify and clarify risk factors and antigenic triggers for psoriasis may lead to strategies for preventing the disease.

CONCLUSIONS

A complex interplay between environmental and genetic factors triggers a cascade of events that leads to the expression of psoriasis. Early upstream events occurring in the disease include activation of DCs and the generation of effector T cells that migrate into the psoriatic skin lesions and expand there. Cross-talk between keratinocytes and immune cells amplifies inflammation and is responsible for chronicity. Recent research has implicated many immunological mechanisms in psoriasis progression, and this has led to the development of new, pathogenesis-based therapies. Although this progress is remarkable, much remains unknown, especially regarding prevention of the condition and how to develop drugs with appropriate risk-benefit and long-term profiles. Future work must take into account these aspects to establish therapeutic and preventive approaches that lead to improved patient outcomes.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Griffiths CE, Barker JN. Pathogenesis and clinical features of psoriasis. Lancet 2007;370:263–71.

2. Bowcock AM. The genetics of psoriasis and autoimmunity. Annu Rev Genomics Hum Genet 2005;6:93–122.
3. Strange A, Capon F, Spencer CC, et al. A genome-wide association study identifies new psoriasis susceptibility loci and an interaction between HLA-C and ERAP1. Nat Genet 2011;42:985–90.
4. Tian S, Krueger JG, Li K, et al. Meta-analysis derived (MAD) transcriptome of psoriasis defines the "core" pathogenesis of disease. PLoS ONE 2012;7:e44274.
5. Onoufriadis A, Simpson MA, Pink AE, et al. Mutations in *IL36RN/IL1F5* are associated with the severe episodic inflammatory skin disease known as generalized pustular psoriasis. Am J Hum Genet 2011;89:432–7.
6. Nestle FO, Kaplan DH, Barker J. Psoriasis. N Engl J Med 2009;361: 496–509.
7. Albanesi C, Scarponi C, Bosisio D, et al. Immune functions and recruitment of plasmacytoid dendritic cells in psoriasis. Autoimmunity 2010;43:215–19.
8. Colonna M, Trinchieri G, Liu YJ. Plasmacytoid dendritic cells in immunity. Nat Immunol 2004;5:1219–26.
9. Albanesi C, Scarponi C, Pallotta S, et al. Chemerin expression marks early psoriatic skin lesions and correlates with plasmacytoid dendritic cell recruitment. J Exp Med 2009;206:249–58.
10. Nestle FO, Conrad C, Tun-Kyi A, et al. Plasmacytoid predendritic cells initiate psoriasis through interferon-alpha production. J Exp Med 2005;202:135–43.
11. Yao Y, Richman L, Morehouse C, et al. Type I interferon: potential therapeutic target for psoriasis? PLoS ONE 2008;3:e2737.
12. Ketikoglou I, Karatapanis S, Elefsiniotis I, et al. Extensive psoriasis induced by pegylated interferon alpha-2b treatment for chronic hepatitis B. Eur J Dermatol 2005;15:107–9.
13. Lande R, Gregorio J, Facchinetti V, et al. Plasmacytoid dendritic cells sense self-DNA coupled with antimicrobial peptide. Nature 2007;449: 564–9.
14. Ganguly D, Chamilos G, Lande R, et al. Self-RNA-antimicrobial peptide complexes activate human dendritic cells through TLR7 e TLR8. J Exp Med 2009;206:1983–94.
15. Nestle FO, Di Meglio P, Qin JZ, et al. Skin immune sentinels in health and disease. Nat Rev Immunol 2009;9:679–91.
16. Zheng Y, Danilenko DM, Valdez P, et al. Interleukin-22, a T(H)17 cytokine, mediates IL-23-induced dermal inflammation and acanthosis. Nature 2007;445:648–51.
17. Costantini C, Calzetti F, Perbellini O, et al. Human neutrophils interact with both 6-sulfo LacNAc+ DC and NK cells to amplify NK-derived IFN-γ: role of CD18, ICAM-1, and ICAM-3. Blood 2011;117:1677–86.
18. Cumberbatch M, Singh M, Dearman RJ, et al. Impaired Langerhans cell migration in psoriasis. J Exp Med 2006;203:953–60.
19. Gudmundsdottir AS, Sigmundsdottir H, Sigurgeirsson B, et al. Is an epitope on keratin 17 a major target for autoreactive T lymphocytes in psoriasis? Clin Exp Immunol 1999;117:580–6.
20. Prinz JC. Psoriasis vulgaris—a sterile antibacterial skin reaction mediated by cross-reactive T cells? An immunological view of the pathophysiology of psoriasis. Clin Exp Dermatol 2001;26:326–32.
21. Lande R, Botti E, Jandus C, et al. The antimicrobial peptide LL37 is a T-cell autoantigen in psoriasis. Nat Commun 2014;5:5621–36.
22. Albanesi C, Pastore S. Pathobiology of chronic inflammatory skin diseases: interplay between keratinocytes and immune cells as a target for anti-inflammatory drugs. Curr Drug Metab 2010;11:210–27.
23. Eyerich S, Eyerich K, Pennino D, et al. Th22 cells represent a distinct human T cell subset involved in epidermal immunity and remodeling. J Clin Invest 2009;119:3573–85.
24. Albanesi C, De Pita O, Girolomoni G. Resident skin cells in psoriasis: a special look at the pathogenetic functions of keratinocytes. Clin Dermatol 2007;25:581–8.
25. Di Meglio P, Villanova F, Navarini AA, et al. Targeting CD8(+) T cells prevents psoriasis development. J Allergy Clin Immunol 2016;138: 274–6.
26. Sestito R, Madonna S, Scarponi C, et al. STAT3-dependent effects of IL-22 in human keratinocytes are counterregulated by sirtuin 1 through a direct inhibition of STAT3 acetylation. FASEB J 2011;25: 916–27.
27. Sano S, Chan KS, Carbajal S, et al. Stat3 links activated keratinocytes and immunocytes required for development of psoriasis in a novel transgenic mouse model. Nat Med 2005;11:43–9.
28. Bata-Csorgo Z, Hammerberg C, Voorhees JJ, et al. Kinetics and regulation of human keratinocyte stem cell growth in short-term primary ex vivo culture. Cooperative growth factors from psoriatic lesional T lymphocytes stimulate proliferation among psoriatic uninvolved, but not normal, stem keratinocytes. J Clin Invest 1995;95:317–27.
29. Gabay C, Towne JE. Regulation and function of interleukin-36 cytokines in homeostasis and pathological conditions. J Leukoc Biol 2015;97: 645–52.
30. Madonna S, Scarponi C, Sestito R, et al. The IFN-gamma-dependent suppressor of cytokine signaling 1 promoter activity is positively regulated by IFN regulatory factor-1 and Sp1 but repressed by growth factor independence-1b and Kruppel-like factor-4, and it is dysregulated in psoriatic keratinocytes. J Immunol 2010;185:2467–81.

MULTIPLE-CHOICE QUESTIONS

1. The age of onset of psoriasis is most commonly:
   - **A.** First three decades of life
   - **B.** Before age 7 years
   - **C.** Middle age
   - **D.** Fifth and sixth decades of life

2. Development of psoriatic lesions at sites of skin trauma is known as:
   - **A.** Koebner phenomenon
   - **B.** Nikolsky sign
   - **C.** Auspitz sign
   - **D.** Dennie-Morgan lines
   - **E.** Triple response of Lewis

3. Early pathogenic events of psoriasis includes which one of the following:
   - **A.** Signal transducer and activator of transcription 3 (STAT3) activation in hyperproliferating keratinocytes
   - **B.** Recruitment of myeloid dendritic cells
   - **C.** Targeting of keratinocytes by proinflammatory cytokines
   - **D.** A cascade of events including formation of complexes formed by keratinocyte-derived nucleic acid and the cathelicidin LL-37

65

Myasthenia Gravis

Arnold I. Levinson

Myasthenia gravis (MG) is a disease characterized by weakness of striated muscles. The weakness is caused by impaired neuromuscular transmission resulting from a reduction in the number of receptors for the neurotransmitter acetylcholine (ACh) at the postsynaptic myoneural junction. In most cases, this reduction is the result of the action of anti–acetylcholine receptor (anti-AChR) antibodies. The disease occurs with a reported prevalence of 0.5–5/100 000 and an incidence of 0.4/100 000/year. Although MG can occur at any age, it typically presents in the second and third decades of life, with a later peak occurring after age 50 years (late-onset disease). A female preponderance (3:1–4:1) has been reported in the first 40 years of life; thereafter, the incidence is comparable between the sexes.

CLASSIFICATION

Patients with MG have traditionally been divided into two categories: those with generalized disease and those presenting with disease limited to the ocular muscles.[1] Within these two groups, patients can be further subdivided on the basis of age of onset. Neonatal MG affects 10–20% of offspring born to mothers with myasthenia. Disease manifestations are those of generalized MG (see below) but are transient, dissipating with the metabolism of maternal anti-AChR antibodies that had been transmitted across the placenta during the third trimester of pregnancy. Several congenital myasthenic syndromes have been described. For the most part, these manifest during the neonatal period, persist into adulthood, and are not considered to have an autoimmune basis.[2] Juvenile MG is said to occur in those patients who present with disease between 1 year of age and puberty. Apart from the age of onset, juvenile myasthenia behaves like adult MG.

Adult patients may present with ocular involvement or signs of more generalized disease. The ocular involvement is characterized by impaired ocular muscle motility and lid weakness, manifesting as diplopia and ptosis, respectively. The vast majority of patients with MG will experience ocular involvement, with roughly 50% of patients presenting with ocular signs at the time of diagnosis. Those generally at risk of disease progression are *(i)* patients with evidence of subclinical disease on electrophysiological testing of limb muscles; and *(ii)* patients who have markedly elevated titers of anti-AChR antibodies. Typically, patients with ocular symptoms for >2 years will not progress to a more generalized form of disease.

In the generalized disease group, patients can be classified into those with mild, moderate, or severe disease on the basis of clinical activity. Any skeletal muscle group can be affected, but typically the palatal, pharyngeal, and upper esophageal muscles are involved. This results in dysarthria, dysphagia, and difficulty handling secretions. Involvement of the diaphragm and intercostal muscles produces dyspnea and may lead to respiratory failure. Involvement of the muscles of the extremities and trunk occurs in 20–30% of patients at initial presentation and causes difficulties with activities of daily living. The hallmark of all muscle involvement in MG is its variability over time, with weakness usually exacerbated by repetitive use.

Furthermore, within the group of patients with generalized disease, patients are further subdivided into subtypes on the basis of age of onset, for example, early-onset MG (EOMG) and late-onset MG (LOMG) with onset before and after age 50 years, respectively. Moreover, patients can be additionally distinguished by their profiles of serum autoantibodies (see discussion of serum autoantibodies) and the presence/absence of thymic pathology (see discussion of thymic pathology).

CLINICAL PEARLS

Telltale Signs of Myasthenia Gravis

- Variable muscle weakness
- Weakness in cranial nerve distribution
- Normal reflexes and sensation

DIAGNOSIS

The differential diagnosis is extremely broad, encompassing neuropathies, primary and secondary myopathies, muscular dystrophy, demyelinating disorders, degenerative diseases, cerebrovascular accidents, mass lesions, and infectious diseases. The clinical features that point to a diagnosis of MG include the variable nature of the muscle weakness, normal sensation, and normal deep tendon reflexes. The diagnosis can usually be confirmed by pharmacological and electrophysiological testing. A decremental pattern is characteristically seen following repetitive nerve stimulation (Fig. 65.1), and this pattern is normalized following treatment with the anticholinesterase agent tensilon. Further confirmation rests on detecting anti-AChR antibodies, which are found in 85–90% of patients with generalized disease. In the standard assay, sera are reacted with a nicotinic AChR preparation labeled with ^{125}I-α-bungarotoxin, a snake venom polypeptide that binds irreversibly to the receptor. Bound antibodies are immunoprecipitated by an antiimmunoglobulin reagent or staphylococcal protein A, and the quantity of antibodies detected is expressed in terms of the amount of α-bungarotoxin bound.

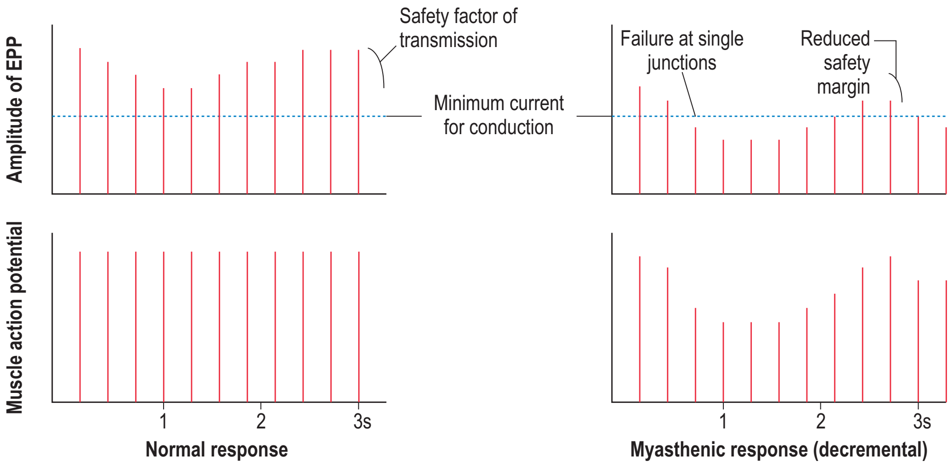


FIG 65.1 Neuromuscular Transmission in Normal and Myasthenic Subjects. With repetitive stimulation there is a reduction in the efficiency of acetylcholine (ACh) release, with a subsequent recovery in efficiency as the train of stimuli continues. Although the endplate potential (EPP) fluctuates at the normal junction, sufficient current is generated to stimulate an action potential of constant magnitude. At the myasthenic junction, however, the amplitude of the EPP in response to a given amount of ACh is reduced. Under conditions of inefficient ACh release, for example, repetitive stimulation, the minimum current for conduction is not generated, resulting in a profile of action potentials that shows a progressive decline or "decrement" with subsequent recovery.

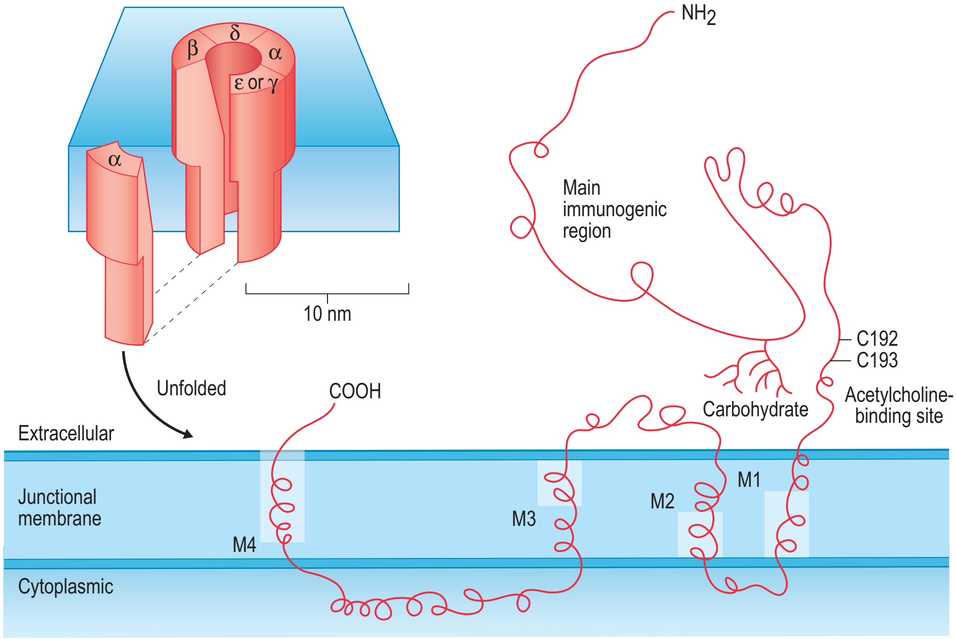


FIG 65.2 The Acetylcholine Receptor. The subunits of the acetylcholine receptor—α, β, δ, and γ or ε—are arranged like barrel staves around the central ion pore. Each subunit winds through the junctional membrane four times (sites M1, M2, M3, and M4). In the unfolded view of the α subunit, the amino-terminal end of the α subunit is extracellular, where it is accessible to acetylcholine, which binds at the site shown (amino acids 192 and 193). In myasthenia gravis, autoantibodies may bind to various epitopes of all subunits, but a high proportion of autoantibodies bind to the main immunogenic region of the α subunit.

ACHR STRUCTURE

The nicotinic acetylcholine receptor (nAChR) is a member of a larger family of ligand-gated ion channels. The muscle-type receptor, which is involved in myasthenia, can be subclassified further into mature junctional receptors and immature, extrajunctional, or denervated receptors. The nAChR at a mature myoneural junction is composed of four subunits, labeled α, β, δ, and ε (Fig. 65.2). In fetal muscle and adult denervated muscle or nonjunctional membrane, a γ subunit replaces the ε subunit found in mature innervated muscle endplates. This form of the receptor differs from the mature junctional form by its lower density (500 receptors/μm^2) and its distribution over most of the surface of the sarcolemma. The immature receptor also has a lower conductance, a longer open time, a more rapid turnover, and a decreased half-life.

The genes for the α, δ, and γ subunits are located on chromosome 2 in humans, and subunits β and ε on chromosome 17. The subunits of the AChR are homologous to each other and to their counterparts across species, with the greatest conservation of sequence being in the α subunit. Two α subunits and one of each of the other subunits are assembled to form an asymmetrical hourglass channel spanning the membrane. Each subunit has a large amino-terminus located extracellularly, four transmembrane regions, and a short cytoplasmic tail formed by a loop between the third and fourth transmembrane domains. The receptor appears as a dimer as a result of disulfide bonding between the δ subunits of two receptors. The two α subunits are not contiguous in each receptor but are separated by another subunit. One ACh-binding site is found on each of the α subunits around the pair of cysteines at amino acids 192 and 193. The binding of ACh to the α subunits is believed to engender a conformational change, possibly resulting in rearrangement of charged groups. The binding of ACh to both α subunits increases the probability of transition of the channel to an open conformation. Binding of curare or α-bungarotoxin to the α subunits blocks this channel.

In normal innervated neuromuscular junctions, there are two forms of the AChR, the predominant form having a long half-life, and a small subset that is rapidly turned over. The rapidly turned-over receptors are the precursors of the stable receptors. It is not clear how these two types differ or how they are regulated. The receptors are concentrated at the top of the folds in the muscle endplate, adjacent to the nerve terminus, at a density of 10 000/μm^2. This localization reflects the action of agrin, a nerve-derived synaptic organizing molecule. The AChRs are organized into clusters by rapsyn, a 43-kilodalton (kDa) cytoplasmic protein. The clustered AChRs are linked to the cytoskeleton by connections between rapsyn and a dystrophin–glycoprotein complex.

NEUROMUSCULAR TRANSMISSION

When an impulse is transmitted along an axon terminal, it results in the release of the neurotransmitter ACh across its presynaptic membrane (Fig. 65.3). ACh diffuses across a 50-nm synaptic cleft, where it interacts with AChRs, which are displayed in greatest density at the tops of the junctional folds of the postsynaptic muscle membrane or endplate. This interaction leads to a local depolarization or endplate potential caused by increased membrane permeability to sodium and potassium. The endplate potential is terminated by acetylcholinesterases, which are present in highest concentrations in the synaptic cleft around the junctional folds. If the summation of endplate potentials attains a prescribed threshold, it produces an action potential that depolarizes the surrounding sarcolemma and causes muscle contraction. In a healthy individual, the arrival of an impulse at the presynaptic membrane of a motor nerve releases considerably more ACh than is required to generate an action potential. This reserve, roughly four times the current needed for propagation of the impulse, is referred to as the *safety factor of neuromuscular transmission.* Because of the severe reduction in receptor number in MG, the electrical threshold for propagation of an action potential cannot be attained and muscle contraction is prevented. With a less severe reduction in receptor numbers neuromuscular transmission may proceed normally unless the efficiency of presynaptic vesicle release is compromised, as occurs with repetitive use of muscles. The combination of decreasing availability of ACh and the reduced number of receptors accounts for the characteristic decremental nerve conduction pattern seen on electromyograms of patients with MG following repetitive nerve stimulation (see Fig. 65.1).

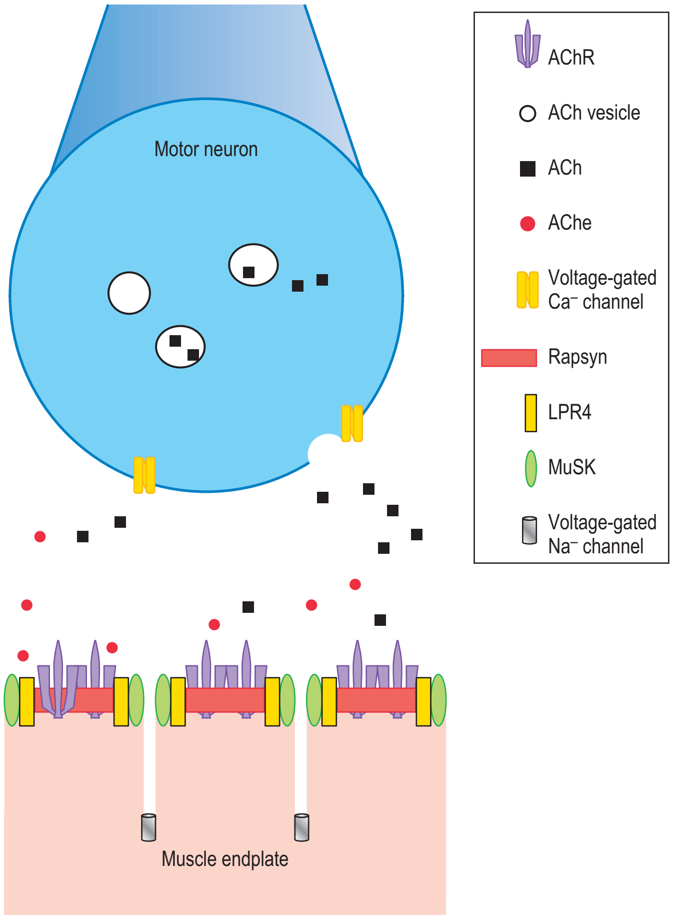


FIG 65.3 **Schematic Representation of the Myoneural Junction.** Vesicles of acetylcholine (ACh) release their contents at active zones across from acetylcholine receptors (AChRs) in response to impulses conducted down nerve axons. ACh diffuses across synaptic cleft and binds to AChRs, with opening of the ion channel and the generation of endplate potential. Action potential is propagated to muscle when sufficient amplitude of summated endplate potentials is attained. MuSK, muscle-specific tyrosine kinase.

IMMUNOPATHOGENESIS OF MG

KEY CONCEPTS

Involvement of Anti–Acetylcholine Receptor (AChR) Antibodies in the Pathogenesis of Myasthenia Gravis (MG)

- AChR antibodies are found in the serum of 85–90% of patients with MG.
- Infants born to mothers with myasthenia sometimes develop MG.
- Immunoglobulin G (IgG) and complement are deposited at the postsynaptic junction.
- Transfer of serum IgG from patients with MG to mice induces neuromuscular blockade.

Anti-AChR antibodies are detected in 85–90% of MG patients and are responsible for the impaired neuromuscular transmission.[3] There are several lines of evidence to support this contention. Immunoglobulin G (IgG), along with C3 and the terminal attack complex (C5–C9), is deposited at AChR-containing areas of the postsynaptic membrane, and anti-AChR/AChR complexes can be extracted from the muscles of patients with MG. Transfer of myasthenic serum from mother to fetus, or from human to mouse, results in symptoms or signs of myasthenia in the recipient. Plasmapheresis, which decreases anti-AChR antibody levels, is associated with clinical improvement.

Properties of Anti-AChR Antibodies and Characterization of B-Cell Epitopes

Anti-AChR antibodies are produced by a small subset of B cells in affected people. The frequency of IgG-producing AChR-specific peripheral blood mononuclear cells is estimated to be 1 in 15 000–70 000. IgG anti-AChR–secreting cells are also found in the peripheral blood of healthy volunteers, albeit in much lower numbers. The proportion of immunoglobulin producing AChR-specific B cells or plasma cells is greater in the germinal centers of hyperplastic thymuses, but still only 1 in 1000–10 000 antibodies produced is AChR specific. Anti-AChR antibodies are predominantly IgG1 and IgG3, but IgG2 and IgG4 isotypes have also been found. IgA and IgM anti-AChR antibodies are present in some patients, but never in the absence of IgG anti-AChR antibodies. The IgA and IgM anti-AChR antibodies tend to appear in patients whose disease is of longer duration and greater severity and in association with high IgG anti-AChR titers.

The pathogenic anti-AChR antibodies in MG are thought to be directed to conformationally dependent structures. Immunization of animals with irreversibly denatured AChR leads to the formation of anti-AChR antibodies capable of binding to native AChR, but the antibodies are not capable of causing disease. This observation indicates that conformationally dependent epitope(s) are important in the induction of disease. Many of the anti-AChR antibodies are directed against the α subunit, particularly to a small region on the extracellular portion referred to as the *main immunogenic region*[4] (see Fig. 65.2). Approximately 60% of the anti-AChR antibodies are directed against this region, which encompasses a set of overlapping epitopes clustered around amino acids 67–76 of the α subunit.[5] The reason for the predominant role of the α chain in the antibody response in myasthenia is not known. Not all disease-producing antibodies in humans or rats appear to be directed to this region. Many patients also have antibodies recognizing the γ-containing embryonic form of AChR. This observation has spawned speculation about a nonmuscle source of sensitization.

Anti-AChR Antibody Levels and Relationship to Disease Activity

The relationship of anti-AChR antibody and disease activity in MG is complicated. In general, serum levels of anti-AChR antibody or anti-AChR/AChR complexes correlate poorly with disease severity. Among patients within one clinical grade, anti-AChR antibody levels can vary by several orders of magnitude. In addition, approximately 10–15% of patients with clinical MG have no anti-AChR antibody by standard assays. Anti-AChR antibodies are more likely to be present if the disease is generalized or severe. In addition, in an individual patient an increase or decrease in anti-AChR antibody levels often accompanies deterioration or improvement, respectively, in clinical activity.

As the quantity of anti-AChR antibodies produced does not fully explain disease severity, studies have also focused on qualitative differences in these antibodies among different patients. These studies, however, have not been able to distinguish properties of anti-AChR antibodies that lead to greater pathogenicity. Differences in specificity and avidity of binding to AChR have not been associated with particular functional effects or disease severity. Anti-AChR antibodies that bind or compete for the same region of AChR can have different functional effects (see anti-AChR blocking antibody discussion). Anti-AChR antibodies show extensive heterogeneity by isoelectric focusing, but this characteristic also does not correlate with pathogenic potential.

Patients previously classified as seronegative have been found to have low-affinity AChR-specific IgG antibodies when their sera were tested with very sensitive assays that rely on tissue substrates on which the receptors are aggregated.[6,7] Antibody binding was facilitated by the high concentrations of AChRs, which compensate for the low affinity of the autoantibodies. Like the anti-AChR antibodies in classic anti-AChR antibody-positive MG, these antibodies belong largely to the IgG1 subclass. Like the conventional antibodies, the low-affinity antibodies also bind complement.

As noted above, 10–15% of MG patients are persistently antibody negative. Approximately 50% of these patients have been discovered to have serum IgG antibodies specific for *mu*scle-*s*pecific tyrosine *k*inase (MuSK).[7,8] This skeletal muscle receptor tyrosine kinase is activated by agrin and is critical for formation of the neuromuscular junction. Originally, anti-MuSK antibodies were found in patients with marked facial and bulbar weakness, including tongue weakness and respiratory involvement with relative sparing of upper- and lower-extremity muscles. Ophthalmoparesis was also seen but was not a first symptom in at least one series of patients. Subsequently, it has become apparent that patients with anti-MuSK antibodies can also have a more traditional clinical phenotype similar to that seen with anti-AChR patients. Patients often respond poorly to anticholinesterase agents and benefit greatly from plasmapheresis but only modestly from intravenous immunoglobulin (IVIG). They typically lack thymus pathology and, not surprisingly, thymectomy is not typically beneficial. Most anti-MuSK antibodies are of the non–complement-fixing IgG4 isotype. Anti-MuSK antibodies do not cause destruction of the muscle endplate, as might be expected because of their inability to activate the classical complement pathway and reduced capacity to bind activating Fcγ receptors on monocytes and macrophages. However, they have been shown to affect AChR clustering on cultured myotubes. The clinical relevance of these autoantibodies is underscored by the appearance of a myasthenic illness in animals either immunized with MuSK or infused with serum antibodies from patients with anti-MuSK antibody–associated MG.

IgG antibodies specific for lipoprotein-related protein 4 (LRP4) have also recently been discovered in the sera of patients with anti-AChR antibody–negative MG.[7,9] LRP4 is an important component of the neuromuscular junction, which serves as a receptor for agrin and which is required for agrin-induced activation of MuSK and AChR clustering. The prevalence of reactivity to LRP4 varied from 2–50% of patients who are negative for anti-AChR and MuSK antibodies (double seronegative), with the differences possibly being related to geography or ethnicity.

Anti-LRP4 antibodies were found in the sera of 3% of patients with anti-MuSK antibodies in one study but otherwise were limited to double-seronegative (anti-AChR$^-$/anti-MUSK$^-$) patients. Anti-LRP4 antibodies were not present in the sera of a large number of patients with other neurological diseases with the exception of neuromyelitis optica spectrum disorder, where 12.5% of patients had anti-LRP4 antibodies. If larger studies demonstrate that anti-LRP4 antibodies are found in patients with MG but without anti-AChR or MuSK antibodies and are rare in other neurological disorders, this test could become important in the diagnosis of MG.

MG sera also contain antibodies reacting with the ryanodine receptor.[7] These receptors, which are critically involved in muscle contraction, are Ca^{2+} release channels located in the sarcoplasmic reticulum of striated muscles. Antibodies to ryanodine receptors are found in 50% of patients with MG who have thymoma, and patients with high levels have a worse prognosis than do antibody-negative patients with MG who have thymoma. *In vitro* studies have suggested a pathogenic role for these autoantibodies in MG.

Antirapsyn antibodies have been detected in a small subset of patients with MG.[7] Seropositive patients with MG are indistinguishable from seronegative patients with regard to clinical and laboratory features of disease. The presence of antirapsyn is not specific for MG, having been detected in the sera of an occasional patient with MS and a majority of the patients with lupus tested.

Antiagrin antibodies have recently been reported, often in MG patients with anti-AChR or anti-MuSK antibodies and occasionally in patients with MG who are “triple seronegative” (AChR$^-$ MuSK$^-$ LRP4$^-$).[7] Studies of larger numbers of patients will be necessary to determine the prevalence of antiagrin antibodies in triple seronegative MG patients and their degree of specificity for MG.

Patients with MG associated with thymoma show distinct patterns of antibody production. Almost all patients with thymoma are anti-AChR antibody positive, and most produce antistriational antibodies.[9] These latter antibodies react with titin, a giant filamentous protein of striated muscle. Titin filaments are involved in muscle assembly and contribute to the muscle’s ability to recoil following stimulation. Such antistriational antibodies are also found in approximately 50% of the sera of older patients with MG who have thymic atrophy but not thymoma; however, they are not frequently detected in patients with early-onset disease and thymic hyperplasia. The finding of antistriational antibodies in a patient with MG who is less than 40 years old strongly suggests the presence of thymoma. There is no evidence that antistriational antibodies are involved in muscle weakness.

Pathogenic Effects of Anti-AChR Antibodies

KEY CONCEPTS

- Effects of anti–acetylcholine receptor (AChR) antibodies in myasthenia gravis pathology
  - Reduced number of receptors
  - Widening of the synaptic cleft
  - Distorted geometry of the synaptic membrane
- Mechanism of damage
  - Complement-dependent damage to muscle endplate
  - Enhanced rate of AChR degradation
  - Block cholinergic binding sites

Complement-Mediated Damage

The critical problem in MG is the anti-AChR antibody–mediated reduction in the number of nAChRs at the myoneural junction. There are several possible mechanisms by which the anti-AChR antibodies could lead to impaired neuromuscular transmission.[10] Ultramicroscopic studies show marked destructive changes in some endplates, particularly at the peaks of the postsynaptic folds, where AChR is usually present in the greatest concentration. The architecture of the muscle endplate is simplified, with loss of junctional folds and widening of the synaptic cleft that contains membrane debris. C3, C9, and the membrane attack complex are deposited at the muscle endplate, suggesting a role for complement in membrane destruction.[11] Indeed, in many patients, anti-AChR antibodies can fix complement *in vitro* when bound to skeletal muscle and can damage cultured rat myotubes with a resultant decrease in AChR content.

Although antibody-directed, complement-mediated destruction is important in the pathophysiology of MG, it is not the entire story. The rapid clinical improvement in MG following certain therapeutic interventions and the lack of destructive changes in many neuromuscular junctions of symptomatic areas despite prominent immunoglobulin deposition suggest that a more readily reversible process is also likely to be involved in the neuromuscular block.

Acceleration of AChR Degradation

In vitro and *in vivo* studies have shown that anti-AChR antibodies can accelerate the rate of degradation of extrajunctional and junctional receptors, respectively.[12] This reaction is complement independent and results from the endocytosis of AChRs via shallow depressions, presumably clathrin-coated pits. Other membrane receptors are not affected. Both stable and rapidly turned-over receptors appear to be affected, thereby explaining the greater than expected antibody-mediated loss of AChRs observed at neuromuscular junctions. The reaction requires cross-linking of adjacent AChRs, as it can be mediated by $F(ab')_2$ fragments, but not Fab fragments, of anti-AChR antibodies. The effect of antibody on the synthesis of new AChR is controversial.

Receptor Blockade

The inhibition of ACh binding has been assessed by studying the effects of MG serum on the binding of the neurotoxin α-bungarotoxin to AChRs. Such blocking antibodies are found in a variable number of MG sera. Blockade has been generally attributed to steric hindrance of the ligand-binding site, rather direct binding to the ACh-binding site.[13] The importance of these antibodies in the pathophysiology of MG remains unclear. However, in one study the functional ability of an individual serum to accelerate degradation and cause blockade of AChRs paralleled most closely the clinical status of the patient. The *in vivo* significance of such antibodies has also been demonstrated by passive transfer of certain rat monoclonal anti-AChR antibodies into chicks. Complete paralysis was observed within 1 hour of the transfer, presumably before there was time for complement-mediated damage. It has been reasoned that in patients with MG, such blocking antibodies could further diminish synaptic function already decreased owing to complement-mediated damage or accelerated receptor degradation. This could result in acute clinical deterioration or a rapid clinical improvement after plasmapheresis, before the repair of damaged membrane and regeneration of new AChRs.

Role of T Cells

An overwhelming body of data indicates that anti-AChR antibody production in patients with MG and rodents with experimental autoimmune myasthenia gravis (EAMG) is dependent on the activity of CD4 T cells. Researchers have used freshly isolated CD4 T cells, T-cell lines and T-cell clones stimulated with AChR purified from the electric organs of *Torpedo californica* (T-AChR), recombinant human AChR, human AChR concentrated on immunomagnetic beads, and synthetic AChR peptides. Most of the T cells derived from patients with MG have responded to the complete AChR or the α subunit, but responses to the δ, ε, and γ subunits have also been demonstrated.[14] Unlike EAMG, a T cell–immunodominant epitope has not been found in humans. However, T cells from most patients recognize a limited number of AChR sequences. In an individual patient, T cells reactive with a particular AChR epitope can show limited T-cell receptor Vβ (TCR Vβ) usage, but the same TCR Vβ usage has not been observed for different patients. A range of major histocompatibility complex (MHC) class II molecules can present AChR epitopes. This is not surprising, given the degenerate binding capabilities of MHC molecules needed to present an extremely large and diverse group of peptides to the immune system. T cells from normal controls respond to some of the same epitopes to which MG patients' T cells respond, albeit in smaller numbers and less vigorously.

EXPERIMENTAL AUTOIMMUNE MYASTHENIA GRAVIS

The serendipitous discovery that rabbits injected with purified AChR developed not only anti-AChR antibodies but signs of MG provided the evidence for the pivotal role of antibodies in this disease and paved the way for the development of an experimental model of human MG.[15] EAMG has been studied in a number of different animal species.[16] It is generally induced by immunizing animals with AChR purified from the electric organs of *T. californica* emulsified in adjuvant. Disease can also be passively transferred by immunoglobulin from affected animals or from patients with myasthenia. Animals suffer from fatigue, hypoactivity, weight loss, paralysis, difficulty breathing, and dysphagia; these signs are reversible with anticholinesterases. There is a decremental response to repetitive nerve stimulation on electromyography in EAMG as in MG. Anti-AChR antibodies are present, and most are directed against the main immunogenic region (see Fig. 65.2). They are deposited along with complement at the muscle endplate. Similar to the pathology in humans, there is simplification of the muscle endplate and a loss of AChR content of muscles in the experimental disease. A major difference between the induced rodent models of EAMG and spontaneous human MG is the absence of thymic pathology in the former.

In murine models, there are disease-resistant and disease-susceptible strains. The MHC determines this susceptibility, in part, with specific H-2 alleles being associated with susceptibility or resistance. However, even in the high-responder strains, only 50–70% of immunized animals manifest disease. This lack of concordance remains to be fully explained. Differences in the level of complement activation and the physiological safety factor of neuromuscular transmission may partly account for the variability in disease expression.

Most immunized animals demonstrate anti-AChR antibodies, whether or not they develop EAMG. Induction of pathogenic antibody production requires the integrity of the three-dimensional structure of the receptor, as immunization with denatured AChR elicits antibody production but not disease, even in susceptible animals. Anti-AChR antibody levels correlate with anti-AChR/AChR complexes and with the AChR content of the muscle, but not with clinical disease severity. No characteristic feature distinguishing disease-producing and nonpathogenic autoantibodies has been defined. Antibodies in rats with and without disease are predominantly of the same immunoglobulin isotope and subclass, express similar clonotypic heterogeneity, and show similar avidities. The percentage of antibodies that recognize the native receptor (as opposed to the foreign eel receptor immunogen) is low (0.2–2%) but similar between the symptomatic and the asymptomatic mice. The percentage of antibody directed against the main immunogenic region of AChR is also similar. The antibodies that develop in asymptomatic animals can bind and cause a decrease in receptor half-life, as well as those that develop in paralyzed animals. In fact, EMG abnormalities can be demonstrated in mice that appear disease free. An analogous situation in humans may be the high frequency of single fiber EMG abnormalities in asymptomatic family members of patients with MG. Importantly, antibodies from affected animals will cause disease in resistant strains, indicating that disease differences between strains is not caused by differences in the AChR receptor itself.

Although EAMG is mediated by autoantibodies, it is a T cell–dependent process. Examination of T-cell responses in rats and mice has shown that EAMG requires the action of CD4 T cells. These cells recognize immunodominant T-cell epitopes located on the α subunit. The T-cell responses to the immunodominant epitope in rats involve a variety of TCRVβ and Jβ gene segments (Chapter 4). However, in susceptible C57BL/6 mice, there is a predominant use of the TCRVβ6 family in the CD4 T-cell response. CD4 T cells in susceptible and resistant strains recognize different immunodominant AChR α-chain epitopes.[17]

Early reports suggested that AChR-reactive T-cell clones did not fit neatly into T-helper-1 (Th1) and Th2 subsets. However, later studies suggested that interferon-γ production by Th1 cells is essential for the development of disease and that IL-4 may subserve a protective role.[16,18] In addition to interferon-γ, tumor necrosis factor (TNF)-α, interleukin (IL)-6, IL-12, IL-18, IL-1, IL-5, and IL-10 have been reported to promote the development of EAMG. However, a more recent publication indicated that C57BL/6 mice genetically deficient in IL-12/IL-23 and IFN-γ are susceptible to experimental EAMG, suggesting a pathogenic role of non-Th1 cells. This result led the authors to suggest that AChR-specific Th17 cells may also contribute to the pathogenesis of this experimental autoimmune disease.[19] Subsequently, Th17 cells were, indeed, shown to be important in the pathogenesis of EAMG.[20]

The roles of a number of costimulatory factors have also been investigated in EAMG. CD28/B7 and CD40/CD40L appear to be required for the development of the primary immune response to AChR. An interaction between ICOS and B7RP-1 may play an important role in the secondary response and/or in the maintenance of the immune response to this autoantigen.

THE THYMUS IN MYASTHENIA GRAVIS

Several lines of evidence have suggested that the thymus is directly linked to the pathogenesis of MG.

KEY CONCEPTS

Pathogenic Roles of the Thymus in Myasthenia Gravis (MG)

- Pathological
  - 65–75% of patients with MG have follicular hyperplasia with germinal centers
  - 10% have thymoma
- Clinical
  - Improvement following thymectomy
- Immunological
  - Acetylcholine receptor (AChR) subunits expressed on myoid cells and thymic epithelial cells
  - AChR-reactive T and B cells localized in the thymus
  - Increased thymic AChRα expression
  - Anti-AChR antibody secreted by thymic B-lineage cells
  - Decreased thymic T regulatory cell (Treg) function
  - Interleukin (IL)1-β, IL-6, CXCL13, CCL21, and B cell–activating factor (BAFF) overexpressed in thymus

Thymic Pathology

Interest in a primary role for the thymus in the pathogenesis of MG has been fueled by pathological, clinical, and immunological lines of evidence, although the nature of its involvement remains to be elucidated.[21-24] The thymus is pathologically abnormal in 80–90% of patients with MG. The majority of patients (65–75%) have thymic follicular hyperplasia with germinal center formation. The architecture of the hyperplastic thymi is generally preserved, with well-demarcated cortical and medullary regions. However, the medulla is crowded by numerous germinal centers that display the architectural features and cellular constituents of germinal centers in the secondary follicles of peripheral lymph nodes from normal individuals. Although these germinal centers in patients are generally thought to occupy an intraparenchymal position, some observers feel that they may actually lie extraparenchymally in the perivascular space.

Thymomas are seen in approximately 10% of patients with MG who tend to be older than those with hyperplastic thymi.[21-24] The thymomas are characterized by a loss of cortico-medullary demarcation and consist largely of neoplastic epithelial cells admixed with thymocytes. The affected epithelial cells belong to the cortical epithelial compartment, and the thymocytes have the immunophenotypic properties of normal immature cortical thymocytes.

Further evidence for a pathogenic role of the thymus comes from the results of empiric trials of thymectomy. Despite the absence of controlled clinical trials until recently,[25] there was general agreement in the past that removal of the thymus, particularly in young patients with follicular hyperplasia, leads to clinical improvement. The underlying basis of this improvement remains unknown.

Intrathymic Factors Possibly Contributing to Local Anti-AChR Antibody Response

The MG thymus, particularly hyperplastic thymus, is characterized by several unique features that strongly suggest its primary role in the immunopathogenesis of MG.[21-24] The thymus contains important constituents necessary for and indicative of an immune response directed against nAChRs. There is considerable evidence that resident cells in the thymus, including myoid cells and medullary epithelial cells, express various subunits of AChR including the α subunit. Additional factors found uniquely in hyperplastic MG thymus include increased expression of chemokines that attract immigrant CD4 T and B cells (CXCL13, CCL21); the presence of nAChR-reactive B and CD4 T cells; cytokines that can facilitate B-cell activation, differentiation, and survival (IL1β, IL6, APRIL, BAFF); anti-AChR antibody–secreting plasma cells; and possibly decreased $CD4^+CD25^+$ T regulatory cell function.[21-24]

As yet, it is not known whether or how such perturbations of the thymus lead to a breach in self-tolerance and the induction of anti-AChR antibodies, although there is mounting evidence that an antecedent inflammatory reaction in the thymic medulla may serve as the initiating trigger. An understanding of this enduring mystery likely holds the key to unlocking the immunopathogenesis of this disease. The reader is referred to other works for an in-depth discussion of this topic.[22,24,26]

ETIOLOGICAL FACTORS

Genetic Factors

As in most autoimmune diseases, the MHC represents an important genetic susceptibility locus for the development of MG. Studies indicate that the extended human leukocyte antigen (HLA)-A1-B8-DR3 haplotype is associated with EOMG and hyperplastic thymus in Caucasian individuals. This haplotype has been associated with the development of other autoimmune disorders. An association with HLA-B7-DR2, although weaker, has also been described in patients with onset of MG occurring after the age of 40 years and associated with atrophic thymic histology. However, the strongest association, focused on a Norwegian population of patients with late-onset MG, was recently shown to be the DRB1*15:01 allele. A recent study found a strong association between the HLA-DQ5 allele and patients with anti-MuSK antibody–positive MG. Interestingly, the MuSK-positive T-cell repertoire appears to be skewed to the usage of this HLA MHC class II allele. In murine studies, the MHC class II molecules I-A^b and I-E^k have been associated with susceptibility to EAMG.[16] In EAMG, the permissive MHC class II molecules are capable of binding AChR peptides that are recognized by antigen-specific CD4 T cells.

Studies have also addressed the potential genetic contributions of other immune system–related genes in the pathogenesis of MG. An association has been reported for a particular IL-1β allele and MG and increased serum levels of this cytokine have been reported. This was most pronounced in patients who lacked disease-susceptible HLA genes. Several groups have reported an association between MG and the presence of particular TNF-α polymorphisms. The expression of a high-transcription TNF-α allotype, TNF-α-308 allele 2, correlated with EOMG. In this regard, patients with MG demonstrate increased serum levels of TNF-α and their peripheral blood mononuclear cells display increased expression of TNF-α messenger RNA (mRNA). Polymorphisms in the IL-10 promoter region have been reported to be associated with distinct patterns of thymic histology. No correlations have been made between IL-4 alleles and MG. Allotypic markers on IgG and FcγRIIA receptors have been associated with the coexistence of MG and thymoma. A single nucleotide polymorphism in the gene encoding the intracellular tyrosine phosphatase PTPN22, which has been associated with the risk of developing other autoimmune diseases, has also been identified in both a subgroup of patients with anti-titin antibody–positive nonthymomatous MG and patients with thymomatous

MG.[27,28] The risk of developing MG has also been linked to a polymorphism in the promoter region of *CHRNA1,* the previously mentioned gene that encodes the α subunit of AChR and *CHRND,* the gene that encodes the δ subunit.[22] It has been proposed that such polymorphisms in genes encoding autoantigens cause reduced expression of AChRα on thymic medullary epithelial cells, thereby impairing central deletion of autoantigen-specific thymocytes. Additional MG-associated risk factors include polymorphisms in *CTLA4*, type II, *IFNII*, *IL12*, *CD86, AKAP12, VAV1, BAFF, TCF19,* and *TNIP* genes.[22,23]

Exogenous Factors

Whether or not sensitization to AChR occurs in the thymus or the periphery, the stimulus for this autoimmune response remains a conundrum. Moreover, it remains to be determined whether the stimulus is a self-antigen (AChR) or a foreign antigen that mimics the receptor's molecular structure. In this regard, several examples of molecular mimicry between AChRα chain and other molecules have been reported. Studies carried out with certain monoclonal anti-AChR antibodies demonstrated epitope sharing between the receptor and several bacteria, including *Klebsiella pneumoniae, Escherichia coli, Proteus vulgaris,* and *Yersinia enterocolitica.*[29] However, for the most part, no difference was observed in the binding of polypeptides from these organisms by either sera from patients with MG or control sera. A computer search of protein banks revealed a sequence homology between AChRα chain and a short peptide in herpes simplex glycoprotein D,[30] although the significance of this finding is unknown. Finally, similarities were reported between idiotypic determinants on anti-AChR antibodies and antibodies reactive with α1,3-dextran. Interestingly, antidextran antibodies were detected in approximately 13% of patients with MG but rarely in normal controls.[31] α1,3-Dextran is found in the cell walls of several common enteric pathogens and thus represents a potential ubiquitous source of immunogen. This type of idiotypic network connectivity led the investigators to postulate that an unregulated antiidiotypic response to anti-α1,3-dextran antibodies might lead, in certain individuals, to an anti-AChR antibody response. Unfortunately, there has been no follow-up to these observations.

A striking association has been reported between the development of MG and treatment with the drug penicillamine,[32,33] particularly in individuals with HLA-DR1. MG developed in patients with rheumatoid arthritis and patients with Wilson disease treated with this agent. After discontinuation of penicillamine, resolution of MG symptoms was reported in some patients but not others. Penicillamine treatment was associated with the development of anti-AChR antibodies that appeared to have the same type of specificity profile as found in idiopathic myasthenia. Additional evidence suggests that penicillamine may directly interfere with neuromuscular transmission. Although penicillamine has been shown to have diverse effects on the immune response in the normal host and has reactive sulfhydryl groups capable of modifying self-antigens, its role in the development of MG remains to be determined.

Currently, considerable attention is being focused on the role of epigenetic mechanisms to explain how environmental factors may promote the development of autoimmune diseases. Although some key epigenomic mechanisms, such as DNA methylation, histone acetylation, and microRNAs, have been demonstrated to possibly play a pathogenic role in several autoimmune disorders, the analysis of these factors is at a rudimentary stage in MG. To date, there is evidence to support an association between miR155 and the development of EOMG and human MG. An interesting story has emerged concerning miR150-P. Increased levels of this miRNA were found in the serum of patients with MG, and these levels fell in association with the clinical improvement that followed thymectomy.[34] A decrease in miR320a was observed in the peripheral blood mononuclear cells of a cohort of Chinese patients.[22] This finding was associated with increased levels of a number of proinflammatory cytokines. Further studies of the role of miRNAs will need to be conducted to elucidate whether they contribute to the immunoregulatory abnormalities seen in MG.

TREATMENT OF MYASTHENIA GRAVIS

THERAPEUTIC PRINCIPLES

- Anticholinesterase agents
- Corticosteroids
- Thymectomy
- Plasmapheresis
- Immunosuppressive agents
- Intravenous immunoglobulin

Therapeutic intervention in MG usually proceeds in a stepwise manner, beginning with anticholinesterase agents.[35] Most of the experience dealing with therapeutic modalities is based on treating patients with anti-AChR-associated MG.

Anticholinesterases

Anticholinesterases are the mainstay of treatment. These agents protect acetylcholine from hydrolysis by cholinesterase, thereby increasing the amount of neutrotransmitter and the number of contacts with the reduced number of receptors at the postsynaptic junction. This, in turn, raises the probability of attaining the necessary threshold for neuromuscular transmission. In addition, some of the anticholinesterase agents have a direct agonist effect at the postsynaptic junction. The three most popular agents in this group are neostigmine bromide (Prostigmin), pyridostigmine bromide (Mestinon), and ambenonium chloride (Mytelase). Although there are only slight differences between these agents, Mestinon remains the most commonly used. It has an onset of action of 30–60 minutes, peak action at about 2 hours, and loss of activity after 4 hours. Adverse effects of these agents are caused by excessive stimulation of nicotinic and muscarinic receptors. Auxiliary drugs that have been purported to have a salutary effect on neuromuscular transmission are ephedrine and xanthine derivatives (theophylline), which are thought to increase the presynaptic release of ACh. The minimal effect of their added benefit has not warranted their common usage. As mentioned previously, anticholinesterase agents often are not beneficial in patients with anti-MuSK-antibody–associated MG and may even exacerbate weakness. These agents neither induce sustained remission of symptoms nor impede disease progression.

Thymectomy

Another mainstay in the therapy of the adult with generalized MG is thymectomy.[36,37] The benefit is greatest in younger patients and those with thymic hyperplasia, although many centers include older patients as well. Over many years, despite the absence of a controlled study, there was general agreement that removal of the thymus leads to clinical improvement, particularly in young patients with follicular hyperplasia. In one study, 90% of patients were asymptomatic or in complete remission within a few years of thymectomy, and 46% were off all medications. An international multicenter controlled trial completed recently compared clinical outcome in patients with MG treated with corticosteroid versus those treated with corticosteroid plus thymectomy and those receiving thymectomy demonstrated a more favorable clinical outcome.[25] With improvements in preoperative care, anesthesia, surgical technique, and postoperative care, thymectomy has become a safe procedure, but its value and safety in children and older patients is less well established. There is still some controversy over what represents the best surgical procedure. The mechanism responsible for the salutary effect of thymectomy remains to be elucidated. No obvious effects on immunoregulatory mechanisms have been demonstrated, although anti-AChR titers tend to fall months after the procedure. Thymectomy is also the recommended treatment for patients of all ages suspected of having thymoma.

Corticosteroids

Corticosteroids are used in patients with generalized MG who fail to respond to anticholinesterase agents or thymectomy and in patients needing optimization of their clinical condition in preparation for thymectomy.[38] They are generally not used as first-line agents to replace thymectomy but are used in patients with ocular myasthenia who fail to respond to anticholinesterases. Corticosteroids are initially given on a daily basis, with therapy initiated in a hospital. This cautious approach is followed because of the fear of clinical deterioration that may occur in some patients during the introduction of corticosteroids. This concern resulted in some groups advocating initiation of alternate-day therapy, which is not typically associated with clinical deterioration and can be carried out on an outpatient basis. Daily corticosteroids are usually started in patients with generalized MG at a dose >1 mg/kg prednisone. Patients should be continued on this dose until clinical improvement is maintained for several days, then gradually weaned off, and switched to alternate-day therapy. With improvement sustained over several months, an effort should be made to reduce the dose (usually in 5-mg decrements) administered on alternate days. Although a Cochrane review underscored the dearth of controlled trials, the improvement rate is generally estimated to be 60–90%.[39] Complete remission is rare, and most patients will require some dose of steroids indefinitely. The physician should be alert to the possibility that anticholinesterase requirements may decrease as the patient responds to corticosteroids.

Plasmapheresis

Plasmapheresis has enjoyed popularity since its introduction as an auxiliary treatment modality in patients with generalized MG in 1976, particularly as a temporizing measure.[40] It appears to be most beneficial in patients in myasthenic crisis and in those experiencing progressive deterioration despite treatment with anticholinesterases and corticosteroids. Plasmapheresis has also proved to be useful in preparing patients for thymectomy when their course is complicated by involvement of the bulbar and respiratory musculature. Such patients may also require short-term plasmapheresis during the postoperative period. Plasmapheresis also appears to be particularly efficacious in anti-MuSK antibody–associated disease. There is no long-term benefit of plasmapheresis when added to prednisone.

Although there are no hard and fast rules, the average exchange is 1–2 L/day for 7–14 days. Improvement is usually observed within a few days of concluding the treatment course, although patients in crisis often benefit more quickly. The mechanism of action most likely involves the removal of the pathogenic autoantibody, as a reduction in titer of anti-AChR antibody correlates with clinical improvement. However, it is also possible that the removal of other phlogistic humoral factors contributes to clinical efficacy.

Intravenous Immunoglobulin

The efficacy of IVIG in the treatment for MG was suggested by several uncontrolled clinical trials. Interest in the use of this biological grew out of its demonstrated efficacy in other autoimmune diseases, most notably autoimmune thrombocytopenia. Subsequent randomized double-blind placebo-controlled trials provided proof of clinical efficacy. Conventional dosing is infusion of 2 g/kg divided over 5 days, although some practitioners prefer to administer this dose over 2 days. In a randomized, controlled trial a total dose of 1 g/kg was found to be as efficacious as a dose of 2 g/kg. IVIG therapy is generally associated with rapid clinical improvement in responsive patients, independent of whether they had undergone thymectomy or were being treated concurrently with corticosteroids or immunosuppressive agents. In some patients, improvement was sustained over a period of several weeks. Improvement has not always been accompanied by a consistent reduction in anti-AChR antibody titers. In general, IVIG and plasmapheresis appear to be equivalent in efficacy.[41] However, there is a general impression that IVIG is preferable because of better tolerance and less cost. The mechanism(s) of this apparent salutary effect is unknown, although there is evidence IVIG contains antibodies directed against the idiotypes of anti-AChR antibodies.

Immunosuppressive Agents

Immunosuppressive drugs have been tried primarily in patients who have failed treatment with anticholinesterases, thymectomy, plasmapheresis, and corticosteroids.[42] Most of the experience has been obtained with azathioprine, which has strong antiinflammatory effects as well as immunosuppressive activity. The dose of azathioprine has varied between 1 and 3 mg/kg/day, with improvement seen between 5 and 20 weeks. The drug is usually started at a lower dose and escalated weekly to achieve the maintenance dose. The patient should be followed with complete blood counts, particularly during the initiation of therapy, as azathioprine has a suppressive effect on the bone marrow. A white blood cell count below 2500 or a neutrophil count below 1500 should prompt a reduction or termination of the dosage. The results of a randomized double-blind placebo-controlled trial indicated that the addition of azathioprine (2.5 mg/kg) to alternate-day prednisolone was associated with a reduction of the prednisolone dose, fewer treatment failures, longer remissions, and fewer side effects.[43]

There is considerably less experience in the treatment of steroid-unresponsive patients with cyclophosphamide, another powerful immunosuppressive agent. It is associated with more adverse effects and does not appear to offer any significant advantage over azathioprine when used in standard dosing regimens. High-dose intravenous cyclophosphamide therapy for patients with MG refractory to conventional immunosuppressive agents has been investigated as an approach to immunoablate bone marrow and allow for subsequent repopulation by endogenous stem cells. Although durable responses were seen in some patients in an early trial, this regimen is not readily available and should only be utilized in refractory patients under treatment in specialized centers.

Methotrexate has been used in some uncontrolled studies, but there is no information to indicate that it is more efficacious or safer than azathioprine and its onset of action may take as long as several months. As is true for corticosteroid therapy, it is the rare patient who enjoys a permanent remission following institution of immunosuppressive therapy, and those who show some improvement often require treatment indefinitely.

Cyclosporine, a potent immunosuppressive agent, has been investigated because it interferes with IL-2–mediated T-cell proliferation and thus would be expected to interfere with the generation of T cells that would "help" the anti-AChR antibody response. A retrospective study suggested that cyclosporine provided benefit in patients whose disease was refractory to corticosteroids and azathioprine. Serious renal toxicity and treatment withdrawal, which plagued earlier studies, were reduced by careful selection of patients. In a 12-month European trial, cyclosporine appeared to be as efficacious as azathioprine in producing clinical improvement. Tacrolimus has a similar mechanism of action to cyclosporine. When used in low dosage, it has proved as effective as cyclosporine as a corticosteroid-sparing agent with fewer side effects. Nevertheless, most practitioners reserve these agents for use in patients whose disease is refractory to the combination of azathioprine and prednisone.

Mycophenolate mofetil (MMF), another immunosuppressive agent that affects both T and B cells, was widely touted following the completion of early trials as a steroid-sparing agent in patients with MG. That experience suggested that it was effective 70–75% of the time, although probably less so in refractory MG. It has an acceptable safety profile with adverse effects largely related to gastrointestinal intolerance. Benefit may require many weeks of administration. However, two randomized, placebo-controlled trials have challenged the early optimism. In one, the addition of MMF treatment at the initiation of a 36-week schedule of prednisone tapering was not found to be superior to placebo in maintaining myasthenia control.[44] In the second, the coadministration of MMF and prednisone provided no better control of myasthenic weakness than prednisone alone in the initial management of generalized MG.[45] However, none of the patients included in these two studies was known to be steroid resistant. Thus whether MMF has long-term benefits with respect to myasthenic weakness or steroid-sparing effects in the population of steroid-resistant patients remains an open question.

POSSIBLE FUTURE THERAPEUTIC OPTIONS

Many possible experimental avenues of investigation that have been opened by studies in EAMG. These are aimed at interrupting the sensitization process of helper CD4 T cells, interrupting their effector function, or interdicting the action of downstream proinflammatory molecules. Studies directed at CD4 T cells include impeding their activation by inhibitors of the costimulatory molecules CD28 and ICOS and the induction of anergy or apoptosis of AChR-reactive T cells.[34] In addition, inhibition of factors that contribute to AChR antibody production, and against which monoclonal antibodies (mAbs) are now available for human use, make this approach possible and, in some cases, are now under clinical trials. This includes agents that abrogate the action of IL-6, IL-17, and BAFF. Additional studies directed at AChR-specific B cells include the use of AChR/Fcγ fusion proteins to induce apoptosis of AChR-specific B cells by cross-linking their B-cell AChR receptors and inhibitory FcγRIIb receptors. In addition, non–cross-linking IgG4 anti-AChR antibodies have been used to inhibit pathogenic IgG autoantibody-induced antigenic modulation. Bortezomib, a proteosome inhibitor approved for usage in multiple myeloma and mantle cell lymphoma, has been found to reduce anti-AChR antibody titers, inhibit damage to the postsynaptic muscle membrane, and lead to clinical improvement. Moreover, this agent was shown to suppress anti-AChR antibody production and deplete plasma cells from cultures of cell suspensions obtained from thymus specimens of patients with MG.[34]

The induction of antigen-specific immune tolerance, the therapeutic holy grail in autoimmune diseases, has received considerable attention in EAMG. Early studies utilized extracellular domain sequences that form epitopes for pathological autoantibodies to induce tolerance to AChR. However, this approach risks provoking autoimmunity rather than suppressing it. To avoid this risk, one group has recently developed a novel vaccine consisting of bacterially expressed human AChR cytoplasmic domains.[46] To date, this group has used this vaccine to prevent development of chronic EAMG when administered immediately after the acute phase of EAMG and to rapidly reverse established chronic EAMG when started during the chronic phase of EAMG. Treatment effects are robust and long-lasting. Although the mechanisms of action of this novel approach have yet to be fully elucidated, they may involve a combination of antibody-mediated feedback suppression and regulatory T cell–mediated active suppression.

Complement inhibition has shown efficacy in the treatment of EAMG and represents a target in patients with MG now that complement inhibitors have demonstrated benefit in the treatment of a number of human disorders. A phase II placebo-controlled cross-over study with a fully humanized monoclonal anti-C5 antibody was initiated only to be terminated because of inadequate recruitment.[34] Nevertheless, given the importance of the terminal complement pathway in the pathogenesis of cases of MG and the continued development of a number of complement inhibitors, it is likely that this strategy will be revisited.

A B-cell targeting agent that has gotten much attention in patients with MG is rituximab, a mAb specific for CD20, a protein expressed on B lymphocytes. Anecdotal reports and small uncontrolled case series using rituximab have shown improvement of MG. Interestingly, responsive patients do not demonstrate a reduction in anti-AChR antibodies, suggesting an effect on nonantibody functions, such as antigen presentation and proinflammatory cytokine secretion. Although most treated patients have been anti-AChR antibody positive, benefit has also been observed in anti-MuSK-positive/anti-AChR antibody–negative patients, in whom longstanding remission has been seen in association with reduction of anti-MuSK antibody titers. A controlled clinical trial of rituximab in anti-AChR–associated MG is currently underway. A number of other mAbs directed at CD20, including fully humanized ones, are under study in other autoimmune diseases and will likely attract attention with regard to MG.

CONCLUSIONS

MG is a prototypic autoimmune disease in which the autoantigen/s are well defined as are the mechanisms that mediate dysfunction of the target tissue, skeletal muscle. EAMG, in which many of the features of MG are faithfully recapitulated, has provided a useful vehicle for elucidating the immunoregulatory abnormalities that underlie the pathogenesis of the human disease, and the

ON THE HORIZON

Development of Immunomodulatory Therapies

- Complement inhibitors
- Newly developed B-cell lineage inhibitors, including anti–plasma cell monoclonal antibodies
- Acetylcholine receptor (AChR)–Fc fusion molecules
- Inhibitors that target key immunopathogenic cytokines
- AChR-specific immunosuppression

development of novel therapies for use in the humans. MG occupies a unique position in the pantheon of autoimmune diseases because of its strong association with thymic pathology. In MG-associated thymic hyperplasia, thymic events, particularly those leading to antecedent inflammation within the medulla, likely serve as a driving force behind the development of the autoimmune process. Indeed, the overarching idea that tolerance is breached in the thymus may be viewed as antithetical, given the established and critical role this organ plays in T-cell central tolerance. Unravelling the link between the thymus and MG will require further study. However, in the meantime, we can expect that continued exploration of immune perturbations and their underlying molecular mechanisms in animals with EAMG and in patients with MG will lead to safer and more robust therapeutic interventions.

ACKNOWLEDGMENT

This work was supported by National Institutes of Health (NIH) grant NS19546.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Lisak RP, Barchi RL. Myasthenia gravis. In: Walton JN, editor. Major problems in neurology, vol. 11. Philadelphia, PA: WB Saunders; 1982. p. 5.
2. Engel A. The investigation of congenital myasthenic syndromes. Ann NY Acad Sci 1993;681:425.
3. Levinson AI, Zweiman B, Lisak RP. Immunopathogenesis and treatment of myasthenia gravis. J Clin Immunol 1987;7:187.
4. Tindal RSA. Humoral immunity in myasthenia gravis: biochemical characterization of acquired anti-receptor antibodies and clinical correlations. Ann Neurol 1981;10:437.
5. Papadouli I, Sakarellos C, Tzartos SJ. High-resolution epitope mapping and fine antigenic characterization of the main immunogenic region of the acetylcholine receptor. Eur J Biochem 1993;211:227.
6. Leite MI, Jacob S, Viegas S, et al. IgG1 antibodies to acetylcholine receptors in "seronegative" myasthenia gravis. Brain 2008;131:1940.
7. Levinson AI, Lisak RP. Myasthenia gravis. In: Detrick B, Schmitz J, Hamilton RG, editors. Manual of molecular and clinical laboratory immunology. 8th ed, ASM Press, 2016.
8. Sanders DB, Juel VC. MuSK-antibody positive myasthenia gravis: questions from the clinic. J Neuroimmunol 2008;201-202:85.
9. Arrli JA, Skeie GO, Mygand A, et al. Muscle striation antibodies in myasthenia gravis. Ann NY Acad Sci 1998;841:505.
10. Drachman DB, Adams RN, Josifek LF, et al. Antibody-mediated mechanisms of ACh receptor loss in myasthenia gravis: clinical relevance. Ann NY Acad Sci 1981;377:175.
11. Engel AG, Sahashi K, Fumagalli G. The immunopathology of acquired myasthenia gravis. Ann NY Acad Sci 1981;377:158.
12. Drachman DB, Angus CW, Adams RN, et al. Myasthenic antibodies cross-link acetylcholine receptors to accelerate degradation. N Engl J Med 1978;298:1116.
13. Richman DP, Wollmann RL, Maselli RA, et al. Effector mechanisms of myasthenic antibodies. Ann NY Acad Sci 1993;681:264.
14. Conti-Fine B, Navaneetham D, Karachunski PI, et al. T cell recognition of the acetylcholine receptor in myasthenia gravis. Ann NY Acad Sci 1998;841:283.
15. Patrick J, Lindstrom J. Autoimmune response to acetylcholine receptor. Science 1973;180:871.
16. Christadoss P, Poussin M, Deng C. Animal models of myasthenia gravis. Clin Immunol 2000;94:75.
17. Bellone M, Ostile N, Lei S, et al. Experimental myasthenia gravis in congenic mice. Sequence mapping and H-2 restriction of T helper epitopes on the α subunits of *Torpedo californica* and murine acetylcholine receptors. Eur J Immunol 1991;21:2303.
18. Karachunski PI, Ostlie NS, Okita DK, et al. Interleukin-4 deficiency facilitates development of experimental myasthenia gravis and precludes its prevention by nasal administration of CD4+ epitope sequences of the acetylcholine receptor. J Neuroimmunol 1999;95:73.
19. Wang W, Milani F, Ostlie G, et al. C57BL/6 mice genetically deficient in IL-12/IL-23 and IFN-γ are susceptible to experimental autoimmune myasthenia gravis, suggesting a pathogenic role of non-Th1 cells. J Immunol 2007;178:7072.
20. Schaffert H, Pelz A, Saxena A, et al. IL-17-producing CD4(+) T cells contribute to the loss of B-cell tolerance in experimental autoimmune myasthenia gravis. Eur J Immunol 2015;45:1339.
21. Levinson AI, Wheatley LM. The thymus and the pathogenesis of myasthenia gravis. Clin Immunol Immunopathol 1995;78:1.
22. Berrih-Aknin S, Lepanse R. Myasthenia gravis: a comprehensive review of immune dysregulation and etiological mechanisms. J Autoimmun 2014;52:90.
23. Marx A, Pfister F, Schalke B, et al. The different roles of the thymus in the pathogenesis of the various myasthenia gravis subtypes. Autoimmunity Rev 2013;12:875.
24. Levinson AI, Zheng Y, Gaulton G, et al. Intrathymic expression of neuromuscular acetylcholine receptors and the immunopathogenesis of myasthenia gravis. Immunol Res 2003;27:399.
25. Wolfe GI, Kaminski HJ, Aban IB, et al. Randomized trial of thymectomy in myasthenia gravis. N Engl J Med 2016;375:511–22.
26. Levinson AI. Modeling the intrathymic pathogenesis of myasthenia gravis. J Neurolog Sci 2013;333:60.
27. Chuang WY, Strobel P, Belharazem D, et al. The PTPN22gain-of-function+1858T(+) genotypes correlate with low IL-2 expression in thymomas and predispose to myasthenia gravis. Genes Immun 2009;10:667.
28. Greve B, Hoffmann P, Illes Z, et al. The autoimmunity-related polymorphism PTPN22 1858C/T is associated with anti-titin antibody-positive myasthenia gravis. Hum Immunol 2009;70:540.
29. Stefansson K, Dieperink ME, Richman DP, et al. Sharing of epitopes by bacteria and the nicotinic acetylcholine receptor: a possible role in the pathogenesis of myasthenia gravis. Ann NY Acad Sci 1987;505:451.
30. Schwimmbeck PL, Dyrberg T, Drachman DB, et al. Molecular mimicry and myasthenia gravis: an autoantigenic site of the acetylcholine receptor α-subunit that has biologic activity and reacts immunochemically with herpes simplex virus. J Clin Invest 1989;84:1174.
31. Dwyer DS, Vakil M, Bradleg RT, et al. A possible cause of myasthenia gravis: idiotypic networks involving bacterial antigens. Ann NY Acad Sci 1987;505:461.
32. Bever CT Jr, Chang HW, Penn AS, et al. Chemical alteration of acetylcholine receptor by penicillamine: a mechanism for induction of myasthenia gravis. Neurology 1982;32:1077.
33. Penn A, Jacques JJ. Cells from mice exposed chronically to D-penicillamine show proliferative responses to D-penicillamine-treated self (macrophage/dendritic cells): a graft-versus-host response? Ann NY Acad Sci 1993;681:319.

34. Guptill JT, Soni M, Meriggioli MN. Current treatment, emergent translational therapies, and new therapeutic targets for autoimmune myasthenia gravis. Neurother 2016;13:118.
35. Kumar V, Kaminski HJ. Treatment of myasthenia gravis. Curr Neurol Neurosci Rep 2011;11:89.
36. Olanow CW, Wechsler AS, Sirotkin-Roses M, et al. Thymectomy as primary therapy in myasthenia gravis. Ann NY Acad Sci 1987;505: 595.
37. Sonnett JR, Jaretzki A III. Thymectomy for Nonthymomatous Myasthenia gravis: a critical analysis. Ann NY Acad Sci 2008;1132:315.
38. Johns TR. Long-term corticosteroid treatment of myasthenia gravis. Ann NY Acad Sci 1987;505:568.
39. Schneider-Gold C, Gadjos P, Toyka K, et al. Corticosteroids for myasthenia gravis. Cochrane Database Syst Rev 2005;(18):CD002828.
40. Seybold M. Plasmapheresis in myasthenia gravis. Ann NY Acad Sci 1987;505:58.
41. Miller RG, Barohn RJ, Dubinsky R. Expanding the evidence base for therapeutics in myasthenia gravis. Ann Neurol 2010;68:776.
42. Sanders DB, Evoli A. Immunosuppressive therapies in myasthenia gravis. Autoimmunity 2010;43:428.
43. Palace J, Newsom-Davis J, Lecky B. A randomized double-blind trial of prednisolone alone or with azathioprine in myasthenia gravis. Myasthenia Gravis Study Group. Neurology 1998;50:1778.
44. Sanders DB, Hart IK, Mantegazza R, et al. An international, phase III, randomized trial of mycophenylate mofetil in myasthenia gravis. Neurology 2008;71:400.
45. The Muscle Study Group. A trial of mycophenylate mofetil with prednisone as initial immunotherapy in myasthenia gravis. Neurology 2008;71:394.
46. Luo J, Lindstrom J. AChR-specific immunosuppressive therapy of myasthenia gravis. Biochem Pharmacol 2015;97:609.

MULTIPLE-CHOICE QUESTIONS

1. The most frequent pathogenic autoantibodies found in patients with myasthenia gravis (MG) are specific for a conformational epitope on which subunit of the neuromuscular acetylcholine receptor?
   A. α
   B. β
   C. δ
   D. ε

2. Patients with MG whose serum is positive for anti- *mu*scle-*s*pecific tyrosine *k*inase (MuSK) antibodies are most resistant to treatment with which one of the following modalities?
   A. Corticosteroid
   B. Intravenous immunoglobulin
   C. Plasmapheresis
   D. Pyridostigmine bromide

3. A 35-year-old man presented with diplopia, difficulty swallowing, and weakness in his upper and lower extremities. On laboratory testing of his serum, he was found to have elevated titers of anti–acetylcholine receptor antibodies and antistriational antibodies. He underwent thymectomy, which very likely showed which of the following pathologies?
   A. Germinal center hyperplasia
   B. Thymic atrophy
   C. Cortical thymoma
   D. Medullary lymphoma

4. An inflammatory process in which of the following anatomical locales is considered to initiate the development of MG?
   A. Skeletal muscle
   B. Thymic cortex
   C. Cervical lymph node
   D. Thymic medulla

Multiple Sclerosis

Benjamin M. Segal

Multiple sclerosis (MS), a chronic inflammatory demyelinating disorder of the central nervous system (CNS), is the most frequent cause of nontraumatic neurological disability among young adults in the Western Hemisphere. Although MS is widely considered a disease of North America and Europe, there is increasing evidence that it is more common in other regions of the world, including Asia and the Middle East, than previously appreciated. The median age at presentation is 28–31 years, which is, in part, responsible for the disproportionately high social and economic tolls of the disease. Furthermore, the incidence of MS is increasing for unknown reasons. Fortunately there have been dramatic advances in the treatment of relapsing forms of MS over the past 20 years, spurred by the introduction of 14 disease-modifying agents (DMAs); and more are actively under development. These drugs significantly decrease the risk of relapse and lesion formation. Consequently, the implications of being diagnosed with relapsing-remitting MS have changed considerably in the span of a generation. Despite this success, significant challenges persist. There is a dire need for treatments that slow, or even halt, disability accumulation in patients with progressive forms of MS, and for interventions that restore lost neurological functions across MS subsets.

CLINICAL SUBSETS AND PHENOMENOLOGY

Relapsing-Remitting MS (RRMS)

In the majority of cases (85–90%), MS presents with a relapsing-remitting course, characterized by discrete episodes of neurological dysfunction (relapses or exacerbations) separated by clinically quiescent periods (remissions). The frequency of relapses can vary widely among patients as well as during different periods in an individual patient's disease course. At present no clinical features or biomarkers that are predictive of relapse rate have been identified. The signs and symptoms that occur during relapses are also diverse and unpredictable, since lesions can form at literally any site in the CNS, spanning the cerebrum, brainstem, cerebellum, optic nerves, and spinal cord. By definition, the peripheral nervous system is spared.

MS lesions are readily visualized in CNS white matter via magnetic resonance imaging (MRI) (Fig. 66.1). Symptomatic lesions generally occur in locations where nerve fibers converge to subserve a common function. Hence, typical presentations of RRMS include optic neuritis with monocular visual deficits (secondary to lesions in the optic nerve), myelitis with weakness and numbness in the extremities, sometimes accompanied by incontinence (caused by spinal cord lesions), and brainstem syndromes manifesting as slurred speech, swallowing difficulties, imbalance, tremor, and/or double vision. Serial MRI studies have demonstrated that the majority of MS lesions are actually clinically silent. This is not surprising considering the abundance of redundant nerve fiber tracts in the CNS and the commitment of large areas of cerebral white matter to subtle personality traits and cognitive skills. Consequently, CNS tissue damage may be inflicted surreptitiously during clinical remissions, making MRI a more sensitive indicator of disease activity compared with the history or the results of neurological examination (see Fig. 66.1).

Patients with MS often recover function following a clinical relapse, either partially or fully, particularly during the early clinical course. However, old symptoms can temporarily reemerge when the core body temperature is elevated as a result of infection or strenuous exercise. This unmasking of latent deficits, referred to as the *Uhthoff phenomenon,* is a consequence of the physiological slowing of axon signal propagation that normally occurs at high core body temperatures. In healthy individuals, the degree of slowing has no clinical consequence, but in MS patients, it may precipitate the decompensation of white matter tracts already compromised by demyelination and axonal drop-out.

Secondary Progressive MS (SPMS)

During the course of RRMS, relapses decrease in frequency over time and sometimes disappear completely. However, in the vast majority of cases, they are replaced by an insidious, gradual accumulation of disability, referred to as the *secondary progressive (SP) stage.* The symptoms and signs that characterize neurological decline during SPMS are diverse. Progressive myelopathy, hemiparesis, and/or gait imbalance are common. Subcortical dementia is increasingly recognized as a feature of the disease. Longitudinal natural history studies conducted before DMAs were widely available found that the majority of patients with RRMS transitioned to the SP stage within 10–20 years of the initial presentation of disease. An epidemiological study of MS patients in British Columbia, published in 2010, found that the median time to SPMS onset was 21.4 years.[1] A number of factors, including male gender, the presence of motor symptoms at clinical onset, and a history of poor recovery from relapses, are associated with both a shorter time to, and younger age at, evolution to SPMS. It has not been definitively determined whether optimal management of RRMS with the use of DMAs can delay, or even prevent, the onset of SPMS. Previous longitudinal observational and retrospective cohort studies that investigated whether treatment with first-generation DMAs alters the time to reach SPMS yielded conflicting results. However, a recent prospective study of 517 actively treated patients found that rates of worsening and evolution to SPMS were substantially

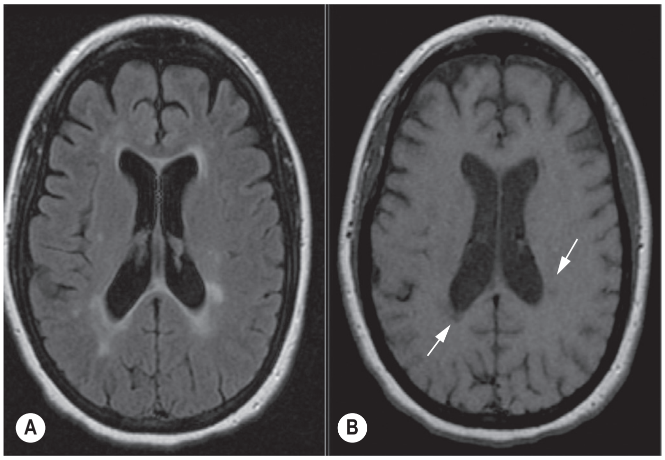


FIG 66.1 (A) T2-weighted fluid-attenuated inversion recovery (FLAIR) magnetic resonance imaging (MRI) scan of the brain of a patient with multiple sclerosis (MS) showing hyperintense lesions located in the periventricular and subcortical white matter. (B) T1-weighted MRI scan showing T1-black holes (arrows), indicative of profound axonal loss, and generalized atrophy with enlarged ventricles.

lower when compared with the findings of earlier natural history studies of untreated patients.[2]

The cellular and molecular mechanisms underlying the conversion from the RR stage to the SP stage are poorly understood. Some investigators have questioned the relevance of neuroinflammation during the SP stage and have posited neurodegeneration as being primarily responsible for the clinical deterioration that ensues (in the form of neuronal death, mitochondrial dysfunction in axons, Wallerian degeneration, and gliosis).[3] Conversely, there is evidence of persistent immune dysregulation in SPMS, although it may differ from RRMS with regard to the cytokine networks and leukocyte subsets involved, as well as the distribution of infiltrating cells within the CNS.[4-6] Nonetheless, DMAs that are therapeutically beneficial in RRMS have generally not been found to be effective in slowing disability accumulation in SPMS.[7] At present, the management of SPMS involves alleviation of symptoms, optimization of residual functions, and prevention of complications.

Primary Progressive MS (PPMS)

PPMS is distinguished from SPMS by the absence of an antecedent RR phase. Otherwise, the clinical features of SPMS and PPMS can be indistinguishable. The most common clinical phenotype is spastic paraparesis, followed by cerebellar dysfunction and hemiplegia. However, there are striking demographic differences between PPMS and RRMS/SPMS. PPMS tends to present at an older age (peaking in the fifth and sixth decades) compared with RRMS (which peaks in the third and fourth decades). Furthermore, RRMS occurs 2–3 times more frequently in females than in males, whereas in PPMS the gender ratio is closer to 50 : 50. Some investigators have noted that the neuropathological and radiological features of PPMS overlap extensively with those of SPMS, leading them to conclude that PPMS and SPMS belong to the same disease spectrum. In support of that viewpoint, familial clusters of MS that include some members with PPMS and others with RR/SPMS have been described. It is possible that, analogous to SPMS, acute inflammatory lesions form during early stages of PPMS, prior to overt clinical progression, but happen to arise exclusively in clinically silent areas. Conversely, some investigators have argued that PPMS is a distinct disease entity driven primarily by neurodegenerative processes from its inception.[3] This viewpoint is supported by the general failure of immunomodulatory agents to attenuate the course of PPMS.[8] Conversely, in a randomized double-blind placebo-controlled trial, treatment with a B cell–depleting monoclonal antibody (mAb) delayed disability progression in a subset of patients with PPMS who were younger and/or had inflammatory lesions.[9] Collectively, these findings suggest that the pathogenesis of PPMS is multifaceted and heterogeneous and that the relative contribution of inflammation and neurodegeneration to clinical outcomes varies among patients.

CLINICAL PEARLS

Clinical Features of MS

In the majority of cases (80–85%), multiple sclerosis (MS) presents with a relapsing-remitting course.

- The symptoms and signs experienced by patients with MS are diverse because lesions can form at any site in the central nervous system (CNS), including the optic nerves, cerebrum, brainstem, and spinal cord.
- The rate, severity, and symptoms of relapses are highly variable and unpredictable.
- Most MS lesions form silently; magnetic resonance imaging (MRI) is a more sensitive tool to gauge disease activity compared with history or neurological examination.
- Acute relapses tend to decrease in frequency over time and are typically replaced by a gradual accumulation of disability. This later phase of disease is referred to as *secondary progressive MS* (SPMS).

DIAGNOSIS

The most widely used guideline for diagnosing MS is the *McDonald criteria*, originally proposed by the International Panel on Diagnosis of Multiple Sclerosis in 2001 and subsequently revised in 2005 and 2010[10] (Table 66.1). RRMS is, by definition, a dynamic multifocal inflammatory demyelinating disease of the CNS. Therefore the demonstration of lesion dissemination in time and space is crucial for its diagnosis. The criteria for dissemination in time can be satisfied by ≥2 distinct clinical exacerbations, by one exacerbation followed by the interval appearance of a new lesion on serial MRI scans, or by the simultaneous presence of asymptomatic gadolinium-enhancing (*i.e.*, acute inflammatory) and nonenhancing MRI lesions at any time. Dissemination in space can be demonstrated by objective clinical evidence of involvement of ≥2 sites in the CNS (based on the neurological examination and/or delayed latencies on evoked potential testing) or by the presence of T2-weighted MRI lesions in at least two of the following areas: periventricular, juxtacortical, and infratentorial, cerebral white matter or spinal cord. In the latest iteration of the McDonald criteria, an MS exacerbation is defined as "patient-reported symptoms or objectively observed signs typical of an acute inflammatory demyelinating event in the CNS, current or historical, with duration of at least 24 hours, in the absence of fever or infection." There are no clinical features or biomarkers that are pathognomonic for MS. Therefore it is essential to rule out competing diagnoses. The presence of unique oligoclonal bands and/or elevated immunoglobulin G (IgG) index in the cerebrospinal fluid (CSF), indicative of

TABLE 66.1 McDonald 2010 Criteria for the Diagnosis of Multiple Sclerosis

Clinical Presentation	Additional Data Needed for an MS Diagnosis
≥ 2 attacks; objective clinical evidence of ≥ 2 lesions or objective clinical evidence of 1 lesion with reasonable historical evidence of a prior attack	None
≥ 2 attacks; objective clinical evidence of 1 lesion	Dissemination in space demonstrated by ≥ 1 T2 lesion in at least 2 or 4 MS-typical regions (periventricular, juxtacortical, infratentorial, or spinal cord)
1 attack; objective clinical evidence of ≥ 2 lesions	Simultaneous presence of asymptomatic gadolinium-enhancing and nonenhancing lesions at any time; or a new T2 and/or gadolinium-enhancing lesion(s) on follow-up MRI, irrespective of its timing with reference to a baseline scan; or Await a second clinical attack
1 attack; objective clinical evidence of 1 lesion (clinically isolated syndrome)	Dissemination in space and time, demonstrated by: *For DIS:* ≥ 1 T2 lesion in at least 2 of 4 MS-typical regions of the CNS (periventricular, juxtacortical, infratentorial, or spinal cord); or Await a second clinical attack implicating a different CNS site; and For DIT: Simultaneous presence of asymptomatic gadolinium-enhancing and nonenhancing lesions at any time; or A new T2 and/or gadolinium-enhancing lesion(s) on follow-up MRI, irrespective of its timing with reference to a baseline scan; or Await a second clinical attack
Insidious neurological progression suggestive of MS (PPMS)	1 year of disease progression (retrospectively or prospectively determined) plus 2 of 3 of the following criteria: 1. Evidence for DIS in the brain based on ≥ 1 T2 lesions in the MS-characteristic (periventricular, juxtacortical, or infratentorial) regions 2. Evidence for DIS in the spinal cord based on ≥ 2 T2 lesions in the cord 3. Positive CSF (isoelectric focusing evidence of oligoclonal bands and/or elevated IgGindex)

From Polman CH, Reingold SC, Banwell B, et al. Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. Ann Neurol 2011; 69(2): 292–302.

primary antibody production in the CNS, supports a diagnosis of MS, but those findings are observed in a wide range of neuroinflammatory conditions, including subacute sclerosing panencephalitis, neurosarcoidosis, Lyme disease, and systemic lupus erythematosus (SLE) with CNS involvement. A diagnosis of PPMS requires 1 year of disease progression plus two of the three following criteria: *(i)* at least one T2 lesion in the periventricular, juxtacortical, or infratentorial cerebral white matter; *(ii)* at least two T2 spinal cord lesions; and *(iii)* positive CSF findings, defined as the presence of oligoclonal bands or elevated IgG index.

RISK FACTORS

Genetic Risk Factors

A monozygotic twin whose cotwin has MS has a 20–30% risk of developing the disease, whereas the corresponding risk for a dizygotic twin is 2–5%.[11] In contrast, the incidence of MS in the general population is approximately 0.1%. Hence, the risk of MS drops with increasing genetic distance, suggesting that genes play a significant role in determining MS susceptibility. The human leukocyte antigen (HLA) region represents the strongest MS susceptibility locus, genome wide, and has been implicated in all ethnic populations thus far studied. It accounts for up to 10.5% of genetic variance underlying risk.[12] The primary signal maps to the *HLA-DRB1* gene in the class II segment of the locus, implicating a role of CD4 T-cell responses in MS pathogenesis. Genome-wide association studies (GWAS) have revealed over 100 non-HLA susceptibility loci, each of which contributes a small amount to MS risk. Strikingly, most of these map to regions containing genes implicated in immunological, rather than neuronal or glial, pathways. Genes involved in T-helper (Th)–cell differentiation are overrepresented, including the interleukin-2 (IL-2) receptor α chain and the IL-7 α chain, which modulate T-cell proliferation and survival. Over one-third of the MS susceptibility loci overlap with regions previously identified in GWAS of other autoimmune diseases, including celiac disease, type 1 diabetes, rheumatoid arthritis, and/or inflammatory bowel disease. Together these data provide support for a primary immunological, as opposed to neurodegenerative, etiology of MS. An unresolved question is whether genetic variants affect the clinical course of RRMS, including such features as relapse rate or severity and time to conversion to the SP stage. Genetic variants that are predictive of responsiveness to DMAs have yet to be identified. A growing area of interest that warrants further investigation is the potential impact of epigenetic modifications in different cell types on MS susceptibility, clinical course, and/or therapeutic responsiveness.

Environmental Risk Factors

As mentioned above, twin concordance rates in MS have been used to highlight the importance of heredity in MS susceptibility. Paradoxically, the same data can be used to argue for the importance of environmental influences. Hence, over 70% of monozygotic twins of individuals with MS do not develop the disease. Despite arduous attempts, no convincing evidence has been produced for genetic, epigenetic, or transcriptome differences that explain MS discordance among monozygotic twin pairs.[13] The risk of MS appears to be predicated on a complex interplay between genetic and environmental factors.

Geographic Prevalence Patterns

One of the most convincing illustrations of the impact of the environment on the development of MS is its geographical distribution. The prevalence of MS is highest in the Scandinavian countries, Canada, and Scotland and lowest in the equatorial regions. Kurtzke et al. were the first to notice a latitudinal gradient in the prevalence of MS across the United States, with the disease being most common in the Northern states and gradually declining toward the South.[14] Similar latitudinal gradients have been observed in Europe, Australasia, and Japan. Population migration studies indicate that the geographical risk of MS is established prior to adolescence. Prepubescent children who migrate assume the risk of their adopted country, whereas adults carry forward the risk of the location where they spent their childhood. An environmental agent that is encountered in childhood and acts as predisposing factor for the development of MS later in life must be responsible, but this agent has yet to be definitely identified.

Vitamin D

The discovery of the protective role of vitamin D in MS may explain, in part, its geographical distribution. Ultraviolet light catalyzes the conversion of vitamin D to its bioactive form, and the prevalence of MS is highest in regions with relatively low annual sunlight exposure. In a large prospective study, the risk of MS decreased with increasing serum levels of 25-hydroxyvitamin D.[15] Therefore interventions that raise the level of 25-hydroxyvitamin D levels might have a prophylactic protective effect on healthy individuals who are predisposed to develop MS, such as the first-degree relatives of patients with MS. Indeed, numerous independent prospective studies have found an inverse relationship between dietary or supplemental vitamin D intake in adults and future risk of MS. Furthermore, there is accumulating evidence that low serum 25-hydroxyvitamin D levels and low dietary vitamin D intake during pregnancy increase the risk of MS in the offspring.[16] Whether vitamin D levels influence the clinical course of individuals with established MS is more controversial. A randomized controlled trial of vitamin D supplementation in patients with MS is currently underway in 16 academic centers (ClinicalTrials.gov identifier: NCT01490502). The results of that trial should bring clarity to the issue.

Infection

In addition to low levels of vitamin D, another prognostic factor for development of MS is primary infection with Epstein-Barr virus (EBV) in adulthood, as indicated by the emergence of EBV-specific IgM antibodies in serum. Analysis of serial serum samples stored in the US Department of Defense Serum Repository revealed that MS risk was extremely low among individuals not infected with EBV but increased sharply in the same individuals following seroconversion to positivity for antibodies against EBV.[17] The odds of MS were recently estimated to be >10 times higher among EBV-positive persons than among EBV-negative persons.[18] One way in which EBV infection could promote MS pathogenesis is via "molecular mimicry," which occurs when a microbial epitope shares sequence similarities with a self peptide, in this case a myelin peptide (Chapter 50). T-cell receptors (TCRs) that cross-react with structurally homologous EBV and myelin antigens have been discovered in the peripheral CD4 TCR repertoire of individuals with MS.[19] Peripheral CD4 T cells that express such cross-reactive TCRs could be activated during EBV infection, enabling them to cross the blood–brain barrier (BBB), encounter their cognate myelin antigen within CNS white matter, and initiate MS lesion formation. An alternative theory involves infiltration of the CNS by EBV-infected B cells. B cells expressing EBV small RNA and proteins have been detected in meningeal follicle–like structures and inflamed cortical lesions in SPMS brain tissue.[20] EBV-driven expansion and activation of meningeal B cells could potentially contribute to the formation of the follicle-like structures, which have been associated with large subpial cortical lesions and a more aggressive clinical course.[6]

Although infectious agents, such as EBV, have been implicated in MS pathogenesis, other pathogens may have a therapeutic effect. It has been suggested that in addition to high levels of sunlight exposure, endemic helminth infection is one of the factors responsible for the low prevalence of MS in tropical regions. Infection with helminths, including *Heligmosomoides polygyrus, Fasciola hepatica, Schistosoma mansoni,* and *Trichinella,* has been shown to be protective in animal models of MS.[21] Several small prospective studies showed that patients with MS naturally infected with different species of parasitic worms had a milder disease course and lower MRI inflammatory activity compared with uninfected patients.[22] Antiparasite treatment was associated with exacerbation of MS. Mechanistic substudies have suggested that parasites modulate MS disease activity by boosting the frequency of IL-10 and transforming growth factor-beta (TGF-β)–producing regulatory T and/or B cells. A phase I safety study of orally administered *Trichuris suis* ova (TSO) has been conducted in RRMS, and a number of clinical trials of TSO or dermally administrated hookworm *Necator americanus* as disease-modifying therapy in MS are planned or underway.

Obesity

Observational studies have found that individuals who are obese in early adulthood, as measured by elevated body mass index (BMI), have an approximately twofold increased risk of MS.[23] Obesity in girls is associated with an increased risk of pediatric MS or clinically isolated syndrome.[24] Furthermore, higher weight in adolescence and young adulthood is associated with an earlier age at onset of MS. A mendelian randomization analysis of large GWAS for MS and BMI, respectively, found that one standard deviation (SD) increase in genetically determined BMI conferred a 41% increase in the odds of MS.[25] A number of theories have been proposed for a mechanistic link between obesity and MS risk. First, obesity is known to cause a systemic proinflammatory state, possibly mediated by an adipose-derived hormone that could create a milieu conducive to the differentiation and/or activation of autoimmune effector cells. Alternatively, there is some evidence that genetically elevated BMI decreases 25-hydroxyvitamin D levels.

Modifiable Habits

A substantial body of literature indicates that cigarette smoking increases the risk of MS. A recent meta-analysis revealed a dose–response relationship between cigarette pack-years and MS risk.[26] On the basis of data collected from the Swedish National MS Registry, it was estimated that each additional year of smoking after diagnosis accelerates the time to conversion to SPMS by 4.7%.[27] There is growing evidence that exposure to passive smoking is also a risk factor for MS. Acrolein, a component of cigarette smoke, was found to exacerbate the clinical course of an animal model of MS, in association with enhanced microglial activation.[28]

Another behavior that has come under scrutiny in relationship to MS is dietary sodium intake. In one study, exacerbation rates were found to be 2.75–3.95-fold higher in patients with RRMS with medium or high sodium intakes, respectively, compared with the low-intake group.[29] In contrast, high salt intake was not associated with decreased time to relapse in pediatric-onset MS.[30] High-salt conditions promote the induction of highly pathogenic myelin-reactive T cells *in vitro* and in animal models *in vivo.*[31]

Sex Hormones

MS relapse rates decline during pregnancy, particularly in the third trimester, and rebound in the first 3 months post partum before returning to the prepregnancy rate. This has led to the hypothesis that certain female sex hormones may play a protective role in RRMS. Estriol is an estrogen unique to pregnancy. It is synthesized by the fetoplacental unit and reaches its highest levels in the last trimester. A randomized, double-blinded, placebo-controlled phase II trial of estriol in combination with glatiramer acetate (GA) versus placebo plus GA showed a reduction in the annualized relapse rate at 2 years in the estriol-treated group.[32] Animal model studies have indicated that estrogens, including estriol, have antiinflammatory and neuroprotective effects through engagement of the estrogen receptors expressed on leukocytes and CNS resident cells, respectively.

Testosterone has neuroprotective effects in animal models of MS, and decreased testosterone levels in males with MS were reported to be associated with disability.[33] In a small, open-label, phase II clinical trial, testosterone treatment appeared to arrest loss of gray matter (and even to reverse atrophy of gray matter in the right frontal cortex), as quantified by using voxel-based morphology, in 10 male patients with MS.[34]

KEY CONCEPTS

Risk Factors

The risk of multiple sclerosis (MS) is determined by a combination of genetic and environmental factors.

- The majority of MS susceptibility loci map to regions containing genes implicated in immunological pathways, including human leukocyte antigen (HLA) class II molecules, the interleukin-2 (IL-2) receptor, and the IL-17 receptor.
- Relapse rates decline during the third trimester of pregnancy, in association with high serum levels of estriol.
- Environmental risk factors include low vitamin D levels, exposure to the Epstein-Barr virus (EBV) in adulthood, cigarette smoking, and childhood obesity.

PATHOLOGICAL FEATURES OF MS

White Matter Lesions

The hallmark of MS pathology is the focal demyelinated lesion, or "plaque," present in the white matter of the optic nerves, brain, and spinal cord. Acute lesions are invariably associated with focal breakdown of the BBB and perivascular inflammatory infiltrates. MS infiltrates are dominated by T cells (with a relatively high CD8/CD4 ratio) and myeloid cells (blood-derived monocytes/macrophages and activated microglia). Macrophages/monocytes and activated microglia are spatially associated with disintegrating myelin sheaths, and they actively take up myelin debris. Apoptosis and loss of oligodendrocytes vary widely among lesions. Frequent sites of lesion formation include subcortical and periventricular cerebral white matter, middle cerebellar peduncles, and the posterior columns of the cervicothoracic spinal cord. In the brain, infiltrates frequently follow the course of pericallosal venules, resulting in "Dawson fingers," which are oblong lesions oriented perpendicular to the long axes of the lateral ventricles (Fig. 66.2).

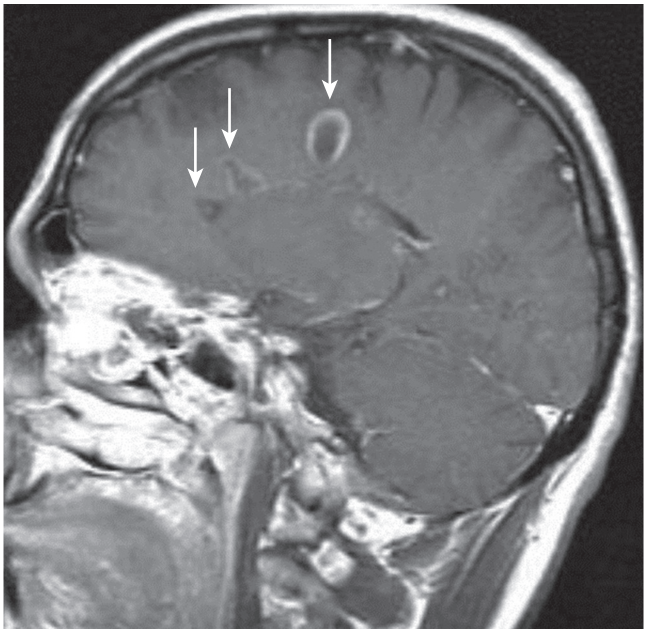


FIG 66.2 A post-gadolinium T1-weighed magnetic resonance imaging (MRI) scan of the brain showing Dawson fingers (arrows).

Classic actively demyelinating plaques are primarily seen during the RR stage of disease and generally decrease in frequency with increasing disease duration. Lesions more typical of progressive forms of MS have been termed "chronic active," or smoldering, and "chronic silent," or inactive. Chronic active plaques are distinguished by a rim of activated microglia and deposits of complement at the lesion edge, surrounding a hypocellular and gliotic core. They are slowly expansive as a consequence of active demyelination at the lesion edge. In contrast, chronic silent plaques have a sharp border. Other characteristics of silent plaques include prominent loss of oligodendrocytes and axons, pronounced astrogliosis, and a paucity of macrophages and activated microglia. Immunopathological changes in the so-called normal-appearing white matter (NAWM), outside of plaques, are pervasive in progressive MS but have also been observed in RRMS. These changes consist of diffuse axonal injury and microglial activation, as well as scattered lymphocytes.

MS is widely classified as a demyelinating disorder. The reason is that a large number of the nerve fiber segments traversing plaques demonstrate myelin loss with relative axonal sparing. However, it is now recognized that axonopathy also occurs and is, in fact, an early and prominent feature of acute MS lesions. Axonal damage results in dysmorphic mitochondria, focal swellings, fragmentation, and frank transections with terminal bulbs at the stumps. Mitochondrial abnormalities and focal swelling have been observed in fully myelinated axons within MS lesions, suggesting that they can occur independent of demyelination.[35] In animal models of MS, axons with abnormal mitochondria are restricted to areas of immune infiltration, and progressive axonal changes correlate with the density of infiltrates.[35] Hence, it is likely that the axonal damage is directly mediated by direct contact with inflammatory cells. Although demyelination can be reversed to some extent by remyelination, axonal transection is irreversible. Clinicopathological investigations have found that permanent motor disability in MS correlates with loss of corticospinal tract axons more so than with degree of demyelination.

Gray Matter Lesions

MS was traditionally considered a white matter disease. It is now established that the gray matter is affected as well. Three types of cortical lesions have been described: leukocortical (which span gray and white matter), intracortical, and subpial. All of these lesions show demyelination and oligodendrocyte loss, microglial activation, neuritic transections, neuronal death, and reduced presynaptic terminals. Subpial lesions are the most common. They can cover long distances of the cortical ribbon and usually extend to the cortical layer III or IV. Cortical lesions are not visible on conventional MRI scans and require special staining to be appreciated in CNS tissue sections. This explains why they were not recognized as common features of MS until recently. In fact, cortical demyelination and gray matter atrophy are evident from the earliest stages of disease, even before a clinically definite diagnosis can be made, and continue to advance at an increasing rate throughout the disease course. Extensive cortical demyelination is evident in the forebrain and cerebellum during progressive MS. Gray matter atrophy in individuals with MS correlates strongly with cognitive deficits and clinical disability.[36]

Meningeal Inflammation

White blood cell (WBC) counts tend to be within normal limits or only slightly elevated in the CSF of most patients with MS. Nonetheless, there is growing recognition that low-grade diffuse meningeal inflammation and focal perivascular meningeal inflammation are common. Meningeal inflammation is most prominent in progressive forms of MS but is prevalent in early MS as well. The meningeal infiltrates are topographically associated with cortical lesions. Lymphoid follicle–like structures, composed of proliferating B cells, T cells, and follicular dendritic cells (FDCs), have been observed in the meninges of up to 40% of autopsied brains from individuals with SPMS.[6] In almost every case, the follicles were found to reside in deep sulci and abut an underlying subpial lesion, suggesting that toxic factors are released by inflammatory cells in the follicles and diffuse into the brain parenchyma. The presence of lymphoid follicles has been associated with a more severe clinical course, shorter disease duration, and younger age at death.

KEY CONCEPTS

Pathology

- The hallmark of multiple sclerosis (MS) pathology is the focal demyelinated lesion, or "plaque," with perivascular inflammatory infiltration and focal blood–brain barrier (BBB) breakdown.
- Axonopathy is an early and prominent feature of acute MS lesions.
- Central nervous system (CNS) damage includes demyelination, apoptosis and loss of oligodendrocytes, and axonal swellings and transections.
- Both gray matter and white matter are affected.
- The pathological features of MS are heterogeneous and evolve over time.

IMMUNOPATHOGENESIS

Animal Models of MS

According to the current dogma, MS is an autoimmune disease mediated by CD4 T cells reactive against myelin antigens. The identification of HLA class II, IL-2Rα, and IL-7Rα as MS susceptibility loci is consistent with a role of CD4 T cells in MS pathogenesis. An autoimmune etiology is further supported by the animal model experimental autoimmune encephalomyelitis (EAE). EAE is a multifocal inflammatory demyelinating disease of the CNS that has striking histological and clinical similarities to MS (Fig. 66.3 and Fig. 66.4). It has been induced in a wide variety of mammalian species (including nonhuman primates, but most commonly in rodents) by vaccination against major histocompatibility complex (MHC) class II–restricted myelin epitopes. EAE can be transferred from myelin-vaccinated mice to syngeneic naïve hosts with purified CD4 T-cell lines or clones. These encephalitogenic myelin-specific CD4 T cells invariably fall within the Th1 or Th17 lineage and produce the proinflammatory cytokines interferon-γ (IFN-γ) and IL-17, respectively, in response to antigenic stimulation[37](Chapter 16). Both Th1 and Th17 cells produce granulocyte macrophage–colony-stimulating factor (GM-CSF), a monocyte mobilizing and growth factor that plays a critical role in many models of EAE. Upon activation in the periphery, myelin-reactive CD4 T cells upregulate adhesion molecules and chemokine receptors, thereby acquiring the ability to cross the BBB. Once having infiltrated the CNS, they are reactivated by local antigen-presenting cells (APCs), such as perivascular macrophages or microglia, which constitutively express MHC class II molecules bound to myelin peptides on their surface. GM-CSF, as well as other Th1 and/or Th17 cytokines, are subsequently released *in situ* and initiate an inflammatory cascade, resulting in the production of chemokines, mobilizing factors and vasoactive substances, upregulation of adhesion molecules on the cerebrovascular endothelium, and thus the recruitment of myeloid cells and lymphocytes from the circulation to the nascent plaque. GM-CSF may drive the differentiation of infiltrating monocytes and CNS-resident microglia into CD11c$^+$ DCs, which are among the most potent APCs. Adoptive transfer studies with labeled donor T cells have demonstrated that the myelin-specific T cells remain clustered in the perivascular space throughout lesion development. A secondary wave of myeloid cells infiltrate deep into the CNS

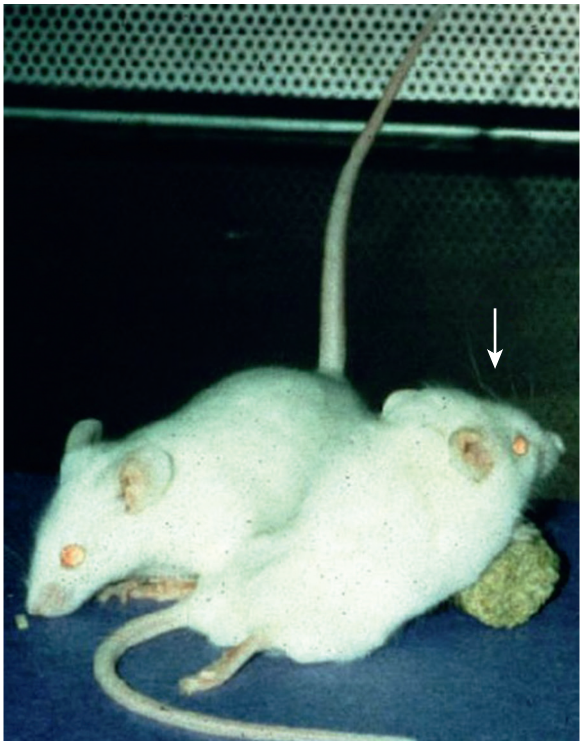


FIG 66.3 A mouse with experimental autoimmune encephalitis (EAE) (arrow) and a healthy littermate. The mouse with EAE has a limp tail and hindlimb weakness.

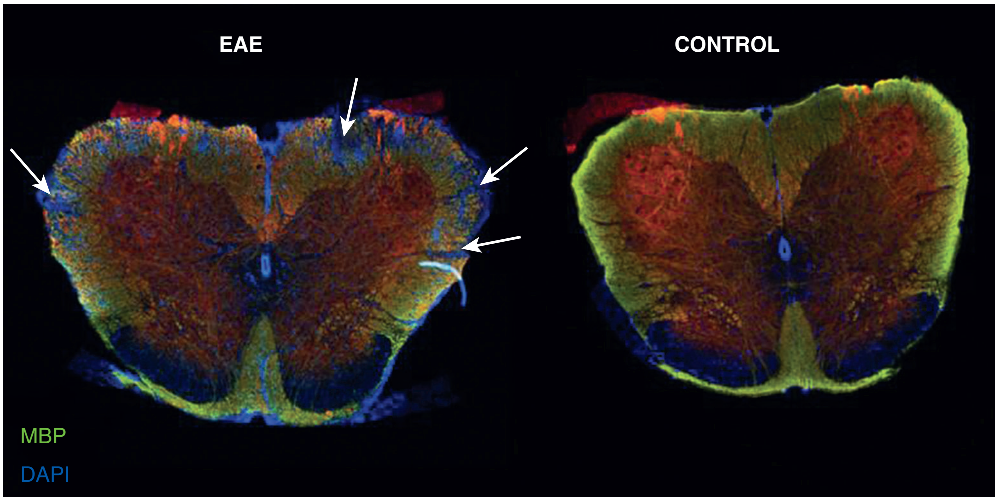


FIG 66.4 Immunofluorescent histology of spinal cord sections from a mouse with experimental autoimmune encephalitis (EAE) (left) and a healthy control (right). White matter tracks were stained with a monoclonal antibody (mAb) specific for myelin basic protein (green). The nuclei of inflammatory cells are stained with DAPI (4′,6-diamidino-2-phenylindole) (blue). The arrows point to areas of demyelination.

white matter, associate with nodes of Ranvier, and directly inflict damage to the myelin sheath and axons[35] (see Figs. 66.3 and Fig. 66.4).

Immune Dysregulation in Patients With MS

Studies on the frequency of myelin-reactive T cells in patients with MS have reported conflicting results; some investigators found a significantly higher incidence of myelin-specific peripheral blood mononuclear cells (PBMCs) in individuals with MS compared with age- and gender-matched healthy controls (HCs), whereas others found no significant difference.[5] Many of the earlier studies that found no differences between patients and HCs used proliferation as a measure of T-cell reactivity and nonhuman myelin proteins for antigenic stimulation. In contrast, several laboratories have found that untreated patients with RRMS have increased frequencies of PBMCs that produce IFN-γ or IL-17 in response to *ex vivo* challenge with human myelin basic protein (MBP), human proteolipid protein (PLP), or their constituent peptides.[4,38,39] IFN-γ responses to PLP peptides were shown to correlate with level of clinical disability.[38] T cells that coexpress IL-17 and IFN-γ were identified in the brain tissue of patients with MS, and circulating lymphocytes obtained from patients with MS were found to have an increased propensity to differentiate into IL-17/IFN-γ double producers.[40] T cells coexpressing IL-17/IFN-γ, and IL-17–producing T cells that convert into IFN-γ producers (the so-called ex-Th17 cells), have been implicated in the pathogenesis of some EAE models. In a recent study, untreated patients with RRMS had a higher frequency of peripheral blood T cells that expressed intracellular GM-CSF in response to stimulation with phorbol myristate acetate and ionomycin, compared with HCs or IFN-β–treated RRMS patients.[41] Approximately 15–20% of both CD4 and CD8 T cells in active MS lesions expressed GM-CSF and the majority coexpressed IL-17 and/or IFN-γ.

There is an increasing appreciation that the cytokine dysregulation that occurs during EAE and MS is heterogeneous. Hence, clinically identical forms of EAE can be induced with stable murine Th1 or Th17 cells independently.[42] Subsets of patients with MS who consistently mounted either IFN-γ– or IL-17–skewed responses to human MBP over the course of a year were recently identified, while others exhibited mixed or oscillating responses.[37] Diversity in the immune pathways that drive MS pathology holds important implications for the development of a more personalized approach to the management of patients in the future. Indeed, Th1- and Th17-mediated forms of EAE show different patterns of responsiveness to the same immunomodulatory drug.[42]

Perhaps the strongest evidence of an autoimmune basis of inflammatory demyelination in humans comes from clinical trials of immunomodulatory agents. Patients with RRMS were treated with an altered peptide ligand (APL) of MBP with the intention of tolerizing MBP-reactive T cells or deflecting their differentiation toward an immunosuppressive, regulatory, or innocuous Th2 phenotype. Unexpectedly, administration of the APL was temporally associated with expansion of circulating MBP-reactive Th1 cells and clinical worsening in a subgroup of patients.[43] In contrast, as will be discussed in detail below, drugs that impede lymphocyte trafficking to the CNS,[44,45] block growth factor signaling into T lymphocytes,[46] or deplete lymphocytes from the periphery[47] suppress MS relapse rates and the accumulation of MRI lesions.

DISEASE-MODIFYING THERAPIES

The approval of IFN-β-1b (Betaseron) by the US Food and Drug Administration (FDA) in 1993 for the management of RRMS marked the beginning of a new era in MS therapeutics. In the intervening 23 years, an additional 13 disease-modifying therapies have been approved, all of which have significantly reduced the annualized relapse rate (in a range from 25% to almost 70%) and the frequency of gadolinium enhancing (acutely inflamed) MRI lesions. Prior to 1993, corticosteroids were the only class of drugs routinely used to treat MS. Although corticosteroids accelerate the rate of recovery from acute exacerbations, there is little evidence that they alter the ultimate clinical outcome or prevent subsequent disease activity. Therefore the introduction of disease-modifying therapies has represented a major advance in the treatment of individuals with RRMS and has had a profound impact by mitigating morbidity and enhancing quality of life.

Recombinant IFN-β

IFN-β is a type I IFN with potent antiviral properties. It has pleiotropic effects on the innate immune system. Recombinant IFN-β was first tested in MS on the basis of the contemporaneous theory that the disease was caused by an active viral infection of the CNS. Although MS is now believed to be an autoimmune disease, recombinant IFN-β therapy was serendipitously found to significantly reduce annualized MS relapse rates. It is manufactured as four different commercial products, all of which are self-administered via either subcutaneous or intramuscular injection. There are two distinct structural forms. IFN-β-1a is produced in mammalian cells and has the same sequence as the naturally occurring compound, while IFN-β-1b is produced in modified *Escherichia coli* and has a Met-1 deletion and a Cys-17 to Ser mutation. IFN-β-1a is glycosylated, and IFN-β-1b is nonglycosylated. Studies of the efficacy of IFN-β in RRMS have yielded remarkably consistent results across different formulations of the drug. In multiple randomized, double-blinded placebo-controlled trials, IFN-β therapy reduced annual relapse rates by 20–35%.[48] However, approximately 30% of patients with MS do not respond to the drug. The mechanism of action of IFN-β in MS has not been definitively elucidated. Proposed mechanisms include induction of the immunosuppressive cytokine IL-10, inhibition of pathogenic Th17 cells, and stabilization of the BBB via direct effects on the cerebrovascular endothelium.

Glatiramer Acetate

GA comprises a mixture of polypeptides of different lengths and sequences, synthesized by the randomized polymerization of four amino acids, namely, glutamic acid, lysine, alanine, and tyrosine. These are the most prevalent amino acids in MBP, a candidate autoantigen in MS. It was originally thought that GA would act as a competitive antagonist of MBP peptides for binding to MHC class II molecules on APCs. Subsequent mechanistic studies did not support that theory. The mechanism of action of GA in MS is currently unknown, but alternative hypotheses that have been proposed include immune deviation of myelin-reactive T cells from a destructive Th1 to an innocuous Th2 phenotype, increase in frequency and function of FoxP3$^+$ regulatory CD4 T cells or regulatory CD8 T cells, and induction of antiinflammatory type II monocytes. A pivotal trial of GA versus placebo in 1995 demonstrated a mean reduction in the relapse rate of approximately 30%.[49] In contrast, a multinational, multicenter, double-blind, placebo-controlled phase III trial of GA in PPMS was stopped after an interim analysis by an independent data safety monitoring board indicated no discernible treatment effect on the primary outcome measure.[8] GA, administered by subcutaneous injection, was approved by the FDA in 1996 for reducing the frequency of relapses, but not for reducing the progression of disability.

Teriflunomide

Teriflunomide, an oral pyrimidine synthesis inhibitor, was approved by the FDA in 2012 for relapsing forms of MS. It has broad immunosuppressive effects, including a cytostatic effect on proliferating T and B lymphocytes. In a pivotal phase III trial, teriflunomide reduced annualized relapse rate by 31% compared with placebo.[50] In an independent phase III rater-blinded comparator trial, teriflunomide was found to have a similar efficacy to IFN-β-1a in reducing the proportion of participants with at least one relapse over 1 year.[51] Teriflunomide is teratogenic in rats and rabbits and is therefore contraindicated in pregnant women and women of childbearing potential not using reliable contraception.

Dimethyl Fumarate

Dimethyl fumarate (DMF) is the methyl ester of fumaric acid. Prior to being tested as a DMA in MS, DMF was used as a biocide in furniture and shoes to prevent the growth of molds during storage or transport. A combination of DMF and three other fumaric acid esters was marketed in Germany as a treatment for psoriasis. In two phase III trials conducted in 2012, oral DMF was demonstrated to decrease the annualized relapse rate of adults with RRMS by approximately 34–50% compared with placebo.[52,53] An additional cohort of subjects treated with GA was included as a reference comparator in one of the trials.[53] There was no significant difference in the effect of twice-daily oral DMF versus daily subcutaneous GA on relapse rate, but significantly fewer new or enlarging hyperintense lesions were noted on T2-weighted MRI images in participants treated with DMF. These results led to FDA approval in 2013. DMF therapy is often associated with lymphopenia. Therapeutic efficacy does not appear to be inversely related to lymphocyte counts. As with IFN-β and GA, the mechanism by which DMF suppresses MS disease activity remains unclear. The most common side effects of DMF are flushing and gastrointestinal upset. In the postmarketing setting, several cases of progressive multifocal leukoencephalopathy (PML), a rare viral infection of the brain, were reported in patients who were taking DMF and had persistent lymphopenia. PML has also been reported as a complication of other DMAs, including fingolimod and, most notably, natalizumab.

Fingolimod

During homeostasis, sphingosine 1 phosphate receptor 1 (S1P1) signaling into lymphocytes drives their egress from lymph nodes into the bloodstream. Fingolimod is an orally administered S1P1 receptor modulator that effectively traps myelin-specific T cells in lymph nodes so that they cannot reenter the circulation and gain access to the CNS. In a 24-month randomized, double-blinded, placebo-controlled trial, fingolimod was shown to reduce annualized relapse rates by 50%.[54] Furthermore, the probability of disease progression at the 24-month follow-up was lower in the fingolimod-treated subjects. In an independent 12-month, double-blind, double-dummy phase III comparator study, fingolimod reduced the annualized relapse rate to a range of 0.16–0.20, as compared with 0.33 for IFN-β-1a, corresponding to a relative reduction of 38–52%.[45] In 2010, fingolimod became the first oral DMA approved by the FDA to reduce relapses and delay disability progression in patients with relapsing forms of MS. However, in a phase III randomized, double-blind, placebo-controlled trial, fingolimod did not slow disease progression in patients with PPMS.[11,12] Fingolimod does not distinguish between pathogenic and protective lymphocytes, and a number of opportunistic infections have emerged as complications of the treatment. Patients taking fingolimod are susceptible herpes virus infections, particularly shingles. In the phase III comparator study with IFN-β-1a, two fatal infections occurred among the fingolimod-treated subjects: disseminated primary varicella zoster and herpes simplex encephalitis.[45] Consequently, patients are screened for immunity to varicella zoster before commencing treatment. In the postmarketing setting, there have been at least three cases of PML in fingolimod-treated patients with no or only distant prior exposure to immunosuppressant drugs. Other potential side effects of fingolimod include macular edema, bradycardia, and atrioventricular block.

Natalizumab

Natalizumab is a humanized mAb reactive against the cell adhesion molecule α_4 integrin, which is widely expressed on lymphocytes and monocytes. Natalizumab is believed to mediate its ameliorative effects in RRMS by blocking interactions between the very late antigen-4 (VLA-4; a heterodimer composed of the α_4 and β_1 integrin chains) on leukocytes with its cognate ligand, vascular cell adhesion molecule (VCAM)-1, on cerebrovascular endothelial cells. VLA-4–VCAM-1 interactions are required for the passage of lymphocytes and monocytes past the BBB. Natalizumab was approved by the FDA for relapsing MS in 2004 following a 2-year phase III trial. Natalizumab reduced the relapse rate at 1 year by 68% and the number of gadolinium-enhancing MRI lesions at both 1 year and 2 years by over 90%, compared with placebo.[44] Natalizumab also reduced the risk of sustained progression of disability by 42% over 2 years. In an independent placebo-controlled trial, the addition of natalizumab to IFN-β-1a suppressed break through disease activity and reduced the risk of sustained disability in patients who were refractory to IFN-β alone.[55] The effects of natalizumab on leukocyte trafficking are not specific to the CNS. In fact, in 2008, the FDA approved natalizumab for the treatment of refractory Crohn disease. The most serious complication of natalizumab treatment is PML; there have been over 400 reported cases. Factors that increase the risk of natalizumab-associated PML include John Cunningham (JC) virus seropositivity, duration of treatment >2 years, and prior exposure to immunosuppressant drugs.

Alemtuzumab

Alemtuzumab is a humanized anti-CD52 mAb that globally depletes circulating T and B lymphocytes via antibody-dependent cell-mediated cytotoxicity, complement-dependent cytolysis, and induction of apoptosis. CD52 is primarily expressed on the cell surface of T and B lymphocytes, but it is also expressed at lower levels on macrophages, eosinophils, and natural killer (NK) cells. The function of CD52 is unknown, although there is some evidence that it is involved in T-cell activation. Hematopoietic stem cells (HSCs) do not express CD52, allowing the subsequent repopulation of circulating lymphocyte pools following alemtuzumab treatment, which occurs at variable rates between lymphocyte subsets. Reconstituted circulating T-cell pools have been found to be enriched in $CD4^{+}CD25^{high}CD127^{low}FOXP3^{+}$ regulatory T cells (Tregs) that might be responsible, in part, for the long-lasting therapeutic effect of alemtuzumab, that has been observed in some individuals.[56] In a 2-year, rater-masked, randomized controlled phase III trial, previously untreated patients with RRMS were randomly allocated to receive intravenous alemtuzumab or subcutaneous IFN-β-1a.[57] Alemtuzumab was given once per day for 5 days at baseline and once per day for 3 days at 12 months. Relapse rates were reduced by 54.9% in the alemtuzumab group compared with the IFN-β-1a group. In an independent study, alemtuzumab reduced relapse rates and the risk of sustained accumulation of disability in patients with RRMS refractory to first-line DMAs.[47] Although alemtuzumab suppressed new MRI lesion formation and superimposed relapses in patients with SPMS, it did not prevent clinical progression or progressive cerebral atrophy.[7]

Despite causing a profound lymphopenia, alemtuzumab therapy has only rarely been associated with opportunistic infections and has not been associated an increased incidence of malignancy. Surprisingly, the principal adverse effect of alemtuzumab in MS is antibody-mediated autoimmune disease, most commonly Graves disease. Idiopathic thrombocytopenic purpura, Goodpasture syndrome, and antiglomerular basement membrane disease have been reported less frequently. The underlying mechanism is not fully understood but may be a consequence of homeostatic T-cell proliferation following lymphoablation, leading to the generation of chronically activated oligoclonal CD4 and CD8 T cells capable of producing proinflammatory cytokines.[58]

Daclizumab

Daclizumab is a humanized mAb reactive to CD25, the α subunit of the IL-2 receptor. Daclizumab was originally proposed as a treatment for RRMS based on the hypothesis that it would block IL-2–dependent expansion of effector T cells. However, mechanistic studies have shown that daclizumab does not inhibit T-cell survival, proliferation, or cytokine production. Some patients treated with daclizumab exhibit an expansion of circulating $CD56^{bright}$ NK cells, which correlates with treatment response.[59] $CD56^{bright}$ NK cells have been shown to possess immunoregulatory properties and to kill activated autologous T cells *in vitro.* Daclizumab might also modulate lymphoid tissue inducer cells in patients with MS and thereby suppress the development of meningeal lymphoid follicles.[60] In a phase III randomized placebo-controlled clinical trial in RRMS, subcutaneous administration of daclizumab reduced the annualized relapse rate by 45%, the proportion of patients who relapsed by 41%, and new or newly enlarged T2-weighted MRI lesions by 54%, compared with intramuscular IFN-β-1a over 144 weeks.[61] It was approved by the FDA in 2016 for the treatment of relapsing MS in adults. Side effects include an increased risk of infections, hepatotoxicity, noninfectious colitis, lymphadenopathy, and cutaneous events, such as rash and eczema. On the basis of its safety profile, daclizumab is generally reserved for patients who have an inadequate response to ≥2 other disease modifying agents.

B Cell–Depleting Monoclonal Antibodies

A role of B cells in the pathogenesis of MS has been suspected since the discovery of unique oligoclonal bands in the CSF of the majority of individuals with MS, indicative of intrathecal antibody synthesis. Nonetheless, B cells are not a prominent component of the perivascular infiltrates in MS lesions. This apparent paradox has been resolved, at least in part, by the discovery of lymphoid follicle–like structures, composed of proliferating B cells, antibody-producing plasma cells, Th cells, and FDCs, in the meninges of some individuals with MS.[6] Although meningeal lymphoid follicles have primarily been detected in autopsy specimens from persons with progressive MS, it is possible that nascent immune cell aggregates develop in the RR stage but are disrupted during the dissection and processing of CNS tissues. The most direct evidence for a role of B cells as effector cells in MS pathogenesis comes from trials of mAbs directed against CD20, a B cell–specific surface molecule. CD20 is expressed on preB cells and mature B cells. It is not expressed on antibody-secreting plasma cells. Rituximab, a genetically engineered chimeric anti-CD20 mAb, induces a rapid depletion of circulating B cells that lasts approximately 6–9 months. In a phase II clinical trial of rituximab in RRMS, subjects in the active treatment arm experienced a significant reduction in clinical relapses and in the formation of new or enhancing MRI lesions compared with subjects in the placebo arm.[62] Ocrelizumab is a next-generation fully human recombinant anti-CD20 mAb that binds to a different epitope from rituximab and with higher affinity. In two large phase III studies, ocrelizumab reduced the annualized relapse rate in subjects with RRMS by nearly 50% compared with IFN-β-1a over a 2-year period.[63] Additionally, ocrelizumab delayed confirmed disability progression by approximately 40% and the total number of gadolinium-enhancing lesions by >90% compared with IFNβ-1a.

Interestingly, some patients with progressive MS might also benefit from the depletion of $CD20^+$ cells, particularly if there is evidence of ongoing activity at the time of treatment initiation. Hawker et al. found that rituximab significantly delayed the time to confirmed disease progression in a subset of patients with primary progressive MS who had gadolinium enhancing lesions on baseline MRI scans of the brain.[9] The benefit was most dramatic in subjects aged <51 years. In a recently completed trial of ocrelizumab in PPMS, the risk of progression of clinical disability was reportedly reduced by 24%.[63] As a result, in early 2016, the FDA granted ocrelizumab "breakthrough therapy designation" for PPMS.

The mechanism by which B cell–depleting therapy alleviates MS does not appear to be reduction in antibody titers. Rituximab depletes B cells from the CSF of patients with RRMS with little effect on CSF IgG levels.[64] B cells are professional APCs, and there is some evidence that they are important for sustaining myelin-specific Th17 responses in MS. The frequency of GM-CSF–producing B cells was found to be elevated in RRMS.[65] Hence, ocrelizumab and rituximab may eliminate an important cellular source of GM-CSF, a monocyte/macrophage–mobilizing cytokine that has been implicated in the pathogenesis of autoimmune demyelinating disease. Disruption of meningeal follicle–like structures is yet another putative mechanism of action that might be particularly relevant to progressive forms of MS.

THERAPEUTIC PRINCIPLES

Pharmaceutical Management of MS

- Corticosteroids accelerate the rate of recovery from multiple sclerosis (MS) relapses, but there is little evidence that they impact the ultimate degree of recovery or the future clinical course.
- The US Food and Drug Administration (FDA) has approved 14 disease-modifying agents (DMAs) that reduce annualized relapse rates in patients with RRMS by approximately 20–70%.
- Newer generation DMAs either deplete lymphocytes, inhibit their expansion, or block their migration to the central nervous system (CNS).
- DMAs differ in efficacy and safety profiles. Opportunistic infection, in particular progressive multifocal leukoencephalopathy, has been observed in association with several DMAs.
- The choice of DMAs must be customized on an individual basis, taking into account disease activity and risk tolerance.

FUTURE DIRECTIONS

Dramatic advances have been made in the management of RRMS, but much remains to be done. The DMAs currently in use have global effects on lymphocytes, as opposed to specifically targeting autoreactive Th1 and Th17 cells. Consequently, protective as well as pathogenic immune responses may be undermined, increasing the risk of infection. Furthermore, DMAs are generally ineffective in progressive forms of disease. Critical goals of the MS research community at present are to better define the cellular and molecular mechanisms that link neuroinflammation to end-organ injury (namely, demyelination and axonopathy) and to elucidate the pathogenic pathways that underlie clinical progression. It is also imperative to develop a deeper understanding of obstacles to endogenous repair pathways in the CNS. This knowledge could ultimately lead to the discovery of laboratory-based surrogate biomarkers of both acute and chronic disease activity, as well as drugs that block, or even slow, disability accumulation in progressive MS and promote remyelination and axonal regeneration across clinical subtypes.

ON THE HORIZON

- The introduction of disease-modifying therapies has represented a major advance in the treatment of individuals with relapsing–remitting multiple sclerosis (RRMS). However, these drugs have global effects on lymphocytes, thereby increasing the risk of infection. A future goal will be to develop drugs that specifically target pathogenic leukocytes and/or autoreactive lymphocytes.
- Disease-modifying therapy is currently focused on the modulation of lymphocytes. Drugs that target innate immune cells may enhance MS therapy in the future.
- Recent studies suggest that B cell–depleting therapy may be beneficial for some patients with primary progressive MS (PPMS). There is a dire need for drugs that slow, or even halt, disability accumulation in progressive forms of MS and that are effective in older patients without evidence of recent neuroinflammatory activity.
- An increased understanding of the mechanisms by which inflammatory cells cause central nervous system (CNS) damage and of the obstacles to endogenous repair pathways in MS may inform the development of neuroprotective and neuroregenerative agents.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Koch M, Kingwell E, Rieckmann P, et al. The natural history of secondary progressive multiple sclerosis. J Neurol Neurosurg Psychiatry 2010;81(9):1039–43. PubMed PMID: 20639385.
2. Cree B, Gourraud PA, Oksenberg JR, et al. Long-term evolution of multiple sclerosis disability in the treatment era. Ann Neurol 2016 Jul 27. PubMed PMID: 27464262.
3. Mahad DH, Trapp BD, Lassmann H. Pathological mechanisms in progressive multiple sclerosis. Lancet Neurol 2015;14(2):183–93. PubMed PMID: 25772897.
4. Huber AK, Wang L, Han P, et al. Dysregulation of the IL-23/IL-17 axis and myeloid factors in secondary progressive MS. Neurology 2014;83(17):1500–7. PubMed PMID: 25253754. Pubmed Central PMCID: 4222856.
5. Segal BM. Stage-specific immune dysregulation in multiple sclerosis. J Interferon Cytokine Res 2014;34(8):633–40. PubMed PMID: 25084180. Pubmed Central PMCID: 4128245.
6. Magliozzi R, Howell O, Vora A, et al. Meningeal B-cell follicles in secondary progressive multiple sclerosis associate with early onset of disease and severe cortical pathology. Brain 2007;130(Pt 4):1089–104. PubMed PMID: 17438020.
7. Coles AJ, Cox A, Le Page E, et al. The window of therapeutic opportunity in multiple sclerosis: evidence from monoclonal antibody therapy. J Neurol 2006;253(1):98–108. PubMed PMID: 16044212.

8. Wolinsky JS, Narayana PA, O'Connor P, et al. Glatiramer acetate in primary progressive multiple sclerosis: results of a multinational, multicenter, double-blind, placebo-controlled trial. Ann Neurol 2007;61(1):14–24. PubMed PMID: 17262850.
9. Hawker K, O'Connor P, Freedman MS, et al. Rituximab in patients with primary progressive multiple sclerosis: results of a randomized double-blind placebo-controlled multicenter trial. Ann Neurol 2009;66(4):460–71. PubMed PMID: 19847908. eng.
10. Polman CH, Reingold SC, Banwell B, et al. Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. Ann Neurol 2011;69(2):292–302. PubMed PMID: 21387374. Pubmed Central PMCID: PMC3084507. eng.
11. Willer CJ, Dyment DA, Risch NJ, et al. Twin concordance and sibling recurrence rates in multiple sclerosis. Proc Natl Acad Sci USA 2003;100(22):12877–82. PubMed PMID: 14569025. Pubmed Central PMCID: 240712.
12. Hollenbach JA, Oksenberg JR. The immunogenetics of multiple sclerosis: A comprehensive review. J Autoimmun 2015;64:13–25. PubMed PMID: 26142251. Pubmed Central PMCID: 4687745.
13. Baranzini SE, Mudge J, van Velkinburgh JC, et al. Genome, epigenome and RNA sequences of monozygotic twins discordant for multiple sclerosis. Nature 2010;464(7293):1351–6. PubMed PMID: 20428171. Pubmed Central PMCID: 2862593.
14. Kurtzke JF. An epidemiologic approach to multiple sclerosis. Arch Neurol 1966;14(2):213–22. PubMed PMID: 4952448.
15. Munger KL, Levin LI, Hollis BW, et al. Serum 25-hydroxyvitamin D levels and risk of multiple sclerosis. JAMA 2006;296(23):2832–8. PubMed PMID: 17179460.
16. Munger KL, Aivo J, Hongell K, et al. Vitamin D Status During Pregnancy and Risk of Multiple Sclerosis in Offspring of Women in the Finnish Maternity Cohort. JAMA Neurol 2016;73(5):515–19. PubMed PMID: 26953778. Pubmed Central PMCID: 4861670.
17. Levin LI, Munger KL, O'Reilly EJ, et al. Primary infection with the Epstein-Barr virus and risk of multiple sclerosis. Ann Neurol 2010;67(6):824–30. PubMed PMID: 20517945. Pubmed Central PMCID: 3089959.
18. Ascherio A, Munger KL, Lennette ET, et al. Epstein-Barr virus antibodies and risk of multiple sclerosis: a prospective study. JAMA 2001;286(24):3083–8. PubMed PMID: 11754673.
19. Lang HL, Jacobsen H, Ikemizu S, et al. A functional and structural basis for TCR cross-reactivity in multiple sclerosis. Nat Immunol 2002;3(10):940–3. PubMed PMID: 12244309.
20. Magliozzi R, Serafini B, Rosicarelli B, et al. B-cell enrichment and Epstein-Barr virus infection in inflammatory cortical lesions in secondary progressive multiple sclerosis. J Neuropathol Exp Neurol 2013;72(1):29–41. PubMed PMID: 23242282.
21. Wu Z, Nagano I, Asano K, et al. Infection of non-encapsulated species of Trichinella ameliorates experimental autoimmune encephalomyelitis involving suppression of Th17 and Th1 response. Parasitol Res 2010;107(5):1173–88. PubMed PMID: 20661746.
22. Correale J. Helminth/Parasite treatment of multiple sclerosis. Curr Treat Options Neurol 2014;16(6):296. PubMed PMID: 24744099.
23. Hedstrom AK, Olsson T, Alfredsson L. High body mass index before age 20 is associated with increased risk for multiple sclerosis in both men and women. Mult Scler 2012;18(9):1334–6. PubMed PMID: 22328681.
24. Langer-Gould A, Brara SM, Beaber BE, et al. Childhood obesity and risk of pediatric multiple sclerosis and clinically isolated syndrome. Neurology 2013;80(6):548–52. PubMed PMID: 23365063. Pubmed Central PMCID: 3589288.
25. Mokry LE, Ross S, Timpson NJ, et al. Obesity and Multiple Sclerosis: A Mendelian Randomization Study. PLoS Med 2016;13(6):e1002053. PubMed PMID: 27351487. Pubmed Central PMCID: 4924848.
26. Poorolajal J, Bahrami M, Karami M, et al. Effect of smoking on multiple sclerosis: a meta-analysis. J Public Health (Bangkok) 2016 May 8. PubMed PMID: 27160862.
27. Ramanujam R, Hedstrom AK, Manouchehrinia A, et al. Effect of Smoking Cessation on Multiple Sclerosis Prognosis. JAMA Neurol 2015;72(10):1117–23. PubMed PMID: 26348720.
28. Gao Z, Nissen JC, Ji K, et al. The experimental autoimmune encephalomyelitis disease course is modulated by nicotine and other cigarette smoke components. PLoS ONE 2014;9(9):e107979. PubMed PMID: 25250777. Pubmed Central PMCID: 4176721.
29. Farez MF, Fiol MP, Gaitan MI, et al. Sodium intake is associated with increased disease activity in multiple sclerosis. J Neurol Neurosurg Psychiatry 2015;86(1):26–31. PubMed PMID: 25168393.
30. Nourbakhsh B, Graves J, Casper TC, et al. Dietary salt intake and time to relapse in paediatric multiple sclerosis. J Neurol Neurosurg Psychiatry 2016;87:1350–3. PubMed PMID: 27343226.
31. Kleinewietfeld M, Manzel A, Titze J, et al. Sodium chloride drives autoimmune disease by the induction of pathogenic TH17 cells. Nature 2013;496(7446):518–22. PubMed PMID: 23467095. Pubmed Central PMCID: 3746493.
32. Voskuhl RR, Wang H, Wu TC, et al. Estriol combined with glatiramer acetate for women with relapsing-remitting multiple sclerosis: a randomised, placebo-controlled, phase 2 trial. Lancet Neurol 2016;15(1):35–46. PubMed PMID: 26621682.
33. Bove R, Musallam A, Healy BC, et al. Low testosterone is associated with disability in men with multiple sclerosis. Mult Scler 2014;20(12):1584–92. PubMed PMID: 24710799. Pubmed Central PMCID: 4188801.
34. Kurth F, Luders E, Sicotte NL, et al. Neuroprotective effects of testosterone treatment in men with multiple sclerosis. NeuroImage Clin 2014;4:454–60. PubMed PMID: 24634831. Pubmed Central PMCID: 3952353.
35. Nikic I, Merkler D, Sorbara C, et al. A reversible form of axon damage in experimental autoimmune encephalomyelitis and multiple sclerosis. Nat Med 2011;17(4):495–9. PubMed PMID: 21441916.
36. Preziosa P, Rocca MA, Pagani E, et al. Structural MRI correlates of cognitive impairment in patients with multiple sclerosis: A Multicenter Study. Hum Brain Mapp 2016;37(4):1627–44. PubMed PMID: 26833969.
37. Carbajal KS, Mironova Y, Ulrich-Lewis JT, et al. Th Cell Diversity in Experimental Autoimmune Encephalomyelitis and Multiple Sclerosis. J Immunol 2015;195(6):2552–9. PubMed PMID: 26238492. Pubmed Central PMCID: 4561206.
38. Moldovan IR, Rudick RA, Cotleur AC, et al. Interferon gamma responses to myelin peptides in multiple sclerosis correlate with a new clinical measure of disease progression. J Neuroimmunol 2003;141(1-2):132–40. PubMed PMID: 12965264. eng.
39. Hedegaard CJ, Krakauer M, Bendtzen K, et al. T helper cell type 1 (Th1), Th2 and Th17 responses to myelin basic protein and disease activity in multiple sclerosis. Immunology 2008;125(2):161–9. PubMed PMID: 18397264. Pubmed Central PMCID: PMC2561132. eng.
40. Kebir H, Ifergan I, Alvarez JI, et al. Preferential recruitment of interferon-gamma-expressing TH17 cells in multiple sclerosis. Ann Neurol 2009;66(3):390–402. PubMed PMID: 19810097. eng.
41. Rasouli J, Ciric B, Imitola J, et al. Expression of GM-CSF in T Cells Is Increased in Multiple Sclerosis and Suppressed by IFN-beta Therapy. J Immunol 2015;194(11):5085–93. PubMed PMID: 25917097. Pubmed Central PMCID: 4433790.
42. Kroenke MA, Carlson TJ, Andjelkovic AV, et al. IL-12- and IL-23-modulated T cells induce distinct types of EAE based on histology, CNS chemokine profile, and response to cytokine inhibition. J Exp Med 2008;205(7):1535–41. PubMed PMID: 18573909. Pubmed Central PMCID: 2442630. Epub 2008/06/25. eng.
43. Bielekova B, Goodwin B, Richert N, et al. Encephalitogenic potential of the myelin basic protein peptide (amino acids 83-99) in multiple sclerosis: results of a phase II clinical trial with an altered peptide ligand. Nat Med 2000;6(10):1167–75. PubMed PMID: 11017150. eng.
44. Polman CH, O'Connor PW, Havrdova E, et al. A randomized, placebo-controlled trial of natalizumab for relapsing multiple sclerosis. N Engl J Med 2006;354(9):899–910. PubMed PMID: 16510744. Epub 2006/03/03. eng.
45. Cohen JA, Barkhof F, Comi G, et al. Oral fingolimod or intramuscular interferon for relapsing multiple sclerosis. N Engl J Med 2010;362(5):402–15. PubMed PMID: 20089954. eng.
46. Gold R, Giovannoni G, Selmaj K, et al. Daclizumab high-yield process in relapsing-remitting multiple sclerosis (SELECT): a randomised,

double-blind, placebo-controlled trial. Lancet 2013;381(9884):2167–75. PubMed PMID: 23562009. eng.

47. Coles AJ, Twyman CL, Arnold DL, et al. Alemtuzumab for patients with relapsing multiple sclerosis after disease-modifying therapy: a randomised controlled phase 3 trial. Lancet 2012;380(9856):1829–39. PubMed PMID: 23122650. eng.
48. Oliver BJ, Kohli E, Kasper LH. Interferon therapy in relapsing-remitting multiple sclerosis: a systematic review and meta-analysis of the comparative trials. J Neurol Sci 2011;302(1-2):96–105. PubMed PMID: 21167504.
49. Johnson KP, Brooks BR, Cohen JA, et al. Copolymer 1 reduces relapse rate and improves disability in relapsing-remitting multiple sclerosis: results of a phase III multicenter, double-blind placebo-controlled trial. The Copolymer 1 Multiple Sclerosis Study Group. Neurology 1995;45(7):1268–76. PubMed PMID: 7617181. Epub 1995/07/01. eng.
50. O'Connor P, Wolinsky JS, Confavreux C, et al. Randomized trial of oral teriflunomide for relapsing multiple sclerosis. N Engl J Med 2011;365(14):1293–303. PubMed PMID: 21991951. eng.
51. Vermersch P, Czlonkowska A, Grimaldi LM, et al. Teriflunomide versus subcutaneous interferon beta-1a in patients with relapsing multiple sclerosis: a randomised, controlled phase 3 trial. Mult Scler 2014;20(6):705–16. PubMed PMID: 24126064.
52. Gold R, Kappos L, Arnold DL, et al. Placebo-controlled phase 3 study of oral BG-12 for relapsing multiple sclerosis. N Engl J Med 2012;367(12):1098–107. PubMed PMID: 22992073.
53. Fox RJ, Miller DH, Phillips JT, et al. Placebo-controlled phase 3 study of oral BG-12 or glatiramer in multiple sclerosis. N Engl J Med 2012;367(12):1087–97. PubMed PMID: 22992072.
54. Kappos L, Radue EW, O'Connor P, et al. A placebo-controlled trial of oral fingolimod in relapsing multiple sclerosis. N Engl J Med 2010;362(5):387–401. PubMed PMID: 20089952. Epub 2010/01/22. eng.
55. Rudick RA, Stuart WH, Calabresi PA, et al. Natalizumab plus interferon beta-1a for relapsing multiple sclerosis. N Engl J Med 2006;354(9):911–23. PubMed PMID: 16510745.
56. De Mercanti S, Rolla S, Cucci A, et al. Alemtuzumab long-term immunologic effect: Treg suppressor function increases up to 24 months. Neurol Neuroimmunol Neuroinflamm 2016;3(1):e194. PubMed PMID: 26819963. Pubmed Central PMCID: 4723135.
57. Cohen JA, Coles AJ, Arnold DL, et al. Alemtuzumab versus interferon beta 1a as first-line treatment for patients with relapsing-remitting multiple sclerosis: a randomised controlled phase 3 trial. Lancet 2012;380(9856):1819–28. PubMed PMID: 23122652.
58. Jones JL, Thompson SA, Loh P, et al. Human autoimmunity after lymphocyte depletion is caused by homeostatic T-cell proliferation. Proc Natl Acad Sci USA 2013;110(50):20200–5. PubMed PMID: 24282306. Pubmed Central PMCID: 3864306.
59. Bielekova B, Catalfamo M, Reichert-Scrivner S, et al. Regulatory CD56(bright) natural killer cells mediate immunomodulatory effects of IL-2Ralpha-targeted therapy (daclizumab) in multiple sclerosis. Proc Natl Acad Sci USA 2006;103(15):5941–6. PubMed PMID: 16585503. Pubmed Central PMCID: 1458677.
60. Perry JS, Han S, Xu Q, et al. Inhibition of LTi cell development by CD25 blockade is associated with decreased intrathecal inflammation in multiple sclerosis. Sci Transl Med 2012;4(145):145ra06. PubMed PMID: 22855463. Pubmed Central PMCID: 3846177.
61. Kappos L, Wiendl H, Selmaj K, et al. Daclizumab HYP versus Interferon Beta-1a in Relapsing Multiple Sclerosis. N Engl J Med 2015;373(15):1418–28. PubMed PMID: 26444729.
62. Hauser SL, Waubant E, Arnold DL, et al. B-cell depletion with rituximab in relapsing-remitting multiple sclerosis. N Engl J Med 2008;358(7):676–88. PubMed PMID: 18272891. eng.
63. Sastre-Garriga J, Wiendl H. Highlights from the 31st ECTRIMS congress - Barcelona 2015. Mult Scler 2016;22(1):7–10. PubMed PMID: 26684283.
64. Cross AH, Stark JL, Lauber J, et al. Rituximab reduces B cells and T cells in cerebrospinal fluid of multiple sclerosis patients. J Neuroimmunol 2006;180(1-2):63–70. PubMed PMID: 16904756. Pubmed Central PMCID: 1769354.
65. Li R, Rezk A, Miyazaki Y, et al. Proinflammatory GM-CSF-producing B cells in multiple sclerosis and B cell depletion therapy. Sci Transl Med 2015;7(310):310ra166. PubMed PMID: 26491076.

MULTIPLE-CHOICE QUESTIONS

1. The etiology of multiple sclerosis is believed to be:
   - **a.** a viral infection of the central nervous system
   - **b.** an autoimmune disorder mediated by myelin-reactive lymphocytes
   - **c.** a primary neurodegenerative disorder
   - **d.** a toxic encephalopathy

2. The mechanism of action of disease modifying therapies currently used to reduce MS relapse rates is:
   - **a.** Blockade of lymphocyte trafficking to the central nervous system
   - **b.** Depletion of lymphocytes
   - **c.** Inhibition of T cell proliferation
   - **d.** Inhibition of pyrimidine synthesis
   - **e.** All of the above

3. To make a diagnosis of relapsing remitting MS it is critical to demonstrate:
   - **a.** Dissemination of lesions in time
   - **b.** Dissemination of lesions in space
   - **c.** Immunoglobulin synthesis in the central nervous system
   - **d.** Worsening of symptoms with increases in core body temperature
   - **e.** a and b

67

Autoimmune Peripheral Neuropathies

Marinos C. Dalakas

Autoimmune peripheral neuropathies (APNs) occur when immunological tolerance to peripheral nerve components (myelin, Schwann cell, axon, and motor or ganglionic neurons) is lost. In some of these neuropathies, there is direct evidence for autoimmune reactivity mediated by specific antibodies or autoreactive T lymphocytes against peripheral nerve. In others, the underlying immune-mediated mechanism is secondary or indirect, and an autoimmune cause is suspected when the neuropathy coexists with another systemic autoimmune disease or viral infection.

This chapter reviews the most common autoimmune neuropathies (Table 67.1), their clinical features and diagnostic criteria, the prevailing autoimmune phenomena that govern each neuropathy, and the most effective therapeutic approach.

ACUTE INFLAMMATORY POLYNEUROPATHY: GUILLAIN-BARRÉ SYNDROME(S)

Guillain-Barré syndrome (GBS) is an acute demyelinating polyneuropathy characterized by acute (within 1 week) or subacute (within 4 weeks) ascending motor weakness, mild or moderate sensory abnormalities, occasional cranial nerve involvement, and muscle or radicular pain.[1-5] Tendon reflexes are reduced but can be normal, especially in axonal GBS. It is a disease of all ages, with an incidence of 0.8–1.9 (median 1.1) per 100 000,[4] and occurs sporadically, although occasional outbreaks have been noted. In typical cases, the maximum deficit is reached by the fourth week, a sign conventionally used to separate GBS from chronic inflammatory demyelinating polyneuropathy (CIDP), in which the disease begins slowly and usually reaches a nadir after at least 2 months. There are, however, patients with CIDP (perhaps up to 16%) with subacute onset and monophasic course and fall between the two time-frames and still others with an even more acute onset, reaching a nadir within 6–8 weeks and resembling GBS.[1-4] Distinguishing GBS from acute-onset CIDP is challenging because the distinction, in most cases, becomes evident in retrospect, although proposed criteria may help separate the two early in the disease course.[1-5]

After reaching the clinical peak, there is a recovery period that varies from weeks to years according to disease subtype or severity at onset; rapid initial progression over <7 days, severity at onset requiring mechanical ventilation, age >60 years, and preceding diarrheal illness are signs associated with incomplete recovery or worse outcome. GBS is not one syndrome, but several syndromes, reflecting the varying degree of involvement of the motor or sensory nerve fibers and the myelin sheath or the axon. The GBS subtypes, or GBS variants,[1-5] include:

- *Acute inflammatory demyelinating polyneuropathy* (*AIDP*), which accounts for the majority (probably 80%) of patients. In classic cases, the weakness starts from the legs and spreads up to the arms, intercostal and diaphragmatic muscles, and facial or bulbar muscles, causing dysphagia and dysarthria. At times the weakness may be limited to one or two limbs or to cranial nerves. Patients need to be monitored for impending respiratory failure, hence the need for early admission to intensive care units (ICUs). Autonomic dysfunction occurs in varying degrees in up to 65% of patients; if severe, it can be life-threatening as a result of cardiac arrhythmias or hemodynamic changes, such as hypertension, orthostatic hypotension, and reduced peripheral vascular tone.[1-5]
- *Acute motor axonal neuropathy* (*AMAN*), which exhibits primary axonal damage caused by massive acute demyelination and inflammation, as occurs in experimental allergic neuritis (EAN) when the animals are immunized with a high dose of myelin antigen,[1-5] or by a primary axonal event mediated by macrophages. These patients have a fulminant course with severe paralysis and complete electrical inexcitability of motor nerves as early as 3–5 days after onset.[1-5] In contrast to AIDP, involvement of cranial nerves is infrequent, and reflexes are normal or increased, especially early in the disease. AMAN is common in Asia and Central South America, accounting for 30–65% of all GBS cases in these regions.[5] Recovery of motor function is variable; some patients recover within days as a result of resolution of conduction block, but others have slow and poor recovery because of excessive axonal degeneration at the root level.[1-5] Infection with *Campylobacter jejuni* appears to be responsible for many of these cases.[1-7] A number of patients also have high levels of antiganglioside GM1 antibodies.[1-7]
- *Acute motor–sensory axonal neuropathy* (*AMSAN*) is similar to AMAN, but with concurrent involvement of the sensory axons, and has a pathomechanism similar to that of AMAN, including frequent antibodies against GM1 and GD1a gangliosides following *C. jejuni* infection.
- *Miller Fisher syndrome* (MFS) is characterized by acute onset of ophthalmoplegia, gait ataxia, normal sensation, and areflexia.[1-7] In a third of patients, there is muscle weakness with pharyngeal, facial, trunk, and respiratory muscle involvement; rarely it can present as an isolated ocular nerve palsy.[1-7] MFS is a variant that is distinct because of the presence of a unique immunoglobulin G (IgG) antibody against GQ1b ganglioside.[1-7]

TABLE 67.1 Common Autoimmune Neuropathies

- Guillain-Barré syndrome(s) (GBS)
- Chronic inflammatory demyelinating polyneuropathy (CIDP) and its variants
- Polyneuropathy associated with paraproteinemias
- Immunoglobulin M (IgM) monoclonal gammopathies
- IgG and IgA monoclonal gammopathy
- Polyneuropathy, organomegaly, endocrinopathy, myeloma, and skin changes (POEMS) syndrome
- Cryoglobulinemic polyneuropathy
- Multifocal motor neuropathy with conduction block
- Paraneoplastic neuropathies associated with anti-Hu antibodies
- Autoimmune autonomic neuropathies
- Vasculitic neuropathies
- Infectious neuropathies (human immunodeficiency virus [HIV], cytomegalovirus [CMV], Epstein-Barr virus [EBV], and herpes virus infections; Lyme disease; leprosy; Chagas disease; diphtheria; others)

- *Sensory ataxic GBS* results from the involvement of roots and ganglionic neurons. Some of these patients have antibodies to GD1b ganglioside, probably forming a continuum with MFS because they share autoantibodies with the same sialic groups.[1-7]
- *Acute pandysautonomic neuropathy,* where the target antigen is seemingly in the ganglionic neurons, although never identified. Autonomic symptoms, however, can coexist in all forms of GBS during the acute, stable, or even the recovery phase.[1-5]

Diagnosis

The diagnosis is often suspected on clinical grounds, but it is confirmed with elevated levels of cerebrospinal fluid (CSF) protein and abnormal results of electrophysiological studies consistent with active demyelination or nerve inexcitability.

CSF protein may be normal in the early phase of the disease, but it can be as high as 1000 mg/dL by the sixth week. The elevation of CSF protein levels may result from the involvement of the roots related to inflammation, but, as the blood–nerve barrier becomes impaired, serum albumin and IgG may enter freely into CSF, contributing further to protein elevation. CSF cell count is normal (or slightly increased <50 cells per microliter); there is, however, lymphocytosis when GBS occurs in conjunction with viral infections, such as human immunodeficiency virus (HIV), cytomegalovirus (CMV), Epstein-Barr virus (EBV), or Lyme disease. When CSF protein is very high, papilledema can develop because of impaired reabsorption of CSF and raised intracranial pressure. Oligoclonal IgG bands can be also seen. Results of nerve conduction studies can be normal early in the disease; they are, however, often helpful to distinguish AIDP from AMAN or AMSAN, although sometimes serial studies are needed in the ensuing weeks. Nerve conduction studies may also have prognostic value because features of demyelination suggest higher chances of needing mechanical ventilation, whereas the presence of low compound muscle axon potentials from the outset (indicative of axonal loss) predicts poor outcome.[4,5] Differential diagnosis of GBS should include other forms of acute flaccid paralysis, such as brainstem stroke; brainstem encephalitis; acute motor neuron involvement caused by poliomyelitis or West Nile virus infection; acute myelopathy; disorders of neuromuscular transmissions, such as myasthenia gravis or botulism; disorders of muscles, such as hypokalemia, inflammatory myopathy (especially the necrotizing autoimmune subtype), acute rhabdomyolysis, or periodic paralysis; and other rare causes of acute neuropathy, such as porphyria, toxins, vasculitis, or critical illness polyneuropathy.

Antecedent Illnesses or Events

Two-thirds of patients with GBS give a history of a flu-like illness or acute dysenteric episodes that precede the development of GBS by 1–3 weeks.[1-5] Among the implicated viruses are CMV, EBV, herpesvirus, hepatitis A, HIV, and now Zika virus. Among bacteria, infection with *Mycoplasma pneumoniae* and, most importantly, *C. jejuni* may be present in >25% of the patients and in some parts of the world up to 50%.[1-5] *Campylobacter* is of special interest because it contains glycoconjugates that share epitopes with the peripheral myelin, as discussed later. Two vaccines—one against rabies and the other against the swine flu A/New Jersey influenza strain that caused an outbreak of GBS in 1976[1-5]—have been convincingly associated with development of GBS. Administration of rabies vaccine that contains brain material is followed by GBS in about 1 in 1000 cases. Apart from these vaccines, however, despite anecdotal reports, there is no convincing evidence that the incidence of GBS is increased in association with other vaccines. Presently there is concern that Zika virus, an arbovirus in the family of Flaviviridae, is emerging in several countries and territories of South America as a cause of microcephaly and GBS. Although the information is still evolving, compelling data from French Polynesia provide evidence that Zika is associated with GBS in the form of AMAN with rapid disease evolution—within 4–6 days—and antiglycolipid antibodies in 31% of the patients.[8] On the basis of the total number of 42 reported cases from October 2013 to March 3014, the risk of GBS was estimated to be 0.24 per 1000 Zika virus infections.[8] Although this is clearly alarming, considering that Zika is a mosquito-borne virus, more data are needed. Surgery can precede the development of GBS in some patients[1-4]; surgical stress, the release of nerve autoantigens, and infections have been proposed as possible explanations for this association. Three drugs—gold, perhexiline, and suramin at high doses—have been causally associated with acute demyelinating neuropathy. GBS has occurred in patients who suffer from neoplasms, especially lymphoma, melanoma, and Hodgkin disease.[1-4] Interestingly, GBS is rarely seen as part of another connective tissue disorder.

Immunopathology of Guillain-Barré Syndrome

GBS is an inflammatory demyelinating polyneuropathy in which the peripheral myelin, the axon, the node of Ranvier, or the Schwann cell are the putative target antigens of an immune attack, possibly triggered by various antecedent events. Both the cellular and the humoral components of the immune system have been implicated.[1-5]

Cellular Factors

Two histopathological features are prominent in typical GBS: perivascular and endoneurial inflammatory infiltrates throughout the nerves, roots, or plexuses[1-4] and segmental demyelination in areas associated with the lymphoid infiltrates, especially macrophages. Macrophages, which are the most prominent cells in contact with nerve fibers, break through the basement membrane of healthy Schwann cells and make direct contact with the outermost myelin lamellae, leading finally to lysis of the superficial myelin sheath (macrophage-mediated demyelination).

Cytokines and chemokines released by the activated T cells or complement activation may increase capillary permeability and facilitate transmigration of additional macrophages or T cells. When the demyelination is extensive or chronic, it is followed by axonal degeneration.[1-5] The degree and effectiveness of remyelination and axonal regeneration dictate the degree of clinical recovery.

The role for a T cell–mediated process in GBS is mostly derived by analogy to the animal model of EAN, which resembles GBS in both its pathology and its clinical course.[1-5] Animals sensitized to whole human nerve or to various myelin proteins, such as P0, P2, and the neutral glycolipid galactocerebroside, develop segmental demyelination with mononuclear cell infiltrates consisting of macrophages and T cells. In EAN, T cells are sensitized against myelin and can passively transfer the disease to healthy animals. Increased levels of interleukin-2 (IL-2) and soluble IL-2 receptors are noted in serum during the acute phase of GBS, suggesting ongoing T-cell activation. Further, lymphocytes from patients with GBS exert myelinotoxic activity when applied to cultures of myelinated axons.

Humoral Factors and Antiganglioside Antibodies

There is much stronger evidence that circulating serum factors are responsible for GBS. On clinical grounds, this is supported by the beneficial effect of plasmapheresis, presumably by removing putative antibodies. On laboratory grounds, it is supported by the variety of autoantibodies detected in patients' sera. Serum from the acute phase of GBS can demyelinate rodent dorsal root ganglionic extracts in a complement-dependent manner. Furthermore, GBS serum injected into rat sciatic nerves causes demyelination and conduction block. Complement-fixing IgM antibodies against a human peripheral nerve myelin glycolipid that contains carbohydrate epitopes as well as high-titer antibodies against various sulfated or acidic glycosphingolipids are present in several patients with GBS.[1-7]

Gangliosides are present in all tissues but are especially abundant in the nervous system. Their lipid portion lies in the cell membrane, and their signature sugar residues are exposed at the extracellular surface bearing one or more sialic acid molecules, such as one sialic acid ganglioside (GM1), two (GD1a), three (GT1a), or four (GQ1b).[1-7] Although they do not form a common "GBS antigen," different gangliosides are involved in different GBS subtypes. They are of pathogenic relevance because immunization of rabbits with GM1 and GD1b induces acute neuropathy with histological features of AMAN.[1-7] Their pathogenicity was also confirmed by an inadvertent experiment in humans who had received ganglioside injections for various maladies and developed AMAN accompanied by anti-GM1 antibodies.[1-7] Additionally, antibodies to GQ1b or GD1a cause conduction block at the motor nerve terminals in a preparation of mouse phrenic nerve.[1-7] Similarly, antibodies to GalNAc-Gd1a from a patient with AMAN blocked neuromuscular transmission in a mouse spinal cord muscle coculture system.

IgG antibodies that react with GM1, GD1a, GalNAc-GD1a, and GM1b are found in 80% of cases with the motor axonal form of GBS (AMAN and AMSAN), but in the most common GBS subtype, AIDP, ganglioside-specific antibodies are uncommon. Among gangliosides, the one that clearly correlates with a specific clinical syndrome is the GQ1b, which is specifically associated with the MFS variant with IgG anti-GQ1b antibodies present in >90% of these patients [1-7] In contrast, anti-GQ1b antibodies of the IgM class can be found in patients with chronic IgM paraproteinemic polyneuropathies,[6] as discussed later. Anti-GQ1b IgG antibodies are also found in postinfectious ophthalmoplegias as well as in GBS with ophthalmoplegia, but not in those without ophthalmoplegia or other autoimmune conditions.[1-7] Of interest, anti-GQ1b antibody binds the paranodal regions of oculomotor nerves III, IV, and VI, suggesting that damage to these regions blocks impulse propagation at the nodes of Ranvier, resulting in a conduction block that is characteristic for GBS. Many patients with antibodies to GQ1b also have antibodies to GD1a.

The reasons for different clinical syndromes in connection with specific gangliosides remains unclear, but distribution, accessibility, and density or configuration of gangliosides at different sites may be critical factors. For example, there is more GM1 in ventral roots than in dorsal roots, hence the predominantly motor neuropathy seen with anti-GM1 antibodies; there is also more GQ1b in the ocular motor nerves, which may explain the involvement of eyes in MFS. How these antibodies induce disease, however, remains unclear. Current evidence suggests that antibodies against different acidic glycolipids or sulfatides may be related to different viral or bacterial antecedent factors, as discussed below.

Molecular Mimicry: Relationship Between *Campylobacter jejuni* and Gangliosides in Acute Motor Axonal Neuropathy

Antecedent infection with *C. jejuni* has been commonly associated with AMAN. The strain of *C. jejuni* associated with AMAN (Penner D: 19 serogroup) is, however, different from those causing common enteritis and is more likely to have the genes for enzymes that synthesize sialic acid in the bacterial wall mimicking ganglioside GM1, GD1a, or GQ1b.[1-5] These patients have a higher incidence of anti-GM1 antibodies, which suggests cross-reactivity between epitopes in the lipooligosaccharide in the bacterial wall and the ganglioside.[1-5] Further, injection of lipooligosaccharides extracted from *C. jejuni* into rabbits has been shown to induce acute neuropathy, with development of anti-GM1 antibodies identical to those found in AMAN.[1-7] In addition, immunization of mice with these lipooligosaccharides generates a monoclonal antibody (mAb) that reacts with GM1 and binds to human peripheral nerve. This GM1 mAb as well as the anti-GM1 IgG extracted from patients with GBS block muscle action potentials in muscle–spinal cord coculture. Carbohydrate mimicry between the bacterial lipooligosaccharide and human GM1 is therefore an important cause of AMAN. Because *C. jejuni* is a common cause of a diarrheal illness worldwide, and diarrhea has been an antecedent event in up to 50% of patients with GBS, *Campylobacter* appears to trigger the disease in many patients, especially in certain parts of the world. Isolation of *Campylobacter* from stools early in acute GBS varies from 44% to 88% of patients, and IgG or IgM *Campylobacter*-specific antibody titers are seen in a higher percentage (36%) of patients with GBS than in controls (10%).

Molecular mimicry may not be limited to *C. jejuni* because GM1 and GQ1b epitopes are also found in the bacteria wall of *Hemophilus influenzae,* which is also a triggering factor in GBS. GBS triggered by CMV infection has been also associated with the presence of IgM anti-GM2 antibodies. Another potential factor for molecular mimicry is *M. pneumoniae,* which precedes GBS in 5% of cases and is known to stimulate antibodies against human carbohydrate antigens, including galactocerebroside, the main glycolipid antigen in peripheral nerves.[1-7]

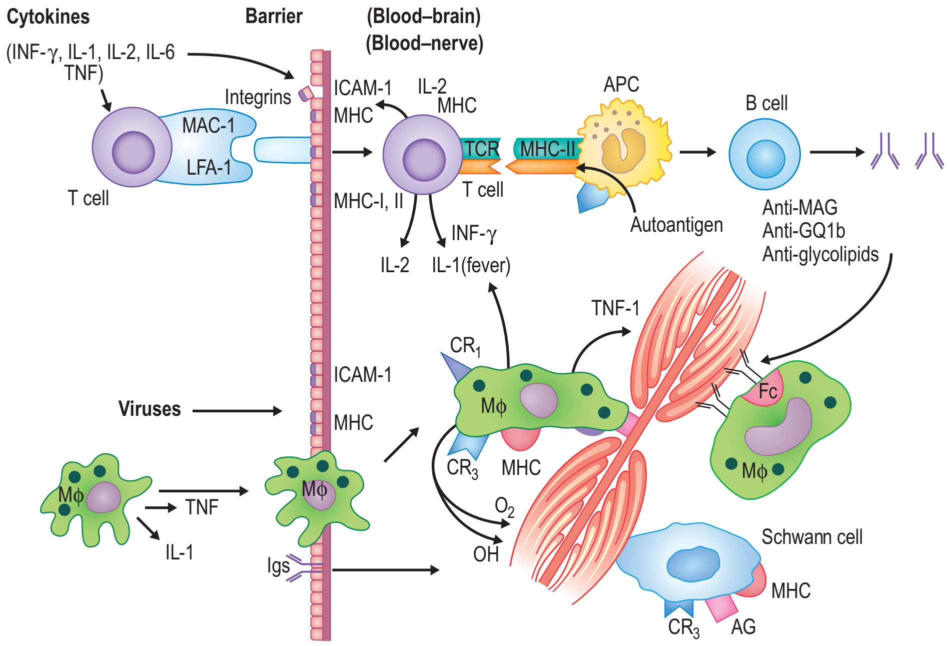


FIG 67.1 Sequence of events in the mechanisms of immune-mediated demyelinating polyneuropathy. Cytokines lead to increased expression of major histocompatibility complex (MHC) class I and intercellular adhesion molecules, allowing the sensitized T cells and macrophages to exit the endothelial cell wall and traffic to the peripheral nerve. There they recognize myelin antigen and induce a macrophage-mediated demyelination. The antigen-presenting cells (APCs; probably Schwann cells or macrophages), in concert with MHC class II expression, interact with CD4 T cells and lead to clonal expansion of B cells, producing antibodies against various peripheral nerve antigens.

Molecular mimicry may also play a role in the Zika virus–associated GBS as antiglycolipid antibodies were found in 31% of these patients.[8]

Molecular mimicry between epitopes of viral proteins (which trigger the disease) and myelin components may result in sensitization of cross-reactive T cells that may stimulate B cells to produce specific antibodies directed against myelin components or may recruit macrophages as effector cells. A combination of cellular and humoral factors therefore seems to participate in the cause of the disease.[1-8] Circulating cytokines triggered by the initiating event (viruses or bacteria) could also upregulate intercellular adhesion molecule (ICAM)-1 expression on the endothelial cells and facilitate the entrance of activated T cells or antibodies to the endoneurial parenchyma. It is relevant that ICAM-1 is increased in patients with GBS.[1-5] A scheme summarizing the immunopathogenic mechanism is shown in Fig. 67.1.

Nodal and Paranodal Alterations and Specific Antinodal Antibodies

Antibodies to nodal and paranodal antigens seen in CIDP subsets, as discussed later, have been also detected in a small number of patients with GBS in Europe. Antibodies against neurofascin, ankyrin G, and contactin could potentially explain the acute conduction block or the paranodal axonal degeneration and the rapid reversibility or slow and incomplete recovery seen in several patients with AMAN [1-5,9,10]

KEY CONCEPTS

Autoimmunity in Guillain-Barré Syndrome (GBS)

Cellular Factors

- The peripheral myelin or the Schwann cells are targets.
- Activated macrophages are the dominant endoneurial cells and lift the outermost myelin lamellae, lysing the superficial myelin sheaths.
- Peripheral blood lymphocytes exert myelinotoxic activity *in vitro.*
- Levels of interleukin-2 (IL-2) and soluble IL-2 receptors are increased during the acute phase of the disease and decline during recovery.

Humoral Factors

- Serum exerts a complement-dependent demyelination *in vitro.*
- Intraneural injections of serum from patients with acute GBS cause demyelination and conduction block.
- Immunoglobulin G (IgG), IgM, and membranolytic attack complex are detected immunocytochemically on the patients' nerves.
- High titers of IgG antibodies against peripheral nerve acidic glycolipids (GM1, GQ1b) are detected in the sera of patients with acute motor axonal neuropathy (AMAN) and Miller Fisher syndrome (MFS). The GQ1b ganglioside is a specific antigen for MFS.
- There is a high incidence of antibodies to *Campylobacter jejuni*, and GM1, with molecular mimicry between *Campylobacter* and nerve gangliosides.
- Injection of lipooligosaccharides extracted from *C. jejuni* cause AMAN and elicit GM1 antibodies in rabbits.
- Antiganglioside antibodies extracted from patients with GBS block muscle action potentials *in vitro.*

CHRONIC INFLAMMATORY DEMYELINATING POLYNEUROPATHY

CIDP is the most common form of chronic APN with prevalence as high as 9/100 000.[2,3,11,12] It is also the most gratifying chronic autoimmune neuropathy because it is treatable in the majority of the cases. It can be considered the chronic counterpart of GBS because of the various clinical, electrophysiological, histological, and laboratory similarities. CIDP differs from GBS predominantly by its tempo, mode of evolution, prognosis, and responsiveness to steroids. First described as a “steroid-responding relapsing polyneuropathy,” CIDP shares with GBS a variety of common autoimmune features.

Clinical Features and Disease Variants

The typical CIDP is characterized by a progressive, symmetrical, proximal and distal muscle weakness; paresthesias; sensory dysfunction; and impaired balance that evolve slowly over at least 2 months.[2,3,11,12] Tendon reflexes are absent or reduced. Cranial nerves rarely may be affected. The course is often monophasic, with stepwise progression; at times the disease has a relapsing course, with spontaneous remissions necessitating the need to periodically evaluate the usefulness of continuing immunotherapeutic interventions. Because the demyelination is multifocal, affecting spinal roots, plexuses, and proximal nerve trunks,[3,13] the clinicopathological picture may be variable, accounting for the different manifestation of symptoms and signs.[2,3,11,12] The most notable CIDP variants include the asymmetrical, unifocal or multifocal, motor–sensory form (Lewis-Sumner syndrome); the pure motor form; the pure sensory form; the sensory ataxic form; and the pure distal form.

Diagnosis

In CIDP, CSF protein is elevated, up to sixfold, without pleocytosis (except if an infection coexists). Nerve biopsy shows demyelination and remyelination, occasional epineurial or endoneurial T cells, and several macrophages that are either scattered or in small perivascular clusters in the endoneurium (Fig. 67.2).[2,3,11,12] Electrophysiological testing is fundamental for the diagnosis by demonstrating the following typical features of demyelination in motor and sensory fibers: *(i)* slow conduction velocity; *(ii)* prolonged distal motor or sensory latencies; *(iii)* prolonged F

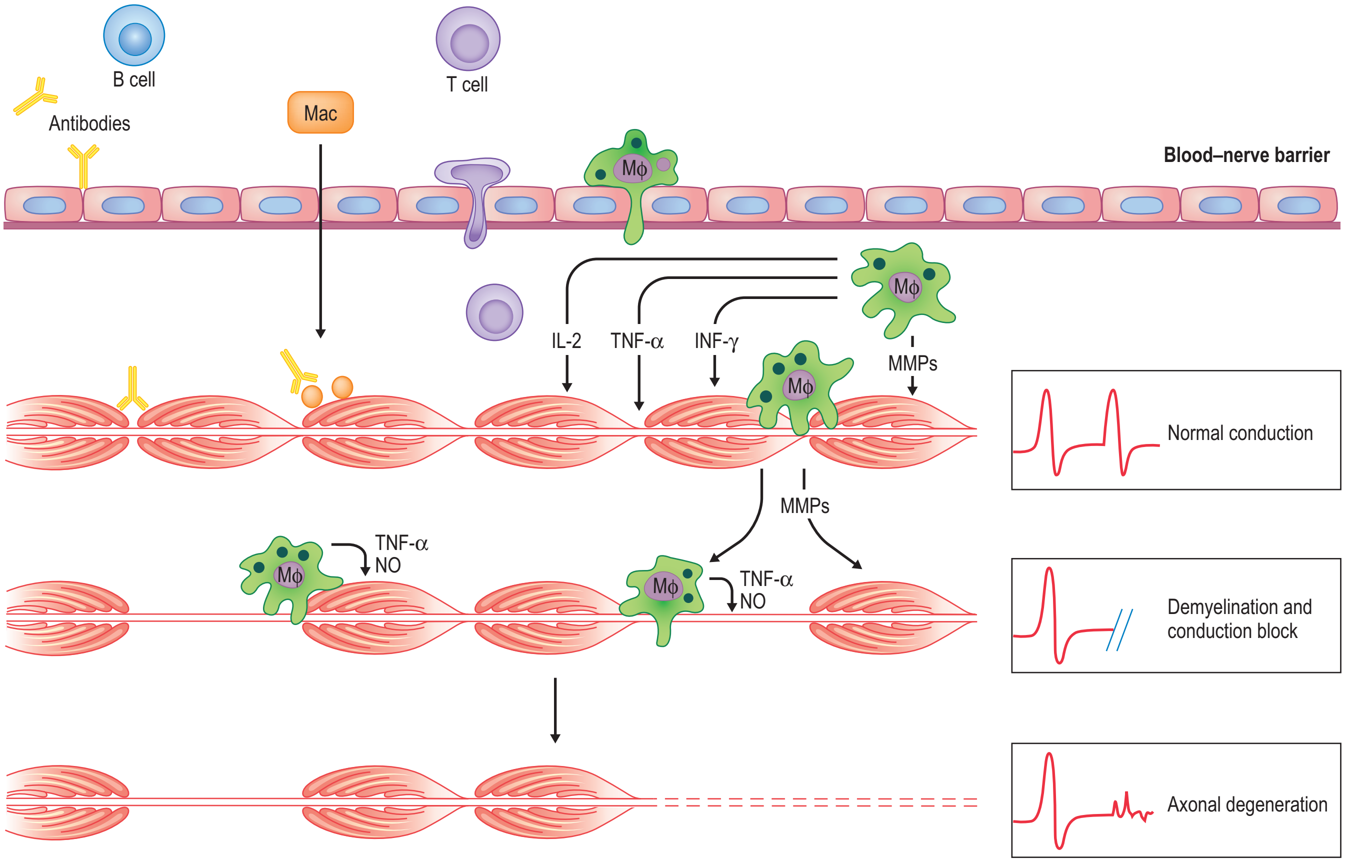


FIG 67.2 Diagrammatic scheme of the main cellular and humoral factors implicated in the demyelinating process of chronic inflammatory demyelinating polyneuropathy (CIDP) leading to axonal loss. Activated macrophages (Mϕ) and T cells cross the endothelial cell wall of the blood–nerve barrier and reach the myelinated fibers. Activated, TNF-α-positive Mϕ invade the myelin sheath, causing Mϕ-mediated segmental demyelination. Axonal loss secondary to demyelination, probably enhanced by TNF-α and metalloproteinases, may become prominent in the chronic phases of the disease. Other cytokines, T cells sensitized to unidentified antigens, and putative antibodies may participate. IL, interleukin; IFN, interferon; MMP, metalloproteinase; Mϕ, activated macrophage; TNF, tumor necrosis factor.

waves latencies; and *(iv)* conduction block with dispersion of the compound muscle action potentials. An associated axonal loss is not unusual in the majority of CIDP cases. A variety of diagnostic criteria have been proposed to capture the most pertinent of the aforementioned features; the revised European Federation of Neurological Societies/Peripheral Nerve Society (EFNS/PNS) guidelines seem the most appropriate because they offer 81% sensitivity and 96% specificity to capture patients more likely to respond to therapies.[2,3,11,12] Routine CSF testing and nerve biopsy are not mandatory for the diagnosis,[2,3,11,12] but they can be helpful when the results of electrophysiological testing are not convincing, or there is a need to exclude hereditary and vasculitic neuropathies. Other diseases causing neuropathy that should be excluded include severe diabetes (although CIDP seems more frequent in patients with diabetes), neoplasms, amyloidosis, IgM paraproteinemia (IgG or IgA monoclonal gammopathy of undetermined significance (MGUS) can be seen in CIDP), myelomas, vasculitis, alcoholism, exposures to neurotoxic drugs, or family history of neuropathy.

KEY CONCEPTS

Autoimmunity in Chronic Inflammatory Demyelinating Polyneuropathy (CIDP)

- Activated macrophages are the predominant endoneurial cell, displacing the Schwann cell cytoplasm, disrupting myelin, and lysing superficial myelin lamellae.
- Complement-fixing immunoglobulin G (IgG) and IgM antibodies are deposited on the myelin sheath.
- IgG antibodies to acidic glycolipids LM1, GM1, or GD1b and against the 28-kDa P0 myelin proteins are detected in the sera of some patients.
- There is upregulation of DR and B-7 costimulatory molecules in Schwann cells and macrophages.
- Serum IgG can induce conduction block when injected into rat nerves.
- Up to 25% of patients with CIDP harbor specific antibodies against antigens in the nodes of Ranvier; in 10% of them these antibodies have been identified as directed against neurofascin-155 and CASPR causing conduction block.

Immunopathogenesis

Because of immunopathological similarities to GBS and the relapsing–remitting EAN, in CIDP, activated T cells, macrophages, complement, and autoantibodies seem to work in concert with each other to induce an immune attack against peripheral nerve antigens[2,3,11,12] (as depicted in Fig. 67.1), but no triggering factors have been identified so far.

The predominant endoneurial mononuclear cells in CIDP are the macrophages that constitute the final effector cells associated with demyelination because they express activation markers probably induced by cytokines released by autoreactive T cells *in situ* or the circulation. These cells sequentially penetrate the basement membrane of the Schwann cell, displace the cytoplasm, and split the myelin lamellae, resulting in focal lysis of the myelin sheath.[2,3,11,12] Macrophages as well as Schwann cells may serve as antigen-presenting cells (APCs) because they express human leukocyte antigen (HLA)-DR and costimulatory molecules B7-1 (CD80) and B7-2 (CD86), whereas their counterreceptors cytotoxic T lymphocyte antigen-4 (CTLA-4) and CD28 are expressed on rare endoneurial CD4 T cells.[2,3,14] B7-2–deficient mice also develop CIDP.[15] Elevated soluble adhesion molecules, chemokines, cytokines, and metalloproteinases are detected in serum and CSF in CIDP, probably facilitating lymphoid-cell transmigration across the blood–nerve barrier. Although T cells are not prominent overall, the few CD8 and CD4 cells found endoneurially have monoclonal or oligoclonal restrictions in their T-cell receptor (TCR) repertoire, implying an antigen-driven T-cell response.[2,3,12,13]

Humoral factors seem to play a predominant role, especially with the recent identification of target antigens at the nodes of Ranvier in >10% of patients with CIDP, even though in the majority of the patients, the pathogenic antigens remain elusive. The beneficial effect of plasmapheresis provides indirect evidence that some circulatory factors are pathogenic. The concept that antibodies may be implicated dates back to >30 years when complement-fixing IgG and IgM were found deposited on a patient's myelin sheath.[16] Antibodies to glycolipids LM1, GM1, or GD1b also have been seen in some patients, but less frequently than in GBS and more frequently than in controls.[2,3,11,12] Overwhelming evidence accumulated in the last 5 years indicates that molecules associated with saltatory conduction at the nodes of Ranvier may be more meaningful targets[3,4,10,17-19] because functional blockade in these regions can best account for the rapid improvement noticeable within days after plasmapheresis or intravenous immunoglobulin (IVIG). Likely antigenic targets include neurofascin-186, moesin, and gliomedin (at the node); neurofascin-155 (NF155), contacting/Caspr 1 (CNTN1), and connexins (at the paranode); and transient–axonal–glycoprotein-1 or potassium channels (at the juxtaparanode).[3,4,10,17-19] Evidence from proteomics, transfected cell lines, teased-nerve fibers, immunocytochemistry, and enzyme-linked immunosorbent assay (ELISA) has shown that IgG4 antibodies against paranodal NF155 and CNTN1 antigens are the most relevant and seemingly pathogenic antibodies that are currently detected in at least 10% of patients.[17-19] Of interest, anti-NF155– and CNTN1-positive patients appear to have a distinct clinical phenotype with more severe disease, axonal involvement, tremors, sensory ataxia, and suboptimal response to IVIG[17-19]; these patients, however, respond better to anti–B cell therapies, such as rituximab.

Overall, the immunopathogenetic scheme proposed for GBS (see Fig. 67.1) is also appropriately applied to CIDP. Molecular mimicry can be implicated in rare cases of CIDP associated with melanoma because the carbohydrate myelin epitopes GM2, GM3, and GD3 are also expressed on melanoma cells, and antibodies against melanoma cells react with myelin glycoproteins.[20] The axonal involvement that accompanies demyelination in both CIDP and GBS, either acutely or after long-standing demyelination, is depicted in Fig. 67.2. The relevant nodal target antigens are presented in Fig. 67.3 separately[2,3,10] to highlight their recognized importance in the field of antibody-mediated peripheral demyelination.

MULTIFOCAL MOTOR NEUROPATHY WITH CONDUCTION BLOCK

MMN is a distinct disease that, although rare with prevalence of 0.6/100 000, should be recognized early because it is treatable. It affects males more than females and is more common in those <50 years of age. MMN presents with progressive weakness, atrophy, and areflexia that often begins in the hands and is prominent in distal muscle groups supplied by many individual peripheral nerves (multifocal).[2,3,21] It differs from vasculitic neuropathy because it is slow and is painless because it affects only the motor nerve fibers. It also differs from the motor variant of CIDP because it is multifocal and asymmetrical. Because some

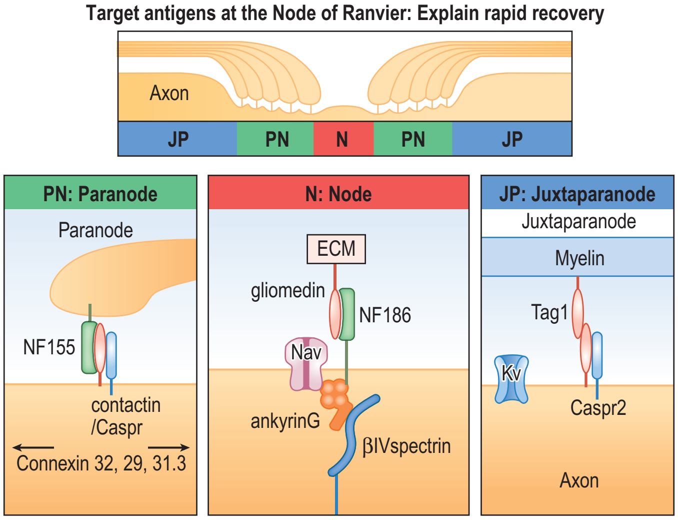


FIG 67.3 Main proteins in the nodal (N), paranodal (PN), and juxtaparanodal (JP) regions in the nodes of Ranvier implicated as antigens in acquired demyelinating neuropathies (Guillain-Barré syndrome [GBS] and predominantly chronic inflammatory demyelinating polyneuropathy [CIDP]). The two main antigenic targets in CIDP are neurofascin-155 and CASPR2 located in the paranodal regions; antibodies against these proteins are detected in the serum of 10% of patients with CIDP, resulting in conduction block and paranodal axonal changes that probably account for resistance to IVIG.[4,13]

patients with MMN can experience cramps and fasciculations, the disease is often erroneously diagnosed as lower motor neuron disease, and treatment is delayed. In contrast to lower motor neuron disease, however, MMN progresses very slowly, the weakness is within distributions of peripheral nerves and not multisegmental, and the cranial musculature is often spared. CSF protein level is normal, an important finding that distinguishes MMN from the motor variant of CIDP.

MMN has a distinct electrophysiological criterion, namely, multifocal conduction block in motor nerves. In contrast to CIDP, in which sensory conduction block is also present, in MMN the sensory conduction remains normal across the nerve segments that have motor block. The reason for a selective motor involvement is unclear. Differences in the antigenic specificities in the myelin components between motor and sensory fibers are suspected because the ceramide composition of gangliosides differs between sensory and motor fibers. Although the immunopathology of the disease remains unclear, patients with MMN respond remarkably well to immunotherapy, as discussed later. Up to 50% of the patients have high antibody titers to GM1 ganglioside, but the role of these antibodies is uncertain. IgM GM1 antibodies can be also seen in other autoimmune neuropathies and in even up to 25% of patients with amyotrophic lateral sclerosis (ALS) who do not have conduction block.

POLYNEUROPATHIES ASSOCIATED WITH MONOCLONAL GAMMOPATHIES OF UNDETERMINED SIGNIFICANCE

A distinct subset of acquired polyneuropathies has been associated with a circulating monoclonal protein (Chapter 80). Although neuropathy occurs in a setting of myeloma, plasmacytoma, or Waldenström macroglobulinemia, the majority of patients with paraproteinemic neuropathies do not have a lymphoproliferative disease, and the monoclonal gammopathy is of undetermined significance (MGUS). A benign monoclonal gammopathy may occur in up to 1% of normal people >50 years of age.[22] The incidence increases to 1.7% above age 70 years and reaches up to 6% above age 90 years. Monoclonal gammopathies, however, are 10 times more frequent in patients with polyneuropathy than in an age-matched control population, and almost 10% of patients with acquired polyneuropathy have MGUS.[23] If these gammopathies are categorized into subclasses, the incidence of polyneuropathy among patients with IgM monoclonal proteins can be as high as 50%,[23,24] implying that almost 50% of patients with IgM MGUS may have or will develop polyneuropathy. At present, polyneuropathies with MGUS comprise 10% of patients with acquired neuropathy,[23-25] and paraproteinemic polyneuropathy is a potentially treatable APN.[25]

Patients with demyelinating polyneuropathy associated with IgG or IgA MGUS are indistinguishable from those with CIDP; the paraprotein in such patients is coincidental and causally unrelated to neuropathy. In contrast, the demyelinating polyneuropathy associated with IgM MGUS is a distinct clinicopathologic entity, and the IgM is considered pathogenic because it is often directed against myelin glycoproteins or glycolipids.[2,3,6,24,25]

Some patients with paraprotein may have an associated amyloidosis derived from the variable region of the Ig light chain, primarily λ. When amyloidosis is present, the neuropathy is painful, and the sensorimotor deficits are accompanied by autonomic symptoms consisting of orthostatic hypotension, impotence, impaired gastric motility, or frequent diarrhea. Amyloid neuropathy is difficult to treat; apart from symptomatic therapy (mostly opiates), immunosuppressive therapies, especially with the various anti–B-cell agents, and bone marrow transplantation are largely applied.

ANTIBODIES TO MYELIN-ASSOCIATED GLYCOPROTEIN IN PATIENTS WITH IgM M MONOCLONAL GAMMOPATHIES OF UNDETERMINED SIGNIFICANCE POLYNEUROPATHY (ANTI-MAG NEUROPATHY)

Most of these patients present with a sensory, large-fiber, demyelinating polyneuropathy that manifests as sensory ataxia.[2,3,24,26] Other patients have a sensorimotor polyneuropathy with mixed features of demyelination and axonal loss. CSF protein is often elevated. Nerve conduction studies have demonstrated slow conduction velocity and a rather characteristic prolonged distal motor and sensory latency indicative of distal demyelination. Sural nerve biopsy demonstrates a diminished number of myelinated axons. On electron microscopy, there is splitting of the outer myelin lamellae, linked to the presence of IgM deposits in the same area of the split myelin sheath.

Sera from approximately 50% of these patients react with myelin-associated glycoprotein (MAG), a 100-kilodalton (kDa) glycoprotein of the central and peripheral nerve myelin, as well as other glycoproteins or glycolipids that share antigenic determinants with MAG.[2,3,24,27,28] The antigenic determinant for the anti-MAG IgM resides in the carbohydrate component of the MAG molecule.[3] The anti-MAG IgM paraproteins coreact with

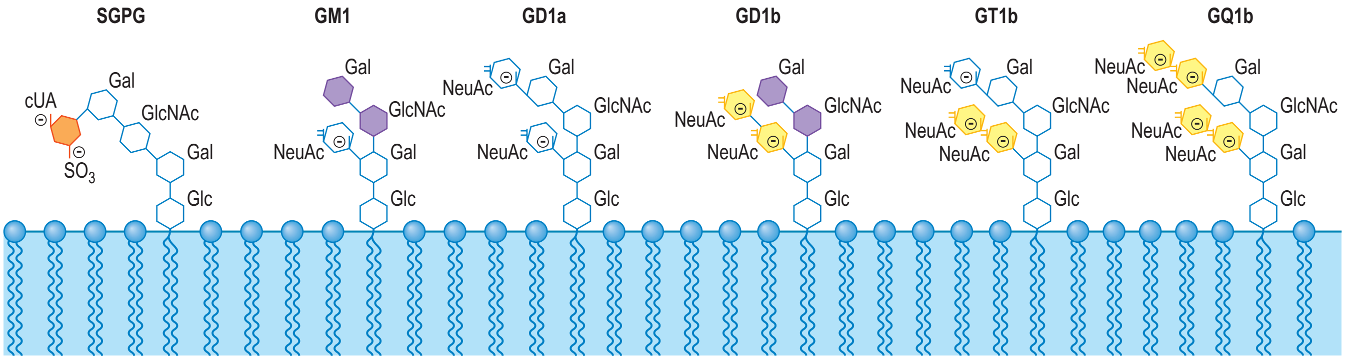


FIG 67.4 Glycolipids implicated as antigens in immune-mediated neuropathies. Sulfate-3-glucuronyl paragloboside (SGPG) is the glycolipid sharing a carbohydrate epitope with myelin-associated glycoprotein (MAG), and the terminal sulfated glucuronic acid is a key part of the epitope. GM1 is the ganglioside implicated in motor nerve disorders, and in most cases the terminal Gal (β1-3) GalNAc epitope, which is shared with GD1b, is involved. The disialosyl moiety implicated in sensory neuropathies consists of NeuAca2—8NeuAc—and is present in GD1b and GT1b gangliosides, as well as the simpler GD2 and GD3 gangliosides (not shown). GQ1b ganglioside, which is the target antigen in Miller Fisher syndrome (MFS), has two disialosyl moieties. Although GD1a ganglioside has two sialic acid residues they are not linked to each other, so antibodies to GD1a do not cross-react with anti-GD1b antibodies. The color-coded sugar moieties represent key aspects of the various epitopes, but carbohydrate sequences recognized by the antibodies may include additional sugar residues. GlcUA, glucuronic acid; Gal, galactose; G1cNAc, *N*-acetylglucosamine; Glc, glucose; GalNAc, *N*-acetylgalactosamine; NeuAc, *N*-acetylneuraminic acid (sialic acid).

an acidic glycolipid in the ganglioside fraction of human peripheral nerves, identified as a sulfoglucuronyl glycosphingolipid (SGPG).[28] In contrast to MAG, which is mostly present in the central nervous system, SGPG is found only in the peripheral nerves. The sera of some patients with IgM-MGUS with sensory ataxia may not react with MAG, but with various gangliosides, most commonly those that contain either a disialosyl moiety, such as GD1b, GQ1b, GT1b, GalNac-GM1b, and GalNAc-GD1a, or two gangliosides that share epitopes with GM2, or a combination of GM2 and GM1, GM1 and GD1b.[2,3,6,24,27,28] More than half of the IgM paraproteins recognize MAG and SGPG, and 75% of the rest recognize ganglioside antigens, indicating that acidic glycolipids are the most common antigenic epitopes.[2,3,6,24,27,28] The glycolipids implicated as antigens in immune-mediated neuropathies are depicted in Fig. 67.4.

Human anti-MAG antibodies can be detected readily in sera with ELISA or preferably with standard Western blot. Because anti-MAG–reacting sera always recognize the SGPG glycolipid, the assay is often performed by using SGPG as antigen instead of purified human MAG. It is preferable, however, to use MAG as the target antigen, rather than SGPG, because the IgM binds to MAG 10–100 times the affinity compared with SGPG—that is, low-affinity anti-MAG antibodies can be missed if SGPG is used as the antigen.

The following factors suggest that these antibodies are related to the cause of the neuropathy:

1. IgM and complement are deposited on the myelinated fibers on the patient's sural nerve biopsy,[2,3] suggesting that activated complement may be needed in the induction of demyelination.
2. IgM recognizes neural cell adhesion molecules and colocalizes with MAG on the areas of the split myelin lamellae, suggesting involvement in myelin disadhesion. Skin biopsies from these patients have also confirmed the presence of IgM, complement C3d, and MAG deposition on the dermal myelinated fibers and the concurrent loss of nerve fibers.[29]
3. Injection of serum from patients with IgM anti-MAG/SGPG paraprotein supplemented with fresh complement into feline peripheral nerve causes complement-dependent demyelination and conduction block within 2–9 days of the injection.[30] The IgM injected intraneurally localizes to the outer layer of the myelin sheath.

KEY CONCEPTS

Autoimmunity in Polyneuropathy With Immunoglobulin G (IgM) Monoclonal Gammopathy

- In more than 50% of patients IgM is an antibody against two antigens, myelin-associated glycoprotein (MAG) and sulfoglucuronyl glycosphingolipid (SGPG).
- In many patients with non-MAG–reacting monoclonal IgM, the IgM recognizes *(i)* gangliosides containing disialosyl moieties, including GM1, GM2, GD1b, GD1a, and LM1; *(ii)* sulfatides; and *(iii)* rarely, chondroitin sulfate.
- Overall, in at least 75% of patients the IgM recognizes gangliosides that appear to be the primary antigenic targets.
- IgM is deposited on the homologous myelin sheath and fixes complement.
- IgM, when deposited on the myelin sheath, results in disadhesion and separation of the myelin lamellae and disruption of normal myelin function.
- Intraneural injection of anti-MAG–reacting IgM or passive transfusion into experimental animals causes segmental demyelination, whereas complement-fixing IgM is immunolocalized to the myelin sheath, causing myelin separation.
- Immunization of cats with purified SGPG causes an ataxic neuropathy, similar to the one seen in humans, with involvement of the dorsal root ganglia.

4. Systemic transfusion of anti-MAG IgM paraproteins produces segmental demyelination in chickens,[31] with deposition of IgM on to the outer lamellae of the myelin along with splitting of the myelin lamellae, similar to that observed in the human neuropathy.
5. Immunization of cats with purified SGPG causes an ataxic neuropathy, similar to the one seen in humans, with involvement of the dorsal root ganglia including inflammation within the ganglionic neurons.[32]

POLYNEUROPATHY, ORGANOMEGALY, ENDOCRINOPATHY, MYELOMA, AND SKIN CHANGES

A subset of patients with malignant IgG or IgA monoclonal proteins have polyneuropathy with osteosclerotic myeloma. Most of them have POEMS syndrome (*p*olyneuropathy, *o*rganomegaly, *e*ndocrinopathy, *M* protein, and *s*kin changes).[33] Not included in the acronym are several features, such as sclerotic bone lesions, giant lymph node hyperplasia (Castleman disease), papilledema, pleural effusion, edema, ascites, and thrombocytosis.[34] More than 50% of patients with osteosclerotic myeloma of the IgA or IgG type have a sensorimotor, symmetric polyneuropathy with mixed demyelinating and axonal features and high (usually 200 mg/dL) CSF protein. Pure axonal neuropathies can be also seen. In POEMS, the neuropathy tends to be associated with edema in the legs, with hyperpigmentation, sclerodermatous thickening, or papular angiomas of the skin, and hypertrichosis with dark hair. Endocrinopathy most often includes gonadal failure, amenorrhea, impotence, gynecomastia, hypothyroidism, diabetes, or elevated prolactin levels. The IgG class is slightly more common than the IgA class, with λ light chain present in the majority of the patients.[34] Bone lesions can be sclerotic, solitary, or multiple, sparing the skull and the extremities. Pathological changes in lymph nodes resemble those of Castleman disease, which can also be associated with polyneuropathy.[34] In POEMS, there is an imbalance of proinflammatory cytokines with increased IL-1β, IL-6, and tumor necrosis factor-α (TNF-α). Vascular endothelial growth factor (VEGF) may play a major role because it induces a rapid increase in vascular permeability and is a growth factor important in angiogenesis and endothelial cells.[34]

In some patients, the neuropathy responds to steroids, tamoxifen, or alkylating agents. In others, it may respond to the removal or irradiation of the solitary sclerotic lesion, suggesting that the tumor may secrete neurotoxic factors. IVIG and plasmapheresis are ineffective. The median survival is 165 months.[35]

Recently, autologous peripheral blood stem cell transplantation has been shown to result in significant improvement with reduction of VEGF, improved nerve conduction velocity, and increased survival.[35]

Cryoglobulinemic Neuropathy

Cryoglobulins are proteins that precipitate in the cold and redissolve when heated (Chapter 58). There are three types of cryoglobulin: type I, which is monoclonal, often of the IgM and IgG class; type II, which is mixed polyclonal, with one monoclonal (often monoclonal IgM with polyclonal IgG); and type III, which is polyclonal (often IgM and IgG). Polyneuropathy occurs most often with mixed cryoglobulinemias and presents as distal, sensory, symmetrical polyneuropathy or as mononeuropathy multiplex. These patients also have purpura, polyarthralgias, cutaneous vasculitis, Raynaud phenomenon, renal involvement with proteinuria, and polyneuropathy. Type II cryoglobulinemia may be associated with an underlying lymphoproliferative process. There is an increased incidence (up to 90%) of hepatitis C virus infection in patients with mixed cryoglobulinemia. The nerve biopsy shows perivascular inflammatory cuffing with axonal degeneration.

PARANEOPLASTIC PERIPHERAL NEUROPATHIES WITH ANTI-HU ANTIBODIES

Peripheral neuropathy is not an uncommon complication of cancer, related either to the systemic effects of the tumor or, more often, to various neurotoxic chemotherapeutic agents. It usually affects the small nerve fibers causing numbness or painful dysesthesias. The most distinct immune-related neuropathy in patients with cancer is paraneoplastic sensory neuronopathy (PSN), often associated with small cell lung cancer, and to a lesser degree with breast cancer or other neoplasms, such as lymphoma or thymoma.[36] It might be the presenting symptom preceding the discovery of the tumor by months. PSN has a unique clinical picture, characterized by burning or aching paresthesias; sensory loss of the large fibers, causing profound sensory gait ataxia and choreoathetotic movements related to loss of proprioception in the feet and hands; normal strength; and areflexia. Some patients may have autonomic dysfunction, encephalopathic symptoms, or cerebellar disturbances. CSF protein is increased, electrophysiological testing shows an axonal sensory neuropathy, and nerve biopsy demonstrates axonal degeneration with rare mononuclear cell infiltrates. PSN is a ganglionopathy (sensory neuronopathy) caused by a variable degree of inflammation in the dorsal root ganglionic neurons. These patients have specific IgG anti-Hu autoantibodies directed against a closely spaced group of proteins with a molecular weight of 35–40 kDa.[36] The antibodies are found in higher titers in CSF, suggesting intrathecal synthesis. The Hu protein is also present in the tumors of patients with PSN. Further, low titers of anti-Hu antibodies can be seen in up to 20% of patients with small cell lung cancer, even without neurological symptoms, suggesting that PSN may be the result of an autoimmune reaction against antigens shared by both the tumor cells and the dorsal root ganglionic neurons. Although the role of anti-Hu antibodies in the causation of PSN is unclear, these antibodies are specific markers to detect an occult small cell lung cancer in patients who present with sensory ataxic neuropathy. Some patients may have other paraneoplastic antibodies, such as collapsing response mediator protein (CRMP-5) often referred to as anti-CV2.[36]

AUTOIMMUNE AUTONOMIC NEUROPATHIES

Autoimmune autonomic neuropathy (AAN) is highlighted by circulating antibodies against the ganglionic nicotinic acetylcholine receptors (AChRs).[37] These patients present with a subacute (within 4 weeks) or chronic (within months) onset of neurogenic orthostatic hypotension, defined as a systolic blood pressure reduction of at least 30 mm Hg or mean blood pressure reduction of at least 20 mm Hg that occurs within 3 minutes of head tilting. The subacute onset is often preceded by a viral infection. In addition, patients demonstrate three or four parasympathetic/enteric symptoms: sicca (dry eyes and dry mouth); abnormal pupillary response to light; upper gastrointestinal symptoms (early satiety, postprandial nausea, and vomiting that lead to severe weight loss); and neurogenic bladder.

Patients with more severe cholinergic impairment have higher-titer antibodies against ganglionic AChRs. Some of these symptoms can be passively transferred to mice injected with the patient's IgG, suggesting that these antibodies may be pathogenic. Further, rabbits immunized with a fragment of ganglionic AChR protein exhibit autonomic failure, similar to the human disease.[37] Because ganglionic AChRs have also been found in small cell lung carcinoma cell lines, cancer may be a potential initiator of ganglionic AChR autoimmunity.

MONONEUROPATHY MULTIPLEX AND LOCALIZED, ISOLATED VASCULITIS OF THE PERIPHERAL NERVES

Polyneuropathy is a common manifestation of systemic vasculitis. It occurs in patients with polyarteritis nodosa; connective tissue diseases, such as rheumatoid arthritis or Sjögren syndrome; hypersensitivity vasculitis; Churg-Strauss syndrome; temporal arteritis; and viral infections, such as those with HIV, human T-lymphocyte virus (HTLV)-1, and hepatitis B and C viruses. It classically presents as mononeuritis multiplex affecting several individual nerves with painful weakness and paresthesias caused by ischemia and infarcts caused by inflammation of endoneurial blood vessels. There is, however, a distinct vasculitic entity localized only to the peripheral nerve, known as *isolated peripheral nerve vasculitis* (PNV). PNV involves the small and medium-sized arteries of the epineurium and perineurium and causes ischemic changes within the peripheral nerve. The presentation is similar to the vasculitic neuropathy seen in systemic vasculitis—the only difference is the lack of systemic organ involvement, slower onset and progression, and negative serology. The diagnosis is confirmed with nerve biopsy of the sural, superficial peroneal, or superficial radial nerve. When nerve biopsy is combined with muscle biopsy, the diagnostic yield is higher. PNV has a better prognosis compared with systemic vasculitides and is a treatable form of neuropathy. An evaluation for PNV should include all the tests needed to exclude systemic vasculitis and cryoglobulinemia, as well as hepatitis B and C infections, which can be associated with PNV.

NEUROPATHY WITH VIRUSES AND HUMAN IMMUNODEFICIENCY VIRUS

Neuropathy can be seen in a setting of infectious, viral, or bacterial processes. In patients with Lyme disease, various neuropathies, including GBS and mononeuritis (Bell palsy), have been noted. Other infections, such as CMV, hepatitis, herpes, leprosy, Chagas disease, and diphtheria, can affect the peripheral nerves, triggering autoimmune peripheral neuropathy.

Common neuropathies seen today in a setting of a viral infection are those associated with HIV and include GBS, CIDP, acute ganglioneuritis, or mononeuritis multiplex; they occur early in the infection, or they are the presenting manifestation of unsuspected HIV infection.[38] HIV has been cultured from the peripheral nerves, and HIV viral RNA has been amplified from sural nerve biopsies of some patients in the author's laboratory. However, there is no convincing evidence that the neuropathy results from direct infection of the peripheral nerves by the virus. Immunocytochemical studies have shown that HIV is present only in rare endoneurial macrophages, but not within Schwann cells or axons. A strong expression of HLA class I and class II molecules on Schwann cells, endothelial cells, and/or macrophages, but sparse presence of CD8 and CD4 T cells, is also noted (Fig. 67.5). It is possible that systemic viral infection or rare HIV-infected endoneurial lymphoid cells may release cytokines that expose new nerve antigens against which there is no self-tolerance, generating a tissue-specific autoimmune attack.

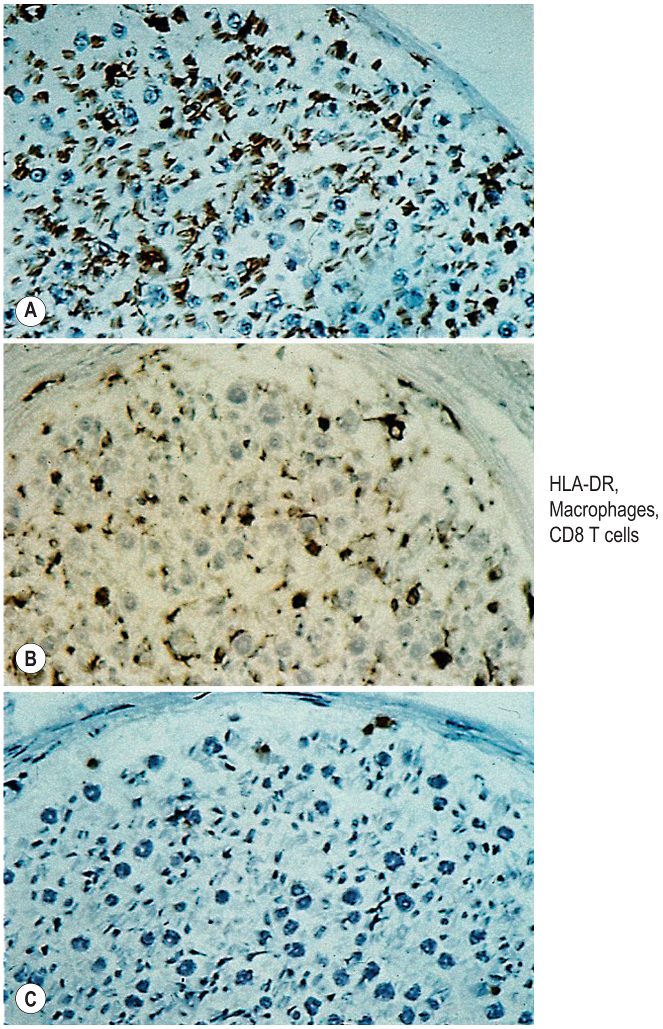


FIG 67.5 Serial sections of a nerve biopsy from a patient with human immunodeficiency virus (HIV)–chronic inflammatory demyelinating polyneuropathy stained for (A) human leukocyte antigen (HLA)-DR, (B) macrophages, and (C) CD8 T shows that the majority of the endoneurial cells are macrophages. Only rare CD8 cells are noted.

A rare neuropathy seen in later-stage HIV infection is a lumbosacral polyradiculoneuropathy related to CMV infection that affects roots and sensory ganglia. It presents with lower-extremity muscle weakness, sacral and distal paresthesias, areflexia, and atrophy, mostly of the legs, associated with sphincteric dysfunction resembling cauda equina syndrome. CMV inclusions can be found within Schwann cells or endothelial cells (Fig. 67.6). Early recognition is important because anti-CMV therapy with ganciclovir or foscarnet can be helpful.

At present, the most common neuropathy in patients with acquired immunodeficiency syndrome (AIDS) is a painful sensory axonal neuropathy that affects up to 70% of adults with AIDS and

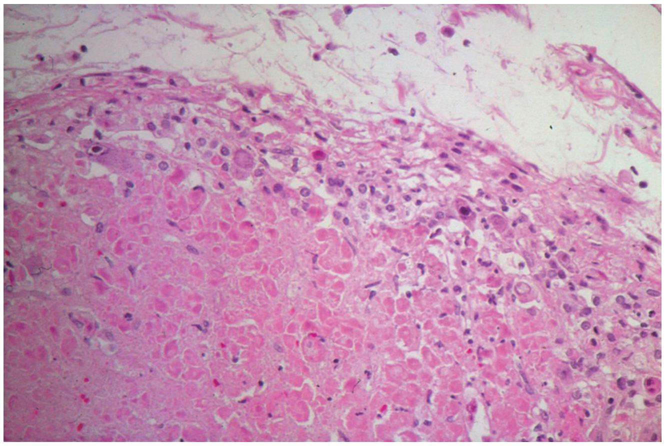


FIG 67.6 Cross-section of a root from a patient with human immunodeficiency virus (HIV)–associated Guillain-Barré syndrome (GBS) shows cytomegalovirus inclusions within the Schwann cell.

occurs later in the disease.[38] It is caused by a cumulative effect, on the peripheral nerves, of various endogenous or exogenous neurotoxins related to a multisystem disease and dysfunction of many organs along with toxicity from various antiretroviral drugs. Clinical findings include distal painful dysesthesias, sensory loss or hypesthesia, areflexia, and, in advanced cases, distal weakness. Despite the relative lack of motor involvement, the severity of neuropathic pain can be disabling.

TREATMENT

APNs are clinically important because they are potentially treatable with various immunosuppressive, immunomodulating, or chemotherapeutic agents. The selection of an effective protocol is based on the results of experimental therapeutic trials, clinical experience, and the risk–benefit ratio of available therapies. The author's approach to the treatment of these disorders is described below.

Guillain-Barré Syndrome

Supportive Care

The dramatic reduction in the mortality of GBS is mainly attributed to the availability of ICUs, improvement of respiratory support, antibiotic therapy, and control of autonomic cardiac dysregulation. A patient with GBS is best monitored in an ICU, even if respiratory compromise is not evident at the time of admission. When forced vital capacity (FVC) drops or bulbar weakness is severe, intubation is necessary.

Plasmapheresis

In several double-blind controlled studies, plasmapheresis has been shown to be effective if performed within the first week from onset of the illness. A series of five or six exchanges, with one exchange every other day, is sufficient. Early relapses can occur in up to 20% of patients, who may require a second series of plasma exchanges.[1-5] Plasmapheresis has been shown to be effective even in mild cases of GBS; two exchanges are sufficient for mild GBS, and four are optimal for moderate cases, but there is no difference between those who receive four plasma exchanges and those who receive six.

High-Dose Intravenous Immunoglobulin

On the basis of the results from two controlled studies,[1,4] IVIG (Chapter 84), given at 2 g/kg over 2–5 days, has been shown to be equally effective as plasmapheresis, with no added benefit when the two procedures were combined. The decision as to which treatment to choose is governed by circumstances, availability of the treatment modality, experience, age of the patient, and other associated conditions. Early relapses can also occur with IVIG, as often as with plasmapheresis. IVIG has become the therapeutic choice worldwide because it is easy to administer and more readily available and because time to initiate treatment is of the essence.

Steroids are ineffective in GBS and may even increase the incidence of future relapses. Combining IVIG with IV methylprednisolone has shown no significant added benefit.

Chronic Inflammatory Demyelinating Polyneuropathy

Prednisone

CIDP has been originally described as a classic steroid-responsive polyneuropathy. The efficacy of steroids was proven in a controlled study, albeit with inadequate blinding, but reconfirmed in another.[2,3,11,12,39] A high-dose regimen of 80–100 mg prednisone daily is preferred, followed by tapering to every-other-day dosing. Azathioprine, cyclosporine, or mycophenolate can be used as steroid-sparing agents, but their efficacy, although not tested in controlled studies, has been disappointing overall. Methotrexate in a controlled study was ineffective.[2,3,11,12,39]

Intravenous Immunoglobulin

In several controlled studies,[2,3,11,12,39] IVIG has been effective in the majority of patients with CIDP. The more chronic the disease and more severe the axonal degeneration that has taken place, the lower are the chances that the recovery will be complete or significant. IVIG, in the largest-ever controlled study, has been proven an effective first-line therapy, and maintenance therapy has prevented relapses.

Plasmapheresis

Plasmapheresis has been also effective in controlled studies.[2,3,11,12] After a series of six plasma exchanges, maintenance therapy, with one exchange at least every 8 weeks, may be required if this therapy is beneficial. IVIG has now replaced plasmapheresis, although in the author's experience, some patients may benefit more from steroids, others more from IVIG, and still others more after plasmapheresis.

Polyneuropathy With Paraproteinemias

Patients with benign IgG or IgA demyelinating polyneuropathies respond in a manner similar to CIDP patients. Patients with malignant paraproteinemias should be treated with chemotherapy, as needed for the underlying disease. When the neuropathy is axonal, treatments are generally disappointing.

For IgM anti-MAG demyelinating polyneuropathies, treatments with prednisone plus chlorambucil, plasmapheresis, and IVIG[2,3,40] has shown a variably marginal benefit. Rituximab, an mAb against CD20, is the most promising therapy,[41] providing efficacy in almost 40% of the patients in a small double-blind study, as confirmed later with a larger study,[33] even though both did not reach significance. Additional, uncontrolled series with many patients, have confirmed that rituximab is effective in 30–40% of these patients.

MULTIFOCAL MOTOR NEUROPATHY

MMN responds very well only to IVIG, which is the treatment of choice based on controlled trials. In difficult cases, cyclophosphamide or rituximab may be promising, but no controlled studies have been carried out.[2,3,21]

Paraneoplastic Anti-Hu Neuropathy

Anecdotally, some of these patients have responded to plasma exchange or IVIG, but overall this neuropathy has not been consistently responsive to available therapies.

Vasculitic Neuropathies

For isolated peripheral nerve vasculitis, a combination of prednisone 1.5 mg/kg/day with cyclophosphamide 2 mg/kg/day orally, or 1 g/m^2 intravenously monthly for 6 months, are the treatments of choice; specific treatment protocol may vary from patient to patient. The administration of cyclophosphamide may not be necessary for >6 months, as in systemic vasculitis. Plasmapheresis has been tried in cryoglobulinemic neuropathies, with variable results.

Human Immunodeficiency Virus Neuropathies

The two demyelinating neuropathies, GBS and CIDP, are treated with the same immunomodulatory therapies as used in HIV-negative patients with IVIG being the preferable choice. Ganciclovir may be helpful in CMV-related polyradiculoneuropathy. Painful sensory neuropathy can be quite disabling because of intractable pain. Tricyclic antidepressants (TCAs), nonsteroidal antiinflammatory drugs (NSAIDs), anticonvulsants (carbamazepine, gabapentin, topiramate), and topical capsaicin in various combinations provide some relief from neuropathic pain.

ON THE HORIZON

- Identification of common causes that break tolerance, either in connection with the gut microbiome or exogenous agents, and expansion of the current limited data on molecular mimicry between nerve glycoproteins and viral or bacterial antigens
- Development of neuroimaging capabilities to image peripheral nerves and dorsal roots *in vivo* and quantify inflammation, demyelination and axonal degeneration, providing an accessible and reliable tool for diagnosis and monitoring response to therapies, in a model similar to magnetic resonance imaging (MRI) in multiple sclerosis
- Performance of proteome studies aimed to identify biomarkers of nerve autoimmunity that may lead to specific therapies
- Trials with target-specific therapies for acute and chronic demyelinating neuropathies, *applied in combination with the existing therapies,* targeting early in the disease process complement activation, regulatory T cell (Treg) functions, key proinflammatory cytokines, and B cells
- Identification of agents and trophic factors that could prevent axonal degeneration from the outset, promote remyelination, prevent axonal loss, or trigger axonal regeneration and reverse, early in the disease process, a seemingly "permanent" clinical deficit

ACKNOWLEDGMENTS

The author thanks the various Neurology Fellows who provided excellent care to his patients; and Drs. Amjad Ilyas and Richard Quarles for help with immunochemistry of MAG and SGPG and gangliosides.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Yuki N, Hartung H-P. Guillain-Barre syndrome. N Engl J Med 2012;366:2294–304.
2. Dalakas MC. Pathophysiology of autoimmune polyneuropathies. Presse Med 2013;42(6 Pt 2):e181–92.
3. Dalakas MC. Pathogenesis of autoimmune neuropathies. Biochim Biophys Acta 2015;1852:658–66.
4. Willison HJ, Jacobs BC, van Doorn PA. Guillain-Barré syndrome. Lancet 2016 Feb 29. pii: S0140-6736(16).
5. Kuwabara S, Yuki N. Axonal Guillain-Barre syndrome: concepts and controversies. Lancet Neurol 2013; 12; 1180-8.
6. Dalakas MC, Quarles RH, editors. Autoimmune ataxic neuropathies (sensory ganglionopathies): are glycolipids the responsible autoantigens? Ann Neurol 1996;39:419–22.
7. Illa I, Ortiz N, Gallard E, et al. Acute axonal Guillain–Barré syndrome with IgG antibodies against motor axons following parenteral gangliosides. Ann Neurol 1995;38:218–24.
8. Cao-Lormeau V-M, et al. Guillain-Barre Syndrome outbreak associated with Zika virus infection in French Polynesia: a case controlled study. Lancet 2016;387(10027):1531–9.
9. Uncini A, Kuwabara S. Nodopathies of the peripheral nerve: an emerging concept. J Neurol Neurosurg Psychiatry 2015;86:1186–95.
10. Stathopoulos P, Alexopoulos H, Dalakas MC. Autoimmune antigenic targets at the node of Ranvier in demyelinating disorders. Nat Rev Neurol 2015;11:143–56.
11. Dalakas MC. Advances in the diagnosis, pathogenesis and treatment of CIDP. Nat Rev Neurol 2011;7(9):507–17.
12. Koller H, Kieseier BC, Jander S, et al. Chronic inflammatory demyelinating polyneuropathy. N Engl J Med 2005;352:1343–56.
13. Schneider-Hohendorf T, Schwab N, Uceyler K, et al. T cell immunity in CIDP. Neurology 2012;78:402–8.
14. Murata K, Dalakas MC. Expression of the co-stimulatory molecule BB-1, the ligands CTLA-4 and CD28 and their mRNAs in chronic inflammatory demyelinating polyneuropathy. Brain 2000;123:1660–6.
15. Salomon B, Rhee L, Bour-Jordan H, et al. Development of spontaneous autoimmune peripheral polyneuropathy in B7-2-deficient NOD mice. J Exp Med 2001;194:677–84.
16. Dalakas M, Engel WK. Immunoglobulin deposits in chronic relapsing polyneuropathies. Arch Neurol 1980;37:637–40.
17. Devaux JJ, Miura Y, Fukami Y, et al. Neurofascin-155 IgG4 in chronic inflammatory demyelinating polyneuropathy. Neurology 2016;86:800–7.
18. Miura Y, Devaux JJ, Fukami Y, et al. Contactin 1 IgG4 associates to chronic inflammatory demyelinating polyneuropathy with sensory ataxia. Brain 2015;138:1–8.
19. Dalakas MC, Gooch C. Close to the node but still far: What antibodies tell us about CIDP and its therapies. Neurology 2016;86:796–7.
20. Weiss MD, Luciano CA, Semino-Mora C, et al. Molecular mimicry in chronic inflammatory demyelinating polyneuropathy and melanoma. Neurology 1998;51:1738–41.
21. Van Asseldonk JT, Franssen H, Van den Berg-Vos RM, et al. Multifocal motor neuropathy. Lancet Neurol 2005;4:309–19.
22. Kyle RA, Therneau TM, Rajkumar SV, et al. Prevalence of monoclonal gammopathy of undetermined significance. N Engl J Med 2006;354:1362–9.
23. Kelly JJ, Kyle RA, O'Brien PC, et al. The prevalence of monoclonal gammopathy in peripheral neuropathy. Neurology 1981;31:1480–3.
24. Latov N, Hays A, Sherman WH. Peripheral neuropathy and anti-MAG antibodies. Crit Rev Neurobiol 1988;3:301–32.
25. Dalakas MC. Pathogenesis and treatment of anti-MAG neuropathy. Curr Treat Options Neurol 2010;12:71–83.
26. Dalakas MC, Engel WK. Polyneuropathy and monoclonal gammopathy: studies of 11 patients. Ann Neurol 1981;10:45–52.

27. Ilyas AA, Quarles RH, McIntosh TD, et al. IgM in a human neuropathy related to paraproteinemia binds to a carbohydrate determinant in the myelin-associated glycoprotein and to a ganglioside. Proc Natl Acad Sci USA 1984;81:1225–9.
28. Ilyas AA, Quarles RH, Dalakas MC, et al. Polyneuropathy with monoclonal gammopathy: glycolipids are frequently antigens for IgM paraproteins. Proc Natl Acad Sci USA 1985;82:6697–700.
29. Lombardi R, Erne B, Lauria G, et al. IgM deposits on skin nerves in anti-myelin-associated glycoprotein neuropathy. Ann Neurol 2005;57:180–7.
30. Willison HJ, Trapp BD, Bacher JD, et al. Demyelination induced by intraneural injection of human antimyelin associated glycoprotein antibodies. Muscle Nerve 1988;11:1169–76.
31. Tatum AH. Experimental paraprotein neuropathy; demyelination by passive transfer of human IgM anti-MAG. Ann Neurol 1993;33:502–6.
32. Ilyas AA, Gu Y, Dalakas MC, et al. Induction of experimental ataxic sensory neuronopathy in cats by immunization with purified SGPG. J Neuroimmunol 2008;193:87–93.
33. Leger JM, Guimaraes-Costa R, Muntean C. Immunotherapy in peripheral neuropathies. Neurother 2016;13:96–107.
34. Dispenzieri A, Kyle RA, Lacy MQ, et al. POEMS syndrome: definitions and long-term outcome. Blood 2003;101:2496–506.
35. D'Souza A, Lacy M, Gertz M, et al. Long-term outcomes after autologous stem cell transplantation for patients with POEMS syndrome (osteosclerotic myeloma): a single-center experience. Blood 2012;120:56–62.
36. Muppidi S, Vernino S. Paraneoplastic neuropathies. Continuum (Minneap Minn) 2014;20:1359–72.
37. Lennon VA, Ermilov LG, Szurszewski JH, et al. Immunization with neuronal nicotinic acetylcholine receptor induces neurological autoimmune disease. J Clin Invest 2003;111:907–13.
38. Dalakas MC, Pezeshkpour GH. Neuromuscular diseases associated with human immunodeficiency virus infection. Ann Neurol 1988;23:38–48.
39. Hughes RA, Donofrio P, Bril V, et al. Intravenous immune globulin (10% caprylate-chromatography purified) for the treatment of chronic inflammatory demyelinating polyradiculoneuropathy (ICE study): a randomised placebo-controlled trial. Lancet Neurol 2008;7:136–44.
40. Dalakas MC, Quarles RH, Farrer RG, et al. A controlled study of intravenous immunoglobulin in demyelinating neuropathy with IgM gammopathy. Ann Neurol 1996;40:792–5.
41. Dalakas MC, Rakocevic G, Salajegheh M, et al. Placebo-controlled trial of rituximab in IgM anti-myelin associated glycoprotein antibody demyelinating neuropathy. Ann Neurol 2009;65:286–93.

MULTIPLE-CHOICE QUESTIONS

1. Which of the following findings predicts an incomplete recovery in patients with Guillain-Barré syndrome (GBS)?
   A. Slow initial progression
   B. Absence of bulbar findings
   C. Presentation in the third decade
   D. Requirement of mechanical ventilation

2. Which is the most common form of GBS?
   A. Acute inflammatory demyelinating polyneuropathy
   B. Acute motor axonal neuropathy
   C. Acute motor–sensory neuropathy
   D. Miller Fisher syndrome

3. Which of the following tests is the most useful in establishing the diagnosis of GBS?
   A. Sural nerve biopsy
   B. Cerebrospinal fluid (CSF) protein levels
   C. Nerve conduction study
   D. Brain magnetic resonance imaging (MRI)

4. Which of the following finding is characteristic of chronic inflammatory demyelinating polyneuropathy (CIDP)?
   A. Asymmetrical weakness
   B. Sensory dysfunction
   C. Normal peripheral reflexes
   D. Cranial nerve involvement

68

Immunological Renal Diseases

Meryl Waldman, Howard A. Austin III, James E. Balow

Compelling clinical, pathological, and experimental data indicate that most forms of glomerular diseases manifest some element of immune-mediated injury.[1] Major advances in our understanding of the etiology and pathogenesis of the various glomerular diseases have occurred in the past decade. However, the primary events or inciting factors that trigger the host immune responses and secondary pathways that are directly or indirectly pathogenic to the kidney are still not fully elucidated. In some conditions, normal or allergic immune responses to exogenous infectious agents or drugs can lead to incidental nephropathic injury (*e.g.,* postinfectious glomerulonephritis, serum sickness). In the case of autoimmune diseases, nephropathic processes range from loss of self-tolerance to specific constitutive renal tissues (*e.g.,* antiglomerular basement membrane [GBM] disease) or to extrarenal tissues and cells (*e.g.*, antineutrophil cytoplasmic autoantibodies [ANCAs]). Other mechanisms include excessive and unregulated polyclonal immune responses that produce nephritogenic immune complexes (*e.g.,* lupus nephritis).[1,2] Recent advances have begun to uncover candidate antigens involved in the pathogenesis of immune-mediated renal diseases (*e.g.,* some subsets of membranous nephropathy [MN]). But it has become increasingly apparent that the immune system does not operate in isolation and that there are extraordinarily complex interactions among vast networks involving components of native and adaptive immune systems, humoral and cell-mediated immune systems, small-molecular-weight mediators derived from lymphoid and other cell types, complement factors and their regulatory proteins, other inflammatory and coagulation system pathways, microvascular biology, and tissue repair, as well as sclerosis and fibrosis reactions. Underlying this constellation are associated genetic risk factors that modulate these responses and predispose to a nephritogenic response. Rapidly evolving molecular and genetic technologies with the tools of modern systems biology and bioinformatics will continue to unravel these complex interactions.[3]

Appropriate evaluation of patients with immune-mediated kidney diseases requires particular attention to the findings of urinalysis, assessment of the type and amount of proteinuria, tests of renal functions, and renal biopsy. Urine microscopy also plays a pivotal role in the assessment of urinary disorders.

HEMATURIA

Red blood cells (RBCs) that emanate from lesions of the calices, ureters, bladder, or urethra tend to maintain their normal morphology; when nondysmorphic RBCs are present, it is usually appropriate to refer the patient to urology for further evaluation. Urinary RBCs that result from glomerular or tubulointerstitial pathology are more likely to appear dysmorphic (abnormal shapes and sizes, fragmented); when these dysmorphic RBCs, called *acanthocytes,* are present, it is usually appropriate to refer the patient to nephrology for further evaluation. Erythrocyte and/or leukocyte casts are indicative of glomerulonephritis (or interstitial nephritis). Cellular casts can be formed from erythrocytes and/or leukocytes that enter the tubular lumen because of glomerular or tubular inflammation (Fig. 68.1).

PROTEINURIA

Glomerular proteinuria results from a loss of the size-selective and/or charge-selective properties of the glomerular capillary wall and disruptions of the glomerular epithelial cells (podocytes), allowing plasma proteins (especially albumin) to leak into the filtrate. Tubulointerstitial nephropathy often leads to impaired tubular absorption of other normally filtered low-molecular-weight proteins; this so-called tubular proteinuria exhibits a characteristic pattern on urine protein electrophoresis (low fraction of albumin), rarely exceeds 2 g/day, and is often associated with other manifestations of tubular dysfunction.

Estimates of proteinuria are based on 24-hour urine collections or on the protein/creatinine or albumin/creatinine concentration ratios in random urine samples. Normally, these ratios should be <0.1. Filtration of abnormal plasma proteins can be responsible for proteinuria (*e.g.,* multiple myeloma or monoclonal immunoglobulin [Ig] light-chain disease). This type of proteinuria (paraproteinuria) may be undetected by albumin-sensitive dipstick screening. Paraproteinemia may also be undetected or underestimated by standard plasma electrophoresis; however, an assay for serum free light chains is more sensitive for detection of low levels of circulating light chain paraproteins. Identification of specific Ig-derived paraproteinuria usually requires urine immunofixation electrophoresis (Fig. 68.2)

NEPHROTIC SYNDROME

Nephrotic syndrome is characterized by substantial degrees of proteinuria (>3.5 g/day or protein to creatinine ratio >2.5) resulting in hypoalbuminemia, edema, hyperlipidemia, and lipiduria. The degree of proteinuria can offer a helpful diagnostic clue because some immune-mediated conditions with typically diffuse glomerular disease (*e.g.,* MN, systemic lupus erythematosus [SLE]) are more likely than others with typically focal disease (*e.g.*, IgA nephropathy, ANCA-associated glomerulonephritis) to be associated with nephrotic syndrome.

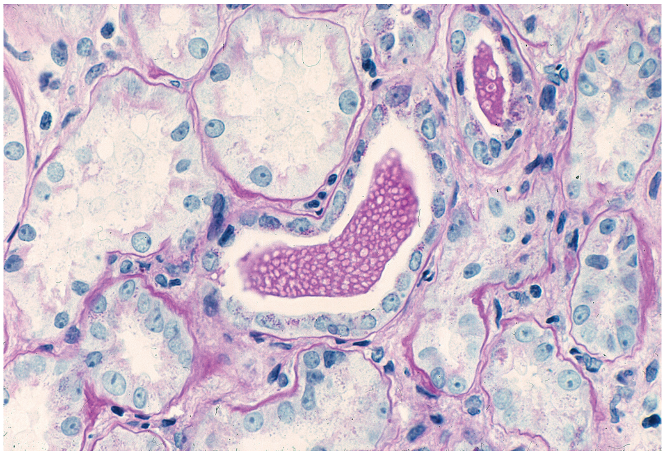


FIG 68.1 **Red Blood Cell Cast.** Cast present *in situ* within the lumen of a distal renal tubule. (Periodic acid–Schiff [PAS] stain.)

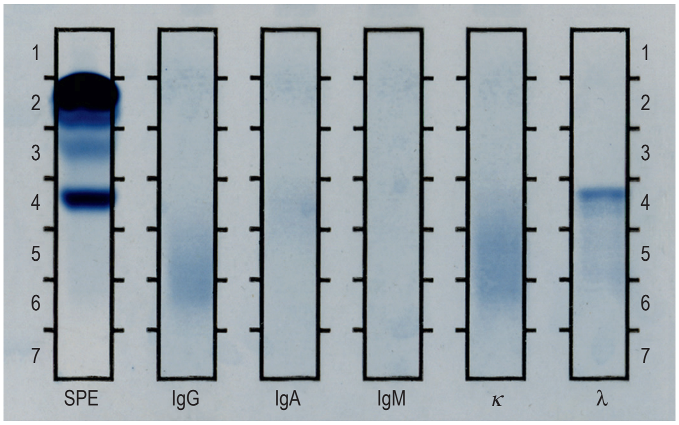


FIG 68.2 **Immunofixation Electrophoresis of Urinary Protein.** Proteins in a concentrated urine protein are separated in six replicate lanes by standard protein electrophoresis. Separated proteins are identified by overlaying specific antisera to immunoglobulin G (IgG), IgA, IgM, κ and λ. In this example, a monoclonal paraprotein composed of λlight chain is identified as an intense narrow band in the far right hand lane.

ACUTE NEPHRITIC SYNDROME

Acute nephritic syndrome is characterized by hematuria (dysmorphic cells), erythrocyte casts, abnormal proteinuria, fluid retention, azotemia, and hypertension. Histologically, this constellation of clinical findings is due to proliferative glomerulonephritis. A variant of this syndrome, called *rapidly progressive glomerulonephritis* (RPGN), is defined by ≥50% loss of glomerular filtration rate over 3 months, along with characteristically >50% of glomeruli showing cellular crescents on renal biopsy..

CHRONIC GLOMERULONEPHRITIS

Broad and waxy casts are features of chronic renal disease that are not likely to be seen in acute or subacute glomerulonephritis.

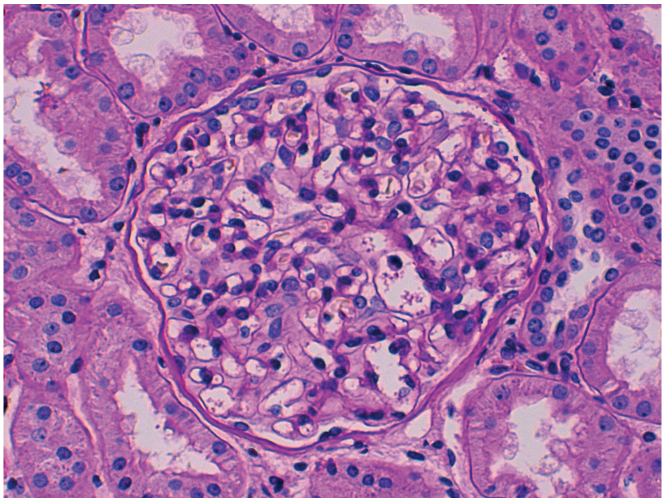


FIG 68.3 **Normal Glomerular Architecture.** The glomerular capillary loops are patent and have normal thickness. Neither increased endocapillary cells nor expanded mesangial matrix encroach upon the patency of the capillary lumina. (Periodic acid–Schiff [PAS] stain).

TABLE 68.1 Indications for Renal Biopsy

1. Active "nephritic" urine sediment
    - Dysmorphic erythrocytes: >10 per high-power field
    - Cellular casts: erythrocyte or leukocyte
2. Proteinuria >2 g/day
3. Abnormal renal function
    - Associated with the above features of active nephritis
    - Particularly important if the duration of renal disease and/or rate of change are unknown
4. Document indications for use of high-risk therapeutic interventions

RENAL BIOPSY

After extensive clinical and laboratory evaluations, renal biopsy may be indicated to establish or confirm a tissue diagnosis, to clarify the precise type of renal involvement, to formulate a prognosis, and to direct therapy. Some of the more important indications for renal biopsy are listed in Table 68.1. Light microscopy appearance of a normal glomerulus is illustrated in Fig. 68.3.

MINIMAL CHANGE DISEASE

KEY CONCEPTS

Minimal Change Nephropathy

- Most common cause of nephrotic syndrome in children
- High rate of response to glucocorticoids
- Cyclophosphamide is useful for frequent relapsers
- Renal prognosis is characteristically excellent
- Subset may evolve to focal segmental glomerulosclerosis

Nephrotic syndrome of childhood is mainly caused by minimal change disease (MCD). The frequency of MCD in adults with nephrotic syndrome is low compared with those of other entities discussed below.[4]

Clinical Features

Minimal change nephropathy characteristically presents with a rather precipitous onset of severe nephrotic syndrome in the absence of signs of a systemic disease. There are no specific tests other than kidney biopsy to establish a diagnosis of minimal change nephropathy, although proteinuria exclusively comprising albumin is characteristic; standard immunological screening tests are usually normal.

Etiology and Pathogenesis

The cause of MCD is largely unknown. There is evidence of immune dysregulation, mainly involving cell-mediated immunity, which is supported by the tendency of the disease to manifest and relapse after viral infections or an allergic reaction, the association of some cases with Hodgkin lymphoma, and its characteristically favorable response to immunosuppressive drugs. Another hypothesis is that a systemic circulating factor of immune origin results in increased glomerular permeability[5] (discussed in the section on "Focal Segmental Glomerulosclerosis")

The basic lesion of MCD is loss or neutralization of the normal high density of anionic proteoglycans in the glomerular capillary loops. Dissipation of the negative-charge barrier allows anion-charged albumin to pass freely. Interdigitating foot processes normally joined by intercellular bridges, called *slit diaphragms,* form a second barrier to the passage of protein into the urinary space. Characteristically, the podocyte foot processes and slit diaphragms are also disrupted in minimal change nephropathy.

Pathology

With the use of light microscopy, renal biopsy results are found to be essentially normal. Electron microscopy shows the characteristic pathological lesion, where there is fusion of the foot processes of the epithelial cells (podocytes) diffusely around glomerular capillaries.

Treatment

MCD is characteristically exquisitely sensitive to glucocorticoids (remitting within the first few weeks of therapy in >90% of children). The response rate to glucocorticoids in adults is somewhat lower and more delayed. A substantial portion of patients with minimal change nephropathy face long-term difficulties because of the disease and the therapy: Some are steroid-resistant from the start; others are steroid-dependent for control of nephrotic syndrome; still others become frequent relapsers and suffer substantial steroid toxicity with repeated treatments. Controlled trials have shown that alkylating agents (*i.e.,* cyclophosphamide) increase response rates and reduce rates of relapse. Cyclosporine (CSA) and tacrolimus are alternatives to prolonged and repeated courses of steroid therapy, but relapses frequently occur with drug withdrawal. Rituximab (a monoclonal antibody [mAb] to CD20 on B cells) has been used with variable success, particularly in those who are steroid-dependent or frequent relapsers. Also, rituximab may have "off target" effects on the podocyte cytoskeletal structure.[6]

The risk of progression to end-stage renal failure is extremely low in true MCD. Renal biopsy sampling errors account for a mistaken diagnosis of MCD in a proportion of cases that are, in fact, focal segmental glomerulosclerosis (FSGS). It is debated whether MCD and FSGS represent different manifestations of one disease (with MCD potentially progressing to FSGS), or if they represent two different diseases.

FOCAL SEGMENTAL GLOMERULOSCLEROSIS

Patients with FSGS have a higher frequency of microscopic hematuria, higher frequency of persistent nephrotic syndrome, a poorer response to immunosuppressive drug therapies, and a higher risk of progression to end-stage renal failure. The incidence of FSGS is clearly increasing as a cause of nephrotic syndrome, particularly among black patients.

KEY CONCEPTS

Focal Segmental Glomerulosclerosis

- Nephrotic syndrome with progressive renal insufficiency
- Glomerular permeability factor in plasma of some cases
- Unpredictable responses to glucocorticoids or cyclophosphamide
- Cyclosporine effective, but relapses common upon withdrawal
- Moderately high relapse rate in renal allografts

Etiology and Pathogenesis

Diverse etiologies may disrupt the podocyte structure and function leading to the histological expression of FSGS. Genetic mutations of key podocyte proteins, such as podocin and nephrin play a role in a subset of FSGS that occurs in children or young adults. Circulating permeability factors have been proposed as mediators of podocyte injury in FSGS (*i.e.,* soluble urokinase-type plasminogen activator receptor). In certain situations, it appears that podocytes have increased surface expression of the transmembrane protein B7-1 (CD80) that stimulates interactions with T cells and leads to foot process effacement.[7] "Secondary" forms of FSGS may result from toxic effects of drugs (*i.e.,* pamidronate), viral infections (*e.g.,* human immunodeficiency virus [HIV]) or maladaptive hemodynamic stress (*i.e.,* resulting from obesity or reduced nephron mass).

Pathology

The characteristic renal pathology includes segmental areas of podocyte damage and detachment, irregular foot process fusion, and collapse of glomerular capillaries associated with marked local increase in matrix and collagen accumulation. Segmental sclerotic areas typically stain nonspecifically with antisera to IgM and C3 (but not to IgG or IgA), particularly in areas of glomerular tuft hyalinosis (representing trapped plasma constituents), but none of these collections should be considered to represent classic immune complexes. Various histological subtypes of FSGS have been described on the basis of the type and location of lesions. One classification scheme divides FSGS into five variants: tip, perihilar, cellular, collapsing, and not otherwise specified (NOS)[8] (Fig. 68.4). The collapsing variant of FSGS, which has been associated with viral infections, notably HIV, tends to follow a particularly aggressive course.

Treatment

Patients with primary FSGS generally have hypoalbuminemia and nephrotic range proteinuria, whereas those with secondary FSGS typically have subnephrotic proteinuria, a normal or near normal serum albumin concentration, no peripheral edema, and focal rather than diffuse foot process effacement on kidney biopsy. The possibility of genetic forms of FSGS should be considered before beginning immunosuppressive therapy, but steroid resistance may also serve as a clue to genetic forms of FSGS.

Treatment of genetic and hyperfiltration-induced forms of FSGS is focused mostly on renoprotection with angiotensin antagonists and lipid-lowering agents (statins). The treatment of other forms of primary FSGS is basically similar to that of MCD and includes immunosuppression with prednisone,

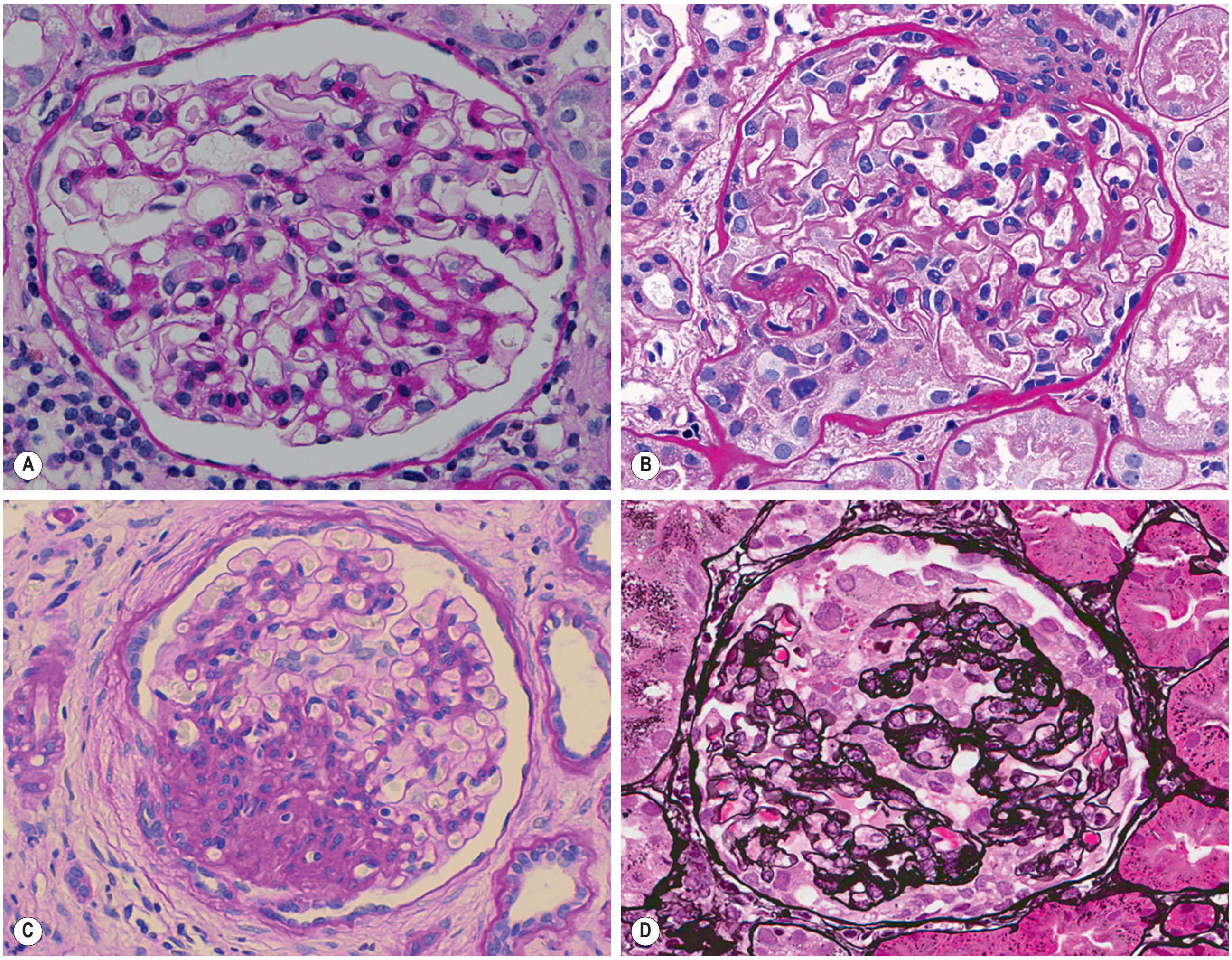


FIG 68.4 **Focal Segmental Glomerulosclerosis (FSGS).** Several forms of glomerular lesions are seen in FSGS, often within the same biopsy: A, minimal abnormality (periodic acid–Schiff [PAS] stain); B, tip lesion manifested by segmental glomerular tuft lesion near the origin of the proximal tubule (PAS stain); C, classical perihilar lesion (PAS stain); and D, collapsing glomerulopathy; the glomerulus is globally contracted with wrinkling of the basement membranes; this is associated with hyperplasia of podocytes surrounding the glomerular capillaries (methenamine silver stain).

cyclophosphamide, calcineurin inhibitors, or mycophenolate. There are mixed outcomes with use of rituximab for FSGS, but it appears less effective for steroid resistant cases.[9,10] Complete remission of proteinuria is less commonly achieved in FSGS than in MCD. Relapses of the disease and progression to end-stage kidney disease (ESKD) remains a major concern despite immunosuppression, particularly in steroid-resistant and frequently relapsing patients. Several novel therapies that target various immunological, inflammatory, and costimulatory pathways (*i.e.*, blockade of CD 80 with cytotoxic T lymphocyte antigen 4-Ig (CTLA-4-Ig) are currently under investigation for FSGS.[11,12]

MEMBRANOUS NEPHROPATHY

MN is identified in approximately 20% of adults who undergo renal biopsy for nephrotic syndrome. It is a less frequent cause of nephrotic syndrome in children. Primary MN is usually a diagnosis of exclusion after considering secondary causes, such as medication reactions, infections, neoplasms, and systemic illnesses (*i.e.*, SLE)

KEY CONCEPTS

Membranous Nephropathy (MN)

- Common cause of idiopathic nephrotic syndrome in adults
- Autoantibodies to phospholipase A_2 receptor (PLA_2R) detected in large percentage of patients with primary MN
- Several secondary causes: systemic lupus erythematosus (SLE), drugs, chronic hepatitis, certain malignancies
- One-quarter of patients have spontaneous remission
- One-quarter of patients develop end-stage renal disease within a decade
- Protracted nephrotic syndrome confers risks of cardiovascular and thromboembolic events
- Therapies: steroids, alkylating agents, calcineurin inhibitors, rituximab, lipid-lowering drugs, angiotensin antagonists

Etiology and Pathogenesis

MN is characterized by subepithelial (epimembranous) immune deposits containing IgG and complement components. The leading hypothesis based on experimental models is that immune

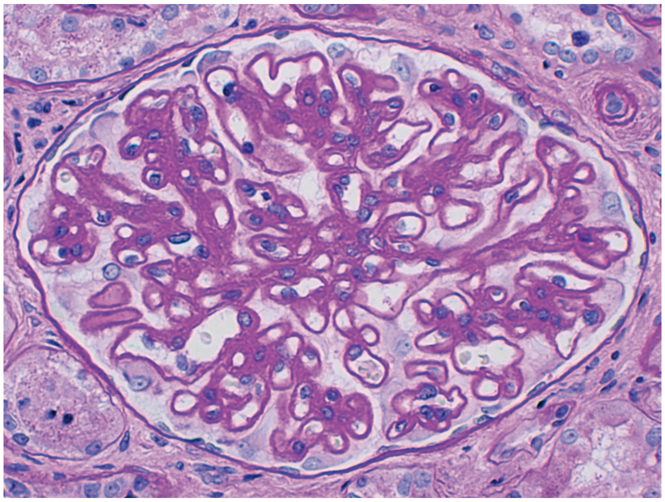


FIG 68.5 **Membranous Nephropathy.** Capillary walls are nearly uniformly thickened but remain widely patent. Cellularity of the glomerulus is normal (Periodic acid–Schiff [PAS] stain).

complexes are formed *in situ* by the interaction of a pathogenic antibody with a constitutive glomerular antigen or with an antigen that had been ectopically planted in the glomerulus. Major advances have been made with the identification of several constitutive podocyte surface antigens, which are the target of autoantibodies in some subsets of patients with primary MN.[13,14] The majority (70-80%) of patients with primary MN have circulating autoantibodies to M-type phospholipase A2 receptor (PLA_2R), a transmembrane receptor that is expressed in glomerular podocytes.[13] Antibodies to thrombospondin type-1 domain containing 7A protein (THSDA7A) and neutral endopeptidase have been identified in smaller subsets of patients.[14]

Clinicopathological Features

Patients with MN characteristically present with nephrotic syndrome. Renal biopsy is usually required to establish the diagnosis, although serological testing for relevant autoantibodies (*i.e.,* anti–PLA_2R) may be informative if renal biopsy is contraindicated.

In most cases, light microscopy shows uniform thickening of the glomerular capillary walls without endocapillary cell proliferation (Fig. 68.5). Characteristic subepithelial (epimembranous) and/or intramembranous deposits are seen on electron microscopy (Fig. 68.6). Identification of PLA_2R in glomerular immune deposits (by immunofluorescence or immunohistochemistry) favors the diagnosis of primary MN; mesangial deposits are often present in secondary MN. But a diagnosis of secondary MN depends on concurrent abnormalities in clinical and laboratory data.

Natural History

The clinical course of primary MN is highly variable. On average, about one-quarter of adult patients progress to ESKD within 10 years. Another quarter experiences a spontaneous remission of proteinuria. The majority is likely to have persistent proteinuria and moderately impaired renal function over an extended period of observation.

Baseline characteristics, such as severe nephrotic syndrome, hypertension, and azotemia, have been associated with poor outcomes. Protracted high-grade nephrotic-range proteinuria is a relatively strong predictor of an adverse renal outcome. Some studies suggest that the titer of autoantibody to PLA_2R at baseline correlates with disease activity such that higher titers may be associated with greater risk of renal function decline, whereas low titers may be associated with greater probability of spontaneous remission.[15]

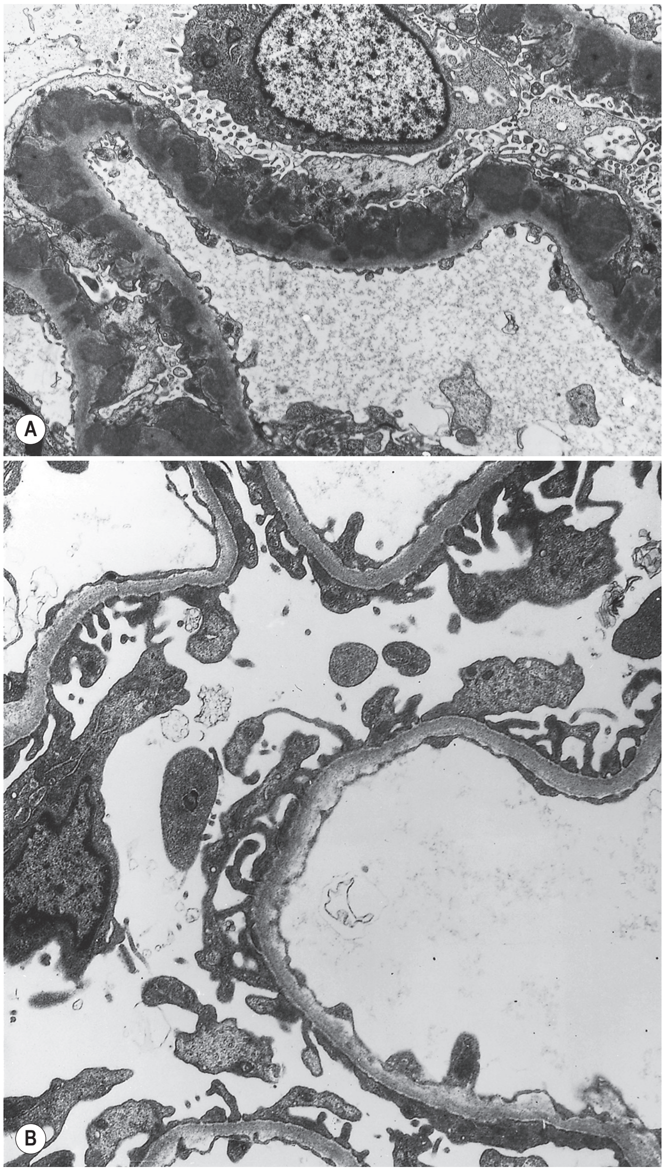


FIG 68.6 **Membranous Nephropathy (Ultrastructure).** A, Electron micrograph demonstrates heavy, dark-staining immune complex deposits along the outer surface of the glomerular basement membrane and beneath the epithelial foot processes (hence the terms *subepithelial* or *epimembranous deposits*). Note the thickening and projections of the gray-staining basement membrane between the electron-dense deposits. B, Ultrastructure of a normal glomerular capillary wall for comparison.

Treatment

Management of patients with MN usually includes diuretics to reduce edema, lipid-lowering drugs (for severe hyperlipidemia), anticoagulant therapy for thromboembolic complications, and

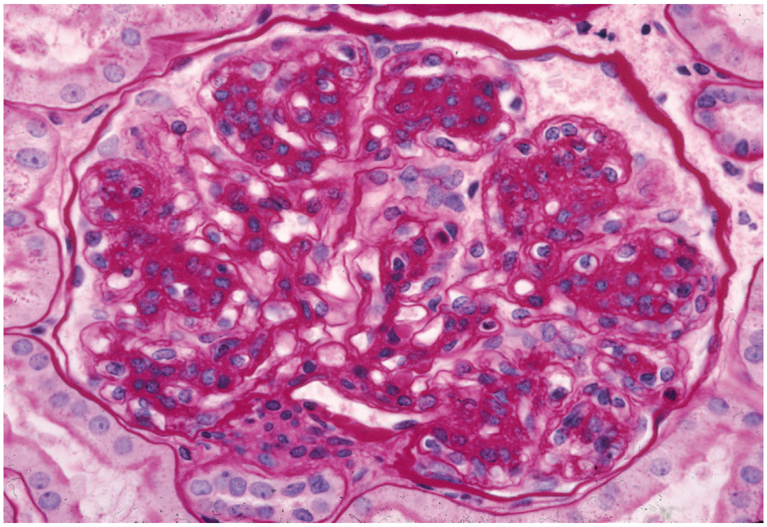


FIG 68.7 **Membranoproliferative Glomerulonephritis (MPGN).** Glomerulus exhibits the typical lobulated appearance of this disease. Markedly increased mesangial cells and matrix in all of the lobules. Mesangium extends outward into the capillary loops and forms double contours with the glomerular basement membrane. (Periodic acid–Schiff [PAS] stain).

antihypertensive agents. Angiotensin antagonists have been shown to have a substantial antiproteinuric effect. Immunosuppressive treatment is generally reserved for patients with persistent high-grade proteinuria (>4 g/day. First-line immunosuppressive therapy consists of cytotoxic drugs (usually cyclophosphamide) in combination with glucocorticoids or a calcineurin inhibitor (CSA or tacrolimus) with or without low-dose glucocorticoids.[16] The most compelling results from controlled trials have shown that patients with MN treated with alternating monthly courses of pulse methylprednisolone and chlorambucil or cyclophosphamide were more likely to experience a remission of the nephrotic syndrome and achieve stable renal function compared with controls. Rituximab may be a useful alternative for treatment of MN based on encouraging results from pilot studies and a good safety profile compared with other immunosuppressive regimens.[16–19]

MEMBRANOPROLIFERATIVE GLOMERULONEPHRITIS

KEY CONCEPTS

Membranoproliferative Glomerulonephritis (MPGN)

- Histologically classified into immune complex–mediated glomerulonephritis or complement-mediated glomerulonephritis based on immunofluorescence staining pattern
- C3 glomerulopathy characterized by C3 accumulation in the glomeruli in the form of electron-dense deposits.
- Genetic or acquired abnormalities in the activation of the alternate alternative complement pathway complement pathway associated with C3 glomerulopathy
- Response to immunosuppressive drug treatment generally poor
- Tends to recur in renal allografts

Membranoproliferative glomerulonephritis (MPGN) is a morphological entity that encompasses a heterogeneous group of diseases with a similar appearance at the light microscopic level. The pattern of injury is characterized by mesangial matrix expansion and increased cellularity and thickening of the glomerular capillary walls with a double contour appearance. These changes give a lobulated appearance to the glomerulus (Fig. 68.7).

Etiology and Pathogenesis

Until recently MPGN was classified into three types based on the location and type of electron dense deposits: type 1 characterized by subendothelial deposits; type II by intramembranous electron dense deposits in a ribbon-like pattern (also referred to as *dense deposit disease* [DDD]); and type III by subendothelial and subepithelial deposits (Fig. 68.8). This older classification scheme also made a distinction between secondary causes when it was possible to identify an etiology

New insight into the pathogenesis of MPGN has recently led to a major paradigm shift in the classification of this disease. Of great importance is the recognition that some cases of MPGN result from the deposition of Igs with secondary complement activation, whereas others arise from primary abnormalities in the control of complement activation.[20] The new classification system now broadly categorizes MPGN into Ig-mediated or complement-mediated disease based on the pattern and composition of the deposits, as assessed by immunofluorescence staining. The presence of both Ig and complement (C3) indicates an Ig-mediated process in which immune complexes are deposited and then secondarily activate the classical complement pathway. In contrast, the presence of complement staining (C3) without significant Ig deposition implies that an antibody-*in*dependent means of complement activation has been triggered and suggests a primary problem with regulation of the alternative complement pathway (Fig. 68.9). Such diseases are now grouped under the umbrella term, "C3 glomerulopathy" (C3G), which encompasses both C3 glomerulonephritis (C3GN) and DDD.[21] A diagnosis of C3G should prompt an evaluation for inherited mutations of complement regulatory proteins (*i.e.,* factor H or factor I) or acquired autoantibodies to regulatory proteins (*i.e.,* C3 nephritic factor or anti-factor H), which lead to dysregulation of the alternative pathway. C3 nephritic factor, an autoantibody to C3

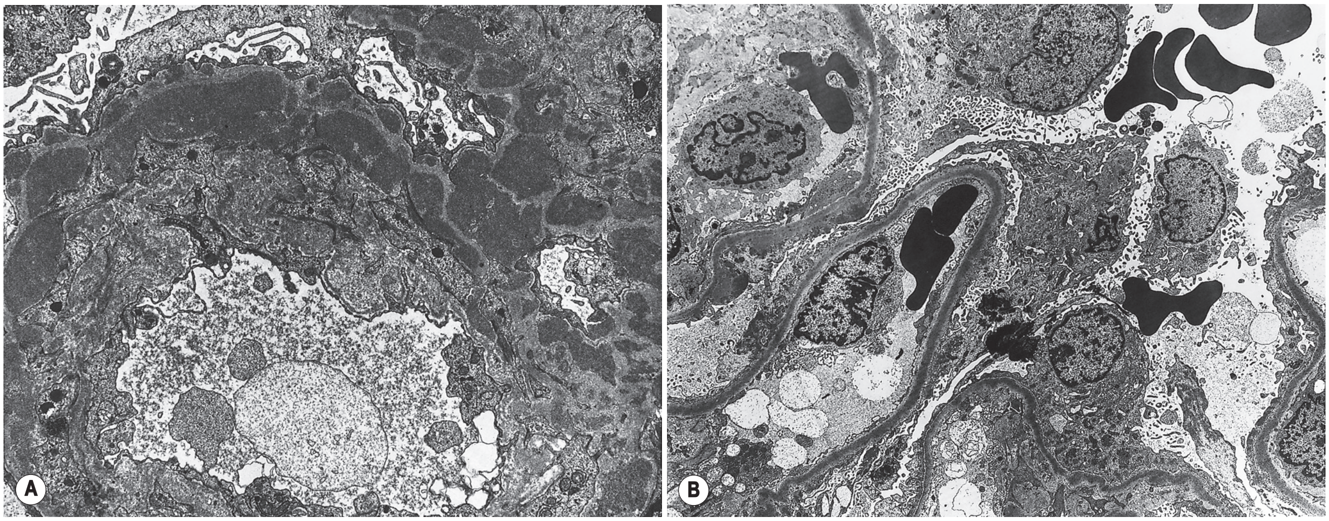


FIG 68.8 **Membranoproliferative Glomerulonephritis (MPGN) Ultrastructure.** A, Capillary wall is markedly thickened and contains heavy, dark-staining, electron-dense immune complexes in the subendothelial space. Mesangium (lighter material) extends into the capillary loop, where it is interposed between the basement membrane and the endothelium; the process gives the appearance of a massively thickened capillary loop on hematoxylin and eosin (H&E) staining, and the split appearance by periodic acid–Schiff (PAS) and silver stains. B, Dense deposit disease; capillary loops contain smooth, continuous, linear dense material within the basement membrane.

convertase, is frequently detected in C3G. C3 nephritic factor stabilizes the convertase, rendering it resistant to control by factor H, thus leading to persistent C3 activation and deposition of alternative pathway activation products in glomeruli.

Pathology

In addition to endocapillary proliferation and findings on immunofluorescence, a double contour appearance to the glomerular basement membrane (GBM) best seen with silver stain represents synthesis of GBM-like material from capillary wall remodeling. Irregular capillary wall subendothelial and mesangial deposits are seen by electron microscopy in Ig- mediated MPGN; often small intramembranous and subepithelial deposits are also present in C3GN. The hallmark of C3GN is dominant bright staining for C3 by immunofluorescence in the mesangium and glomerular capillary walls (see Fig. 68.9).

Double contours of the GBM are also present in conditions associated with chronic endothelial injury, including thrombotic microangiopathy, transplant glomerulopathy, and preeclampsia, and show a histological appearance of MPGN by light microscopy. However, there are no associated immune deposits and immunofluorescence is negative for Ig and C3.

Clinical Presentation

Although it is characteristically a chronic low-grade nephropathy, some patients experience nephrotic syndrome and even rapidly progressive and crescentic disease. Hypocomplementemia is common in all types of MPGN. In immune complex-mediated MPGN, serum C4 levels are characteristically low, consistent with classical pathway activation. Serum C3 concentration may be normal or mildly decreased. In complement-mediated MPGN, serum C4 levels are normal, and C3 is typically low, consistent with alternative pathway activation, but a normal serum C3 does not exclude the diagnosis. The renal outcome in C3 glomerulopathy is variable with up to 30% of patients progressing to ESKD. The prognosis of DDD is worse, with almost half the patients progressing to kidney failure. Idiopathic MPGN and MPGN resulting from complement dysregulation tend to recur in renal allografts; recurrences of both types of MPGN have a detrimental effect on graft survival.

Treatment

It is important to recognize that existing treatment trials for MPGN predate the new classification system and our appreciation of the different pathogenesis. Thus their relevance is questionable because of the heterogeneity of patient populations. It is anticipated that the new diagnostic categories will better define groups of patients with similar pathogenesis, thus enabling a more rational approach to targeted therapy in the future.

There is no consensus regarding treatment of "idiopathic" immune complex–mediated MPGN, but various immunosuppressants have been used with variable success, including prednisone alone or in combination with either mycophenolate or cyclophosphamide. There are limited data on the role of rituximab in this disease.

Optimal treatment of C3GN remains unclear but will likely need to be individualized on the basis of proper identification of the defect in the alternative complement system (*i.e.*, genetic vs acquired autoantibodies). Chronic infusions of fresh frozen plasma to replace missing complement factors may be beneficial in some cases. Immunosuppression with corticosteroids, rituximab, and mycophenolate is of theoretic benefit in cases resulting from pathogenic antibodies to complement regulatory proteins but is unproven. A more targeted approach using eculizumab, a mAb that blocks activation of C5 complement (C5b-9), has been tested in small series, but the data are inconclusive. Additional new complement blocking agents may also hold promise.[22]

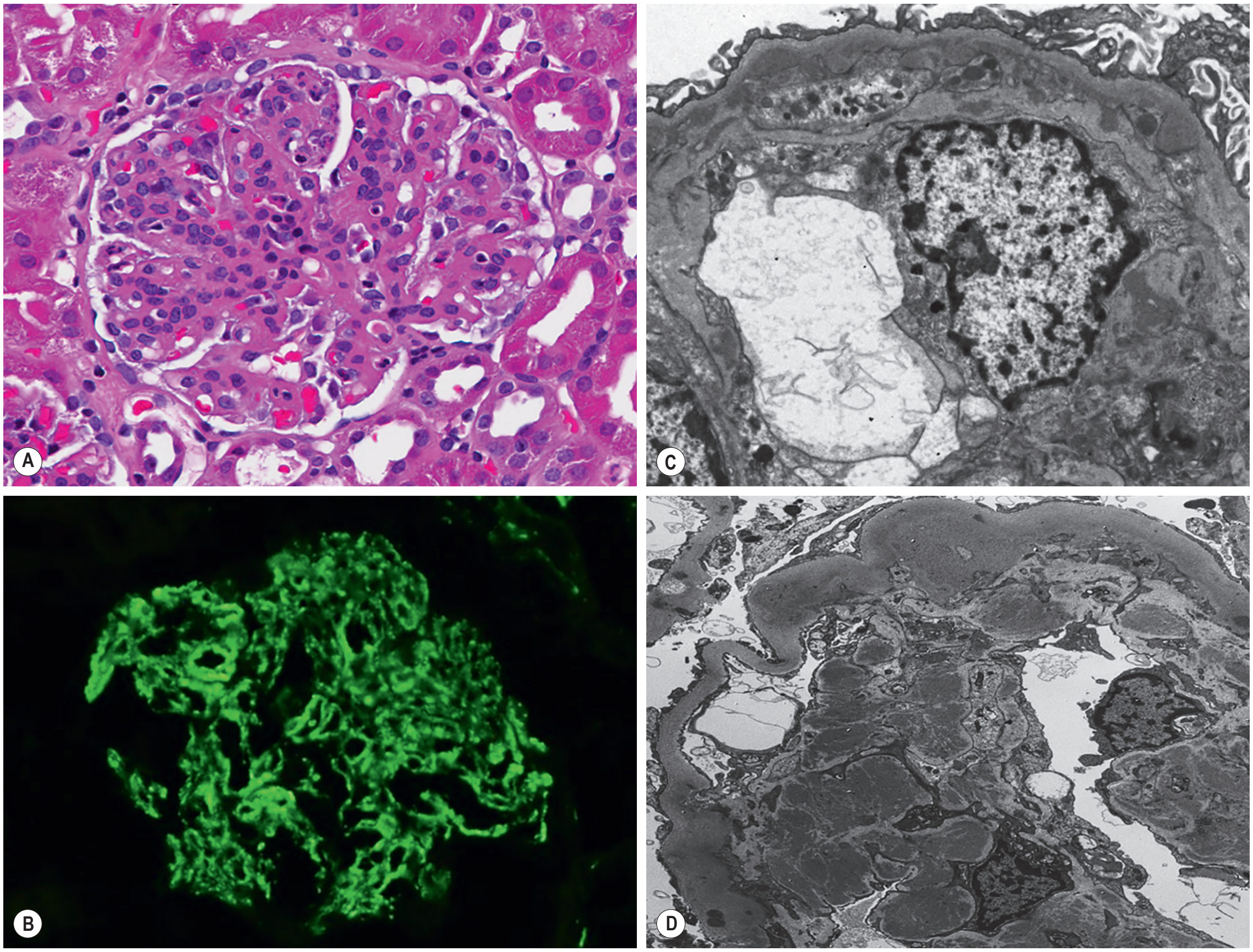


FIG 68.9 **C3 Glomerulopathy.** A, Glomerulus exhibiting the characteristic membranoproliferative pattern (hematoxylin and eosin [H&E] stain). B, Immunofluorescence showing prominent glomerular staining for C3 complement. C and D, Electron microscopy ultrastructure showing mesangial interposition giving a double contour of the glomerular capillary wall, as well as interposed electron dense deposits in the subendothelial space and the mesangium.

POSTINFECTIOUS NEPHROPATHIES

An ever-expanding list of infectious agents has been implicated in the pathogenesis of certain forms of immune-mediated nephropathies.

Viral Infections

Hepatitis B

Hepatitis B–associated nephropathy most frequently presents as nephrotic syndrome accompanied by microscopic hematuria; renal biopsy most commonly shows MN. Immunofluorescence studies usually show Ig, C3, and some IgM staining in the subepithelial region along the glomerular basement membranes and in some cases viral antigens can be detected within the glomerulus. Electron microscopy shows subepithelial and intramembranous deposits, but there may also be mesangial and even some subendothelial deposits.

Therapy for hepatitis B renal disease has focused on antiviral drugs because glucocorticoids and cytotoxic agents may promote viral replication. Treatment generally consists of interferon-α (IFN-α) or nucleoside/nucleotide analogues, such as lamividine, entacavir, adefovir, tenofovir, and telbivudine. Treatment with nucleoside/nucleotide analogues are generally used for at least 1 year and continued for at least 6 months after HBeAg seroconversion. Immunosuppression (corticosteroids with or without cyclophosphamide or rituximab) may cautiously be considered in patients with rapidly progressive glomerulonephritis, but concomitant antiviral therapy is required.

Hepatitis C

Hepatitis C virus (HCV) is a major cause of both transfusion-associated and sporadic chronic hepatitis. The most common HCV-associated renal disease is usually, but not invariably, MPGN in the context of type II mixed cryoglobulinemia. The pathogenesis of the renal disease is caused by deposition within glomeruli of immune complexes containing HCV, anti-HCV antibody and virus-related (or unrelated) cryoglobulins.

Treatment of HCV infection has greatly improved in the last few years. Historically, treatment was mainly based on IFN-α and ribavirin. These agents are now being replaced by HCV-specific antiviral drugs that are used in combination (*i.e.,* sofosbuvir plus ledipasvir), which are highly effective and often curative.[23] The evidence for antiviral treatment in HCV-related kidney disease is based mostly on IFN-based regimens, which have reported remission of proteinuria and hematuria and improvement of renal function.[24,25] There are limited data regarding use of the newer antiviral agents in HCV-associated glomerulonephritis, but these hold considerable promise.[26] For patients with progressive renal disease and/or nephrotic-range proteinuria, treatment with rituximab or pulse intravenous steroids and cyclophosphamide may be warranted in conjunction with antiviral therapy. Treatment with rituximab and plasma exchange may provide additional benefit in patients with HCV-associated severe cryoglobulinemia refractory to antiviral therapy.[27,28]

Human Immunodeficiency Virus

HIV-associated nephropathy (HIVAN), the first kidney disease to be associated with HIV infection, is a collapsing form of focal segmental glomerulosclerosis accompanied by associated tubular microcysts and interstitial inflammation. It was first described in patients with acquired immune deficiency syndrome (AIDS) but also is occasionally diagnosed in less advanced cases of HIV infection and before acute HIV seroconversion has been identified. HIVAN classically presents with significant proteinuria and rapidly progressive kidney disease. HIVAN displays a striking racial predilection for black patients, which may indicate the importance of genetic influences.[29]

A spectrum of other renal abnormalities have also been described in patients with HIV infection, including IgA nephropathy, lupus-like glomerulonephritis, postinfectious glomerulonephritis, MPGN, MN, cryoglobulinemic glomerulonephritis, fibrillary and immunotactoid glomerulopathy, and thrombotic microangiopathy.[30,31] Tubulointerstitial changes related to drug toxicity, acute interstitial nephritis, or superimposed viral, fungal, or mycobacterial infections may also be present.

The mainstay of treatment of HIVAN is combination antiretroviral therapy (cART) regardless of the CD4 lymphocyte count. Highly effective modern therapies for HIV have reduced the frequency of developing HIVAN and have greatly improved the otherwise dismal renal prognosis of HIVAN. This has greatly improved the dismal renal prognosis of HIVAN. The evidence for initiating cART in other HIV-associated immune complex glomerular diseases is inconclusive, but use of cART is a rational approach. The use of standard immunosuppressive drugs in an immunocompromised population is controversial.

Bacterial Infections

Poststreptococcal Glomerulonephritis

Poststreptococcal glomerulonephritis is the best-studied and classic immune complex–mediated glomerulonephritis. It is the result of skin or throat infection with nephritogenic strains of group A streptococci. The latent period between upper respiratory infection and nephritis is 7–10 days, and 2–4 weeks after skin infection. Antistreptococcal antibody titers are usually measured to demonstrate the existence of a preceding streptococcal infection. Antistreptolysin O titers and anti-DNase B titers are the most frequently elevated in upper respiratory and skin infections, respectively.

Poststreptococcal glomerulonephritis is characterized by a nephritic syndrome consisting of smoky or rust-colored urine, generalized edema, hypertension, and nephritic urine sediment. Proteinuria is typically mild. Patients have rising titers of antistreptolysin and depressed C3 levels early in nephritis but normal or minimally depressed C4 levels, indicating activation of the alternative complement pathway.

KEY CONCEPTS

Infection-Related Nephropathies

- *Viral:* Hepatitis B—membranous nephropathy; hepatitis C—cryoglobulinemic membranoproliferative glomerulonephritis; human immunodeficiency virus (HIV)—focal segmental glomerulosclerosis (HIV nephropathy)
- *Bacterial (mainly gram-positive):* Nephritogenic streptococcal infections, prosthetic device (shunt) infections, subacute bacterial endocarditis, chronic deep tissue abscesses—mainly diffuse or focal proliferative glomerulonephritis

Proliferative glomerulonephritis with polymorphonuclear leukocyte and monocyte infiltration, granular immune deposits of IgG and C3, and dome-shaped electron-dense subepithelial deposits (humps) are characteristic. The prognosis is excellent in children; with supportive care, almost all will recover. Progressive renal failure accompanied by severe hypertension appears to be more common in adults. Kidney biopsy is rarely needed in the child but may be warranted if there is an atypical presentation or evolution. .

The classic childhood form is still seen in developing countries but is now rare in developed countries. However, there has been an increase in the incidence of nonstreptococcal postinfectious glomerulonephritis or "infection-related" glomerulonephritis, which tends to be seen in older patients with multiple comorbidities, especially diabetes, HIV infection, and malignancy. These clinical variants are usually related to other infective agents, such as *Staphylococcus aureus,* both methicillin-resistant (MRSA) and methicillin-sensitive *S. aureus,* and may be characterized by an IgA-dominant glomerular immune complex deposition.

Etiology and pathogenesis. Glomerular injury results from passive glomerular trapping of circulating immune complexes composed of nephritogenic bacterial antigens and IgG antibody or by the *in situ* formation of immune complexes. This is followed by immune cell recruitment, production of chemical mediators and cytokines, and local activation of the complement and coagulation cascades that drive an inflammatory response.[1]

Current therapeutic strategies rely on culture-guided systemic antibiotics, especially in older patients, in whom MRSA may be the causative agent. Steroids may be used in selected cases in which crescents and severe interstitial inflammation are present.

IgA Nephropathy

KEY CONCEPTS

Immunoglobulin A Nephritis (IgAN)

- Common cause of asymptomatic microscopic hematuria, recurrent macroscopic hematuria, and/or low-grade proteinuria
- Spectrum of disease, including idiopathic IgA nephritis and Henoch-Schönlein purpura nephritis; IgA in skin and renal biopsy samples
- Mostly benign prognosis, especially in children
- Patients with progressive renal insufficiency and/or crescentic glomerulonephritis warrant trial of glucocorticoids and/or cytotoxic drug therapy

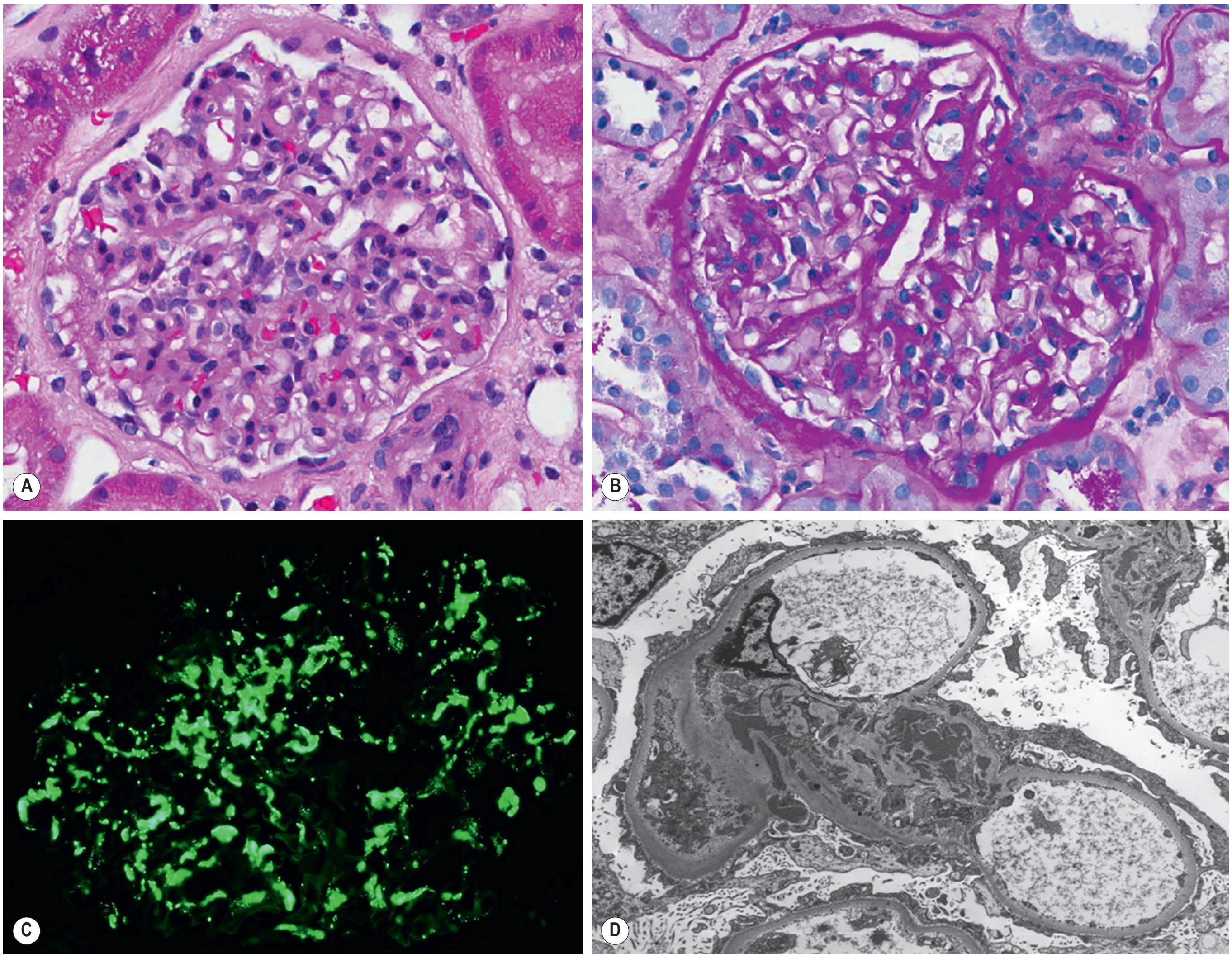


FIG 68.10 **Immunoglobulin A Nephropathy (IgAN).** (A) hematoxylin and eosin (H&E) stain and (B) periodic acid–Schiff (PAS) stain: Glomeruli manifest modestly increase mesangial hypercellularity and expanded mesangial matrix. C, Immunofluorescence showing predominantly IgA deposition within the mesangium. D, Electron microscopy showing deposits mainly within the mesangium but not along peripheral capillary loops.

IgA nephropathy (IgAN) is the most common primary glomerulonephritis worldwide.[32] IgAN can affect patients of all ages, especially children and young adults, with a male preponderance.[34] There are geographical and ethnic differences in the prevalence of IgAN. The highest frequency is found among East Asians; IgAN is very uncommon in individuals of African ancestry. This observation and examples of familial clustering of IgA nephropathy favor an important element of genetic susceptibility.

IgAN may be discovered during evaluation of asymptomatic microscopic hematuria.[34] Alternatively, patients (especially children) can present with recurrent episodes of macroscopic hematuria that occur within 24–48 hours after an intercurrent infection, usually an upper respiratory or gastrointestinal (GI) tract infection. A transient elevation in serum creatinine has been associated with macroscopic hematuria in about one-third of cases. This has been attributed to tubular injury and obstruction caused by intraluminal RBC casts. A small percentage of patients present with either nephrotic syndrome or an acute RPGN picture characterized by edema, hypertension, renal insufficiency, and hematuria.

Pathology

The characteristic findings on light microscopy are mesangial cell proliferation and mesangial matrix expansion.[32] Electron microscopy typically reveals electron-dense deposits that are primarily limited to the mesangium, but a few deposits may also be present in subendothelial and subepithelial locations. The pathognomonic finding on immunofluorescence microscopy is globular deposits of IgA (often accompanied by C3 and IgG) in the mesangium and, to a lesser degree, along the glomerular capillary wall (Fig. 68.10).

Etiology and Pathogenesis

Aberrant glycosylation of *O*-linked glycans in the hinge region of IgA1 resulting in increased serum levels of galactose-deficient IgA1 (Gd-IgA1) plays a pivotal role in pathogenesis of IgAN.[34] The aberrantly glycosylated IgA1 is recognized by antiglycan antibodies and leads to circulating IgA immune complexes that preferentially deposit in the mesangium, provoking local injury.

A genetic predisposition to IgAN has been linked with polymorphisms involving innate and adaptive immunities and the alternative complement pathway. A "second hit" may be needed in predisposed individuals. Infections may play a role because episodes of macroscopic bleeding often coincide with mucosal infections, including upper respiratory tract (synpharyngitic) or GI infections.

Natural History

Patients with IgAN who have low-grade proteinuria (<1 g/day) have a good renal prognosis and usually have a low risk of progression. However, it is recognized that at least one-third of patients with IgAN eventually progress to ESKD.[34] Twenty years after apparent disease onset, the probability of renal failure is 25%, and the probability of some renal dysfunction is 50%. Hypertension occurs frequently as the disease progresses and forebodes a poor prognosis. Other clinical features that have been associated with a poor prognosis include older age at disease onset, persistent proteinuria (>1 g/day), and persistent azotemia.

Treatment

The most effective treatment of progressive IgAN remains undefined and controversial.[34] Angiotensin antagonists are recommended to achieve blood pressure control, to reduce proteinuria, and to slow the rate of deterioration of renal function. There are conflicting data about the value of fish oil dietary supplements (eicosanoids) in preventing renal progression in patients with IgAN. The practice of tonsillectomy in IgAN has not been confirmed. A large trial (STOP-IgA nephropathy) showed that immunosuppression (glucocorticoid monotherapy or a regimen that included prednisolone, cyclophosphamide, and azathioprine) did not have a significant beneficial effect on preservation of kidney function.[33] However, a course of oral steroids may be considered in patients with high-grade proteinuria, and cytotoxic drugs are indicated in a small subset of patients with crescentic rapidly progressive IgAN. The lack of effective therapy have provided the impetus for several new phase II/III clinical trials testing novel therapies, such as B-cell activating factor (BAFF) inhibition, proteasome inhibition, and B-cell inhibition.[34,35]

Renal Vasculitis Associated With Antineutrophil Cytoplasmic Antibodies

KEY CONCEPTS

Antineutrophil Cytoplasmic Antibodies (ANCAs)–Associated Renal Vasculitis

- Renal vasculitis with glomerular involvement includes microscopic polyangiitis (MPA), granulomatosis with polyangiitis (GPA), and necrotizing crescentic glomerulonephritis (renal-limited vasculitis)
- Associated with ANCAs
- Rapidly progressive glomerulonephritis is common; early treatment includes pulse methylprednisolone, cyclophosphamide, rituximab, or possibly plasma exchange
- Maintenance therapy: azathioprine, rituximab

ANCAs are associated with a distinct form of vasculitis that can affect many different types of vessels and any organ in the body. ANCA-associated vasculitis (AAV; Chapter 58) is categorized into four main types: microscopic polyangiitis (MPA), granulomatosis with polyangiitis (GPA, formerly called *Wegener granulomatosis*), eosinophilic granulomatosis with polyangiitis (EGPA, formerly Churg-Strauss syndrome), and renal limited vasculitis. As shown in Table 68.2, the patterns of ANCAs differ, depending on the type of vasculitis.

TABLE 68.2 Prevalence of Antineutrophil Cytoplasmic Antibodies (ANCAs) in Renal Vasculitis

	ANCAS TEST POSITIVITY (%)	
Type of Renal Vasculitis	P-ANCAs or Anti-Mpo	C-ANCAs or Anti-Pr3
Polyarteritis nodosa	10–20	10–20
Microscopic polyangiitis	50–80	10–20
Granulomatosis with polyangiitis (Wegener granulomatosis)	10–20	80–90
Necrotizing and crescentic GN	50–80	10–20

MPO, myeloperoxidase; PR3, proteinase 3.

Renal involvement is common in AAV but varies in type and severity. It occurs more frequently in MPA (90%) and in GPA (80%) and less frequently in EGPA (45%). Clinically RPGN is a common manifestation of these renal vasculitides characterized by the presence of hematuria, proteinuria, active urinary sediments, and renal failure.

Pathology

The glomerular abnormalities are similar among the subtypes of ANCA-associated glomerulonephritides (Fig. 68.11). The glomerular lesions are characteristically focal and segmental in distribution, with fibrinoid necrosis and crescent formation. Breaks in the GBM may be seen. There may be accompanying necrotizing arteritis. In contrast to immune complex–mediated vasculitis, ANCA-associated vasculitis has little or no Ig deposition in injured glomerular vessels, with minimal or negative staining by immunofluorescence, a so-called pauci-immune pattern. Patients with large percentage of cellular crescents (>50%) typically present with severely reduced renal function but have a good chance for recovery of renal function with treatment, whereas those with a greater percentage of globally sclerotic glomeruli are less likely to recover renal function.[36]

Treatment and Prognosis

ANCA-associated renal vasculitis tends to be severe and fulminant, making early detection critically important in management. Even with early diagnosis, approximately one-third of patients will progress to renal failure within 5 years. Relapsing courses are common in patients with microscopic polyangiitis, particularly GPA. This underscores the importance of prompt induction treatment followed by maintenance therapy. Glucocorticoids are important in the early treatment of renal vasculitis. Cyclophosphamide-based regimens are very effective in inducing remission in patients with AAVs. Most advocate daily oral cyclophosphamide regimens, but intermittent pulse cyclophosphamide may be substituted to reduce the toxicity of extended therapy. In patients with severe pulmonary hemorrhage or rapidly progressive glomerulonephritis caused by renal vasculitis, pulse methylprednisolone, followed by prednisone and daily cyclophosphamide, is clearly indicated. Adjunctive plasma exchange is commonly used in cases of aggressive pulmonary–renal syndrome. Rituximab has become a valuable alternative to

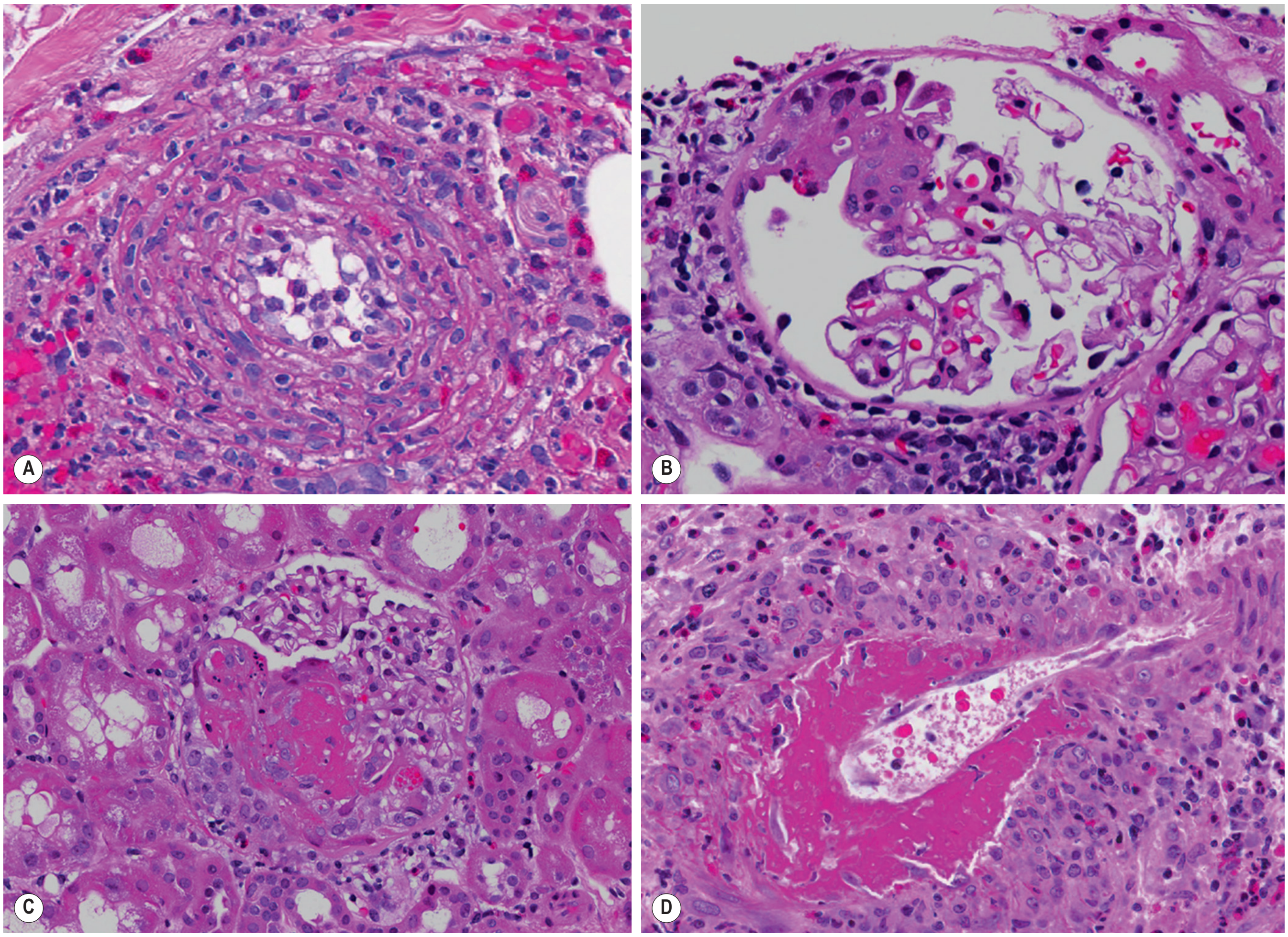


FIG 68.11 Antineutrophil Cytoplasmic Antibody (ANCA)–Associated Systemic Vasculitis. A, Skin biopsy showing leukocytoclastic vasculitis of a small dermal artery (hematoxylin and eosin [H&E] stain). B, Glomerulus illustrating a characteristic segmental proliferative lesion forming an adhesion to the Bowman capsule in an early and mild case (H&E stain). C, Glomerulus with classical fibrinoid necrosis and an associated cellular crescent of a more severe and rapidly progressive case (H&E stain). D, Necrotizing vasculitis in a small renal artery in an extreme case of ANCA-associated renal vasculitis; large numbers of eosinophils are present in the perivascular inflammatory infiltrate (H&E stain).

cyclophosphamide for induction treatment based on the results of two pivotal prospective randomized controlled trials.[37,38] Maintenance therapy is usually given for 12–18 months after achieving remission. Azathioprine has been the mainstay of maintenance therapy in AAVs after remission induction, but rituximab is also likely to have a prominent role in maintenance therapy, as indicated by favorable results from preliminary trials.[39]

Anti-GBM Antibody-Mediated Nephritis: Goodpasture Disease

KEY CONCEPTS

Goodpasture Disease

- Circulating anti–glomerular basement membrane (GBM) antibody
- Associated with pulmonary hemorrhage
- Rapidly progressive glomerulonephritis with cellular crescents and linear deposits of immunoglobulin G (IgG)
- Treated with high-dose steroids, cyclophosphamide, plasma exchange

Goodpasture disease is a rare but classic immune-mediated cause of severe pulmonary–renal syndrome. The cardinal pathogenic factor is an autoantibody to a component of type IV collagen (noncollagenous domain 1 of the α_3 chain subunit) within the GBM or the alveolar basal membrane.[40,41] These autoantibodies lead to GBM rupture with development of crescentic glomerulonephritis (Fig. 68.12) and account for the pathognomonic finding of linear deposition of IgG along the glomerular capillaries as assessed by immunofluorescence microscopy.

The genesis of anti-GBM antibodies in sporadic cases is unknown. It is postulated that an inciting event(s) *(i.e.*, viral infection or exposure to environmental factor, such as hydrocarbons, tobacco) perturbs the normal conformation of collagen in the basement membrane, exposing previously cryptic epitopes on the α_3 subunit, eliciting the pathogenic autoantibody response. Iatrogenic cases have occurred when normal kidneys have been transplanted into patients with hereditary Alport nephropathy (which is caused by mutations in the α_5 chain of type IV collagen). The absence of expression of α_5 (and associated lack of α_3 and

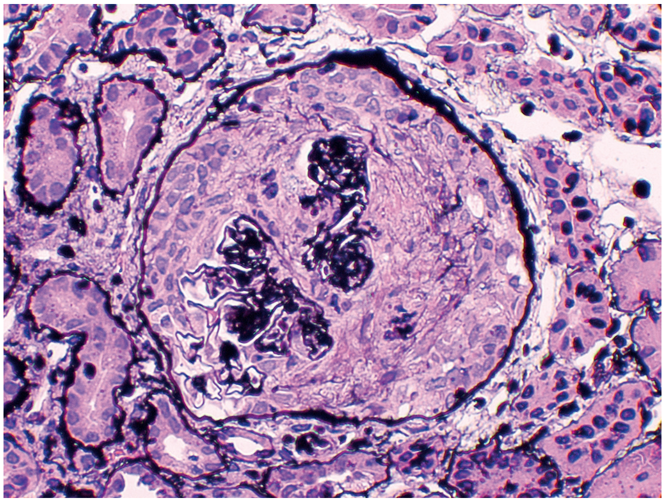


FIG 68.12 Goodpasture Disease. Circumferential cellular crescent fills the Bowman capsule and compresses the glomerular tuft (Silver stain).

α_4) means that these molecules within the renal allograft are regarded as foreign antigens. Thus recipients mount an anti-GBM antibody response to the new donor antigens.

Clinically malaise, weight loss, fever, or arthralgia may be the initial features of anti-GBM disease. Some patients present with isolated renal involvement, but more typically, pulmonary hemorrhage with hemoptysis accompanies acute renal failure. The devastating nature of the disease warrants aggressive treatment. Pulse methylprednisolone, cyclophosphamide, and plasma exchange are indicated early in the course of Goodpasture disease.[42] The role of rituximab in this disease requires further study. Reversibility of renal disease is unlikely if renal function is substantially impaired or oliguria ensues before treatment is started. Immunosuppressive treatment is normally continued until the patient has been in sustained clinical remission and anti-GBM titers are minimal or absent for at least 3 months.

Lupus Nephritis

KEY CONCEPTS

Lupus Nephritis

- *Class I:* normal glomeruli by light microscopy in patients with SLE.
- *Class II, mesangial:* Immunosuppressive treatment is usually not indicated unless patient has proteinuria >1 g/day despite renin–angiotensin–aldosterone (RAA) system blockade, especially if urine sediment is nephritic
- *Class III, focal nephritis and class IV, diffuse nephritis:* Mycophenolate mofetil (MMF) or intravenous cyclophosphamide (IVCY) plus glucocorticoids for induction followed by MMF or azathioprine plus low-dose corticosteroids as maintenance therapy
- *Class V, Membranous nephropathy:* Alternate-day prednisone with MMF, or alternatively bimonthly pulse cyclophosphamide or low-dose daily cyclosporine

Glomerular disease affects the majority of patients with SLE (Chapter 51), but lupus nephritis has a wide spectrum of disease expression and outcomes among different patient populations. Lupus nephritis occurs more often and is associated with less favorable outcomes among Hispanic, Asian, Native American, and especially African American populations compared with Caucasians. Nephritis is a major cause of morbidity and mortality and accounts for a large portion of all hospital admissions in patients with SLE.

Pathogenesis

Several different mechanisms appear to be involved in the pathogenesis of lupus nephritis, resulting in a wide spectrum of renal lesions. Deposition of immune complexes from the circulation into the kidney appears to be the initiating event in proliferative lupus nephritis; however, only a subset of immune complexes appears to be nephritogenic. DNA and anti-DNA antibodies are known to be concentrated in glomerular deposits in the subendothelial location and are likely to play a central role in the pathogenesis of proliferative lupus nephritis. Unfortunately, there are fewer insights into the pathogenesis of lupus MN with its characteristic epimembranous immune deposits. Although T cells are almost certainly involved in autoantibody production, it is unknown whether they have a direct role in the pathogenesis of lupus nephritis.

Clinical Features

Asymptomatic hematuria or proteinuria may be the presenting features, but they often progress to nephritic and/or nephrotic syndrome. Hypertension, azotemia, nephritic urine sediment (with hematuria and cellular casts), hypocomplementemia and high anti–double-stranded DNA (dsDNA) titers are more commonly found in patients with proliferative lupus nephritis. RPGN is usually associated with the appearance of cellular crescents and may be superimposed on severe proliferative or membranous forms of lupus nephritis.

Pathology

The former classification of renal biopsy in lupus nephritis by the World Health Organization (WHO) was revised by an international committee in 2004 (Figs. 68.13 to 68.15). A summary of the histological features in each class of lupus nephritis can be found in Table 68.3.

Treatment

In 2012, the American College of Rheumatology (ACR), the European League Against Rheumatism (EULAR), and the European Renal Association–European Dialysis and Transplantation Association (ERA-EDTA) published their recommendations for the management and treatment of patients with lupus nephritis.[43,44] Immunosuppressive treatment of mesangial classes of lupus nephritis (classes I and II) is usually not indicated (ACR). However, the distinction between early mesangial lesions that are in transition to more ominous classes from those that reflect mild and stable nephropathy is difficult. Consequently, treatment with prednisone alone or in combination with azathioprine has been recommended for patients with proteinuria that exceeds 1 g/day despite blockade of the renin–angiotensin–aldosterone (RAA) system, especially if the patient also has a nephritic urinary sediment (EULAR/ERA-EDTA). For patients with focal or diffuse proliferative glomerulonephritis (classes III and IV), ACR and EULAR/ERA-EDTA recommended mycophenolate mofetil (MMF) or intravenous cyclophosphamide (IVCY) plus glucocorticoids. Low-dose IVCY (500 mg IV every 2 weeks in six doses) offers a favorable balance of efficacy and relatively low toxicity for Caucasian patients with Western or Southern European ancestry. Higher dose, monthly IVCY for 6 months, plus 3 daily IV infusions of methylprednisolone initially, followed by

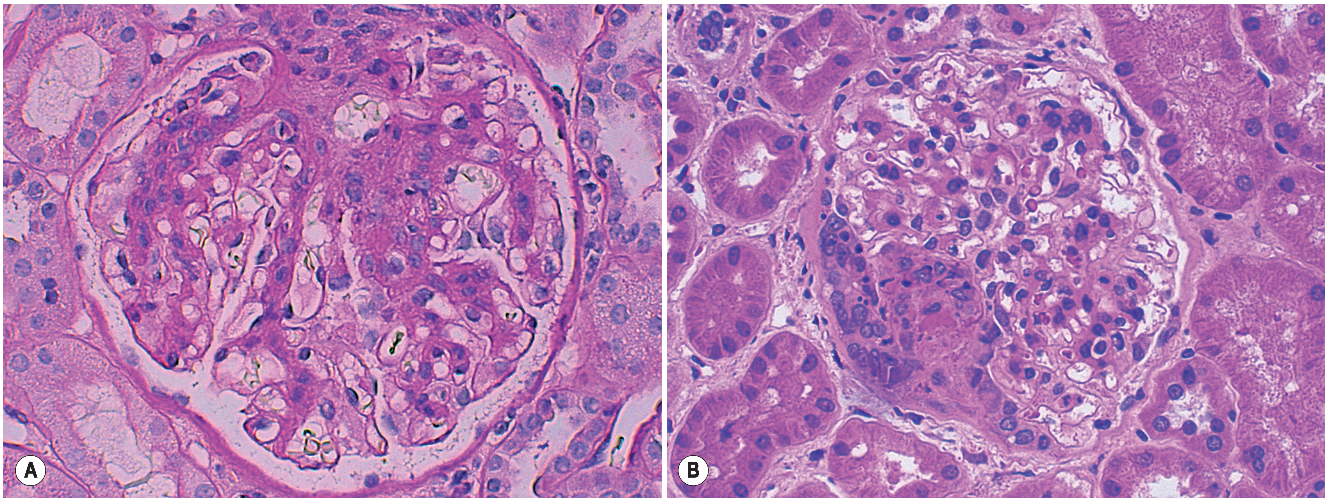


FIG 68.13 **Classes of the Pathology of Lupus Nephritis (1).** A, Class II, mesangial proliferative lupus nephritis; mesangial areas are expanded by cells and matrix but the peripheral capillary loops remain widely patent (periodic acid–Schiff [PAS] stain). B, Class III, focal lupus nephritis; solid lesion at the lower right portion of this glomerulus demonstrates segmental fibrinoid necrosis. Note the nuclear fragments (karyorrhexis) in the fibrinous exudate (hematoxylin and eosin [H&E] stain.)

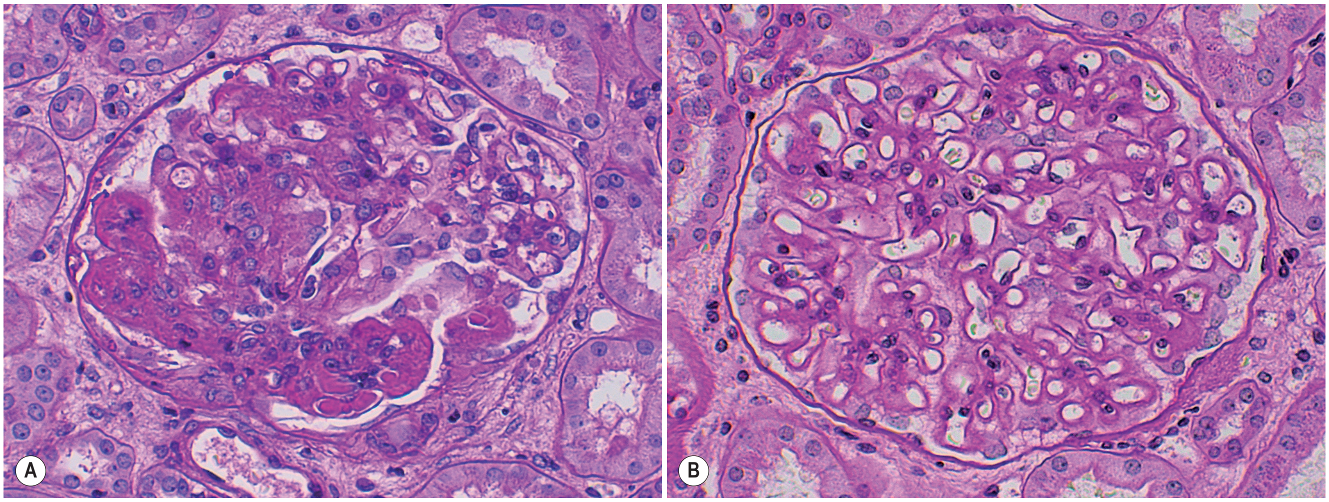


FIG 68.14 **Classes of the Pathology of Lupus Nephritis (2).** A, Class IV, diffuse lupus nephritis; glomerulus with irregular but nearly global changes, including obliteration of many capillary loops resulting from endocapillary hypercellularity, "wire loop" thickening and hyaline thrombi (periodic acid–Schiff [PAS] stain). B, Class V, membranous lupus nephritis; glomerulus shows minimally increased mesangial cellularity with thickened but widely patent capillary loops (PAS stain).

prednisone has also been recommended for patients with ominous clinical and histological prognostic indicators, including cellular crescents and fibrinoid necrosis. ACR and EULAR/ERA-EDTA recommended either azathioprine (AZA) or MMF as maintenance therapy for patients showing a favorable response after initial immunosuppressive therapy. Although AZA and MMF appeared to be equally effective maintenance therapies in a European study, patients randomized to maintenance therapy with MMF had more favorable outcomes than those randomized to AZA in a larger study conducted worldwide. The duration of maintenance immunosuppressive therapy involves careful consideration of the risks of another renal flare-up versus the risks of drug toxicity. EULAR/ERA-EDTA recommended at least 3 years of therapy in patients showing improvement after initial therapy. In general, we have offered a comparable recommendation that treatment should continue for at least 1 year after remission of renal disease to prevent exacerbations.

Neither IVCY nor MMF is universally effective in the management of lupus nephritis, hence the search for more efficacious treatment regimens, including rituximab, belimumab (binds to BAFF), immunomodulators (*e.g.*, laquinimod), cytokine inhibitors, immunoablation without or with stem-cell reconstitution, and immunological costimulation inhibitors (*e.g.*, CTLA-4-Ig), continues.[35,45] The importance of these ongoing efforts to

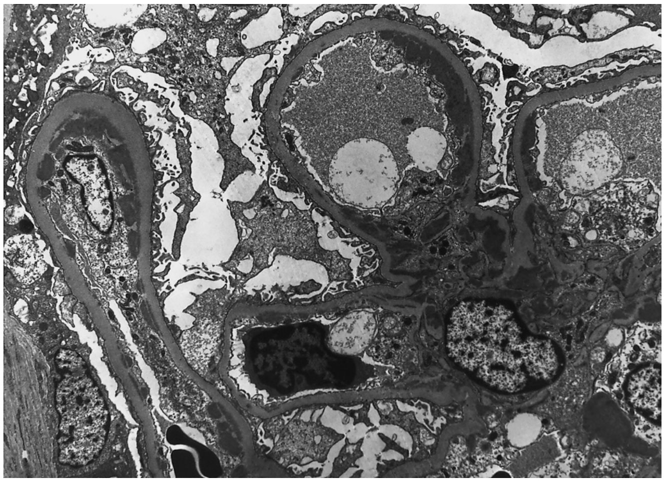


FIG 68.15 Ultrastructure of Proliferative Lupus Nephritis. Electron micrography demonstrates the characteristic mesangial deposits (dark materials interspersed within the centrally located amorphous, gray mesangial matrix) and subendothelial deposits (dark materials extending along the peripheral capillary loops).

TABLE 68.3 International Society of Nephrology/Renal Pathology Society 2004 Classification of Lupus Nephritis

Class	Histologic Features/Comments
I. Minimal mesangial	Normal light microscopy (LM); mesangial deposits by immunofluorescence (IF) and electron microscopy (EM)
II. Mesangial proliferative	Pure mesangial hypercellularity and matrix expansion *IF* and *EM:* mesangial immune deposits
III. Focal	Glomerular capillary obliteration in <50% of nephrons as a result of proliferation or sclerosis *LM:* Increased numbers of mesangial, endothelial, and/or hematogenous cells. Active inflammatory lesions (karyorrhexis, fibrinoid necrosis, adhesion to the Bowman capsule, cellular crescents, interstitial inflammatory infiltrates). Wire loop lesions. Hyaline thrombi *IF* and *EM:* Mesangial and peripheral capillary loop (subendothelial) immune complex deposits
IV. Diffuse	Qualitatively similar histologic lesions as in class III. Glomerular capillary obliteration involving >50% of nephrons. Subsets defined as primarily global (class IV-G) or primarily segmental (class IV-S) involvement
V. Membranous	*LM:* Regular thickening of the peripheral capillary loops of the glomerulus. Mesangial expansion *EM:* Subepithelial, intramembranous, mesangial (but no or very rare subendothelial) immune complex deposits
VI. Advanced	>90% global sclerosis without residual active lesions

investigate novel approaches to treatment is underscored by observations that although the risk of renal failure caused by lupus nephritis decreased from the 1970s to the mid-1990s (coincident with the increased use of cyclophosphamide), that risk has shown a reverse trend and increased slightly in the last two decades.[46]

For patients with "pure" class V lupus MN and nephrotic-range proteinuria, both the ACR and the EULAR/ERA-EDTA have recommended oral prednisone and MMF. A prospective controlled trial showed that both IVCY and CSA were more effective than steroids alone in inducing remission of proteinuria in lupus MN but that relapse of nephrotic syndrome occurred significantly more often in the CSA group than in the IVCY treatment group.

Scleroderma (Systemic Sclerosis) Renal Disease

KEY CONCEPTS

Nephropathies of Selected Connective Tissue Diseases

- *Scleroderma renal crisis:* Predominantly renal vasculopathy; moderate to severe (high renin) hypertension with progressive renal failure—treated with angiotensin-converting enzyme inhibitors (ACEIs); additional antihypertensive agents may be needed
- *Sjögren syndrome:* Distal renal tubular acidosis, nephrogenic diabetes insipidus, interstitial nephritis, hypokalemia, and/or renal calculi; glomerulonephritis rare

The most common and potentially devastating renal manifestation of systemic sclerosis (Chapter 55) is scleroderma renal crisis (SRC).[47] Most cases of SRC occur within 4 years of the onset of systemic sclerosis in patients with diffuse cutaneous scleroderma affecting the proximal extremities and the trunk. Several features, including rapid progression of skin thickening, palpable tendon friction rubs, anti-RNA polymerase III antibody, recent-onset cardiac events (*e.g.,* pericardial effusion or heart failure), new-onset anemia (especially if associated with microangiopathic hemolysis and thrombocytopenia), and recent treatment with high-dose corticosteroids, help identify patients at increased risk for developing SRC. A classic clinical presentation may obviate the need for renal biopsy. However, renal biopsy may be necessary in atypical cases. For example, about 20% of SRC cases occur before the diagnosis of scleroderma has been established. Furthermore, patients with scleroderma have rarely developed other renal diseases, such as ANCA-associated vasculitis, which are important to recognize because they require treatments different from those usually recommended for SRC.

Scleroderma renal crisis is characterized by abrupt onset of renin-mediated moderate to severe hypertension, rapid deterioration of renal function, and proteinuria (usually nonnephrotic). Associated findings may include microangiopathic hemolysis, hypertensive encephalopathy (including seizures), and hypertensive retinopathy, acute left ventricular failure, and pulmonary edema. Normotensive renal crisis occurs infrequently and may be recognized by the presence of microangiopathic hemolysis and/or unexplained azotemia. The primary pathogenic process appears to be a renal vasculopathy involving predominantly the interlobular arteries and arterioles. Marked intimal thickening with an attendant "mucoid" appearance, and fibrinoid necrosis in the absence of vasculitis, are common and characteristic of the disease (Fig. 68.16). Immune deposits are rarely observed by fluorescence or electron microscopy studies.

Although a variety of treatments have been proposed for patients with scleroderma, none has been proven to be consistently efficacious. The most significant therapeutic advance in the treatment of SRC is the use of angiotensin-converting enzyme (ACE) inhibitors (ACEIs), which have dramatically

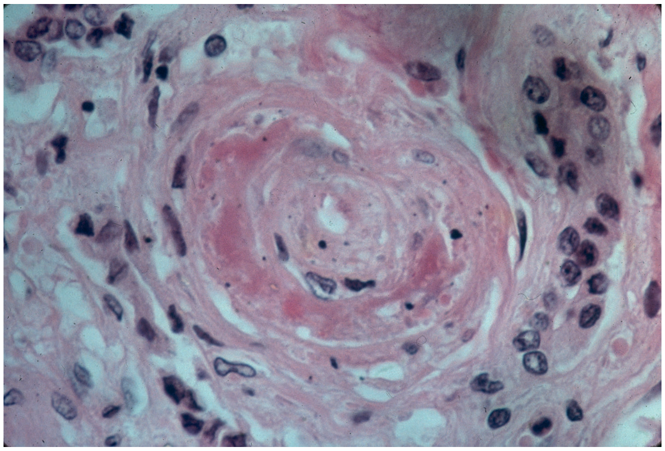


FIG 68.16 **Scleroderma Renal Crisis.** Renal arteriole demonstrates extensive fibrin deposition (dark material) within multiple layers of its wall. The lumen is further compromised by severe swelling and intimal hyperplasia (Masson trichrome stain).

increased the 1-year survival of patients with SCR. Prompt treatment with ACEIs is recommended for hypertensive and normotensive SRC because these agents may reverse the process by interfering with angiotensin II–mediated vasoconstriction and by inhibiting the degradation of the potent vasodilator bradykinin by ACE. Despite this impressive impact of ACEIs on clinical outcomes, approximately one-third to one-half of patients still progress to early death or renal failure. This has prompted continued investigation of additional treatment strategies, including blockade of the RAA system at multiple sites (using, for example, a direct renin inhibitor) as well as addition of potent vasodilators (*e.g.*, prostacyclin or an endothelin receptor antagonist).

Renal Disease of Sjögren Syndrome

A variety of renal manifestations occur in approximately one-third of patients with primary Sjögren syndrome (Chapter 54), including tubular dysfunction (proximal or, more often, distal tubular acidosis, Fanconi syndrome, hypokalemia that may be profound and associated with paralysis, Bartter syndrome, Gitelman syndrome, nephrogenic diabetes insipidus), as well as renal calculi, nephrocalcinosis, interstitial nephritis, pseudolymphoma, necrotizing vasculitis, and glomerulopathy.[48] Studies of renal pathology among patients with primary Sjögren syndrome have shown that acute and/or chronic tubulointerstitial nephritis is the predominant lesion. A diverse range of glomerular diseases, including MN and proliferative glomerulonephritis (focal or diffuse, and membranoproliferative), have been reported in primary Sjögren syndrome. In such cases, the possibility of overlap with SLE should be considered. Immunosuppressive therapy for renal manifestations of primary Sjögren syndrome should be individualized on the basis of activity and severity of the glomerular and/or tubulointerstitial nephritis. Patients with serious complications of tubular dysfunction may warrant immunosuppressive therapies based on preliminary evidence that autoantibodies (against, for example, carbonic-anhydrase-II) as well as cell-mediated injury likely contribute to at least some of those disorders.

ON THE HORIZON

- The advent of new genetic and molecular techniques and new disease models has led to exciting progress in our understanding of the biology of the glomerulus, the pathogenesis of many glomerular diseases, and the influence of genetic variants on disease predisposition and progression.
- It is anticipated that these insights should lead to better noninvasive diagnostic techniques, biomarkers, and predictors of prognosis and relapse and facilitate a more personalized approach to therapy rather than a one-size-fits-all approach.
- Novel targeted therapies are on the horizon that will interrupt or modulate the underlying pathophysiology of the individual diseases and also halt the downstream pathways of injury and fibrosis common to all of the glomerular diseases.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Couser WG, Johnson RJ. The etiology of glomerulonephritis: roles of infection and autoimmunity. Kidney Int 2014;86:905–14.
2. Mastroianni-Kirsztajn G, Hornig N, Schlumberger W. Autoantibodies in renal diseases - clinical significance and recent developments in serological detection. Front Immunol 2015;6:221.
3. Floege J, Amann K. Primary glomerulonephritides. Lancet 2016;387:2036–48.
4. Waldman M, Crew RJ, Valeri A, et al. Adult minimal-change disease: clinical characteristics, treatment, and outcomes. Clin J Am Soc Nephrol 2007;2:445–53.
5. Maas RJ, Deegens JK, Wetzels JF. Permeability factors in idiopathic nephrotic syndrome: historical perspectives and lessons for the future. Nephrol Dial Transplant 2014;29:2207–16.
6. Fornoni A, Sageshima J, Wei C, et al. Rituximab targets podocytes in recurrent focal segmental glomerulosclerosis. Sci Transl Med 2011;3:85ra46.
7. Yu CC, Fornoni A, Weins A, et al. Abatacept in B7-1-positive proteinuric kidney disease. New Engl J Med 2013;369:2416–23.
8. D'Agati VD, Alster JM, Jennette JC, et al. Association of histologic variants in FSGS clinical trial with presenting features and outcomes. Clin J Am Soc Nephrol 2013;8:399–406.
9. Fernandez-Fresnedo G, Segarra A, Gonzalez E, et al. Rituximab treatment of adult patients with steroid-resistant focal segmental glomerulosclerosis. Clin J Am Soc Nephrol 2009;4:1317–23.
10. Iijima K, Sako M, Nozu K, et al. Rituximab for childhood-onset, complicated, frequently relapsing nephrotic syndrome or steroid-dependent nephrotic syndrome: a multicentre, double-blind, randomised, placebo-controlled trial. Lancet 2014;384:1273–81.
11. Malaga-Dieguez L, Bouhassira D, Gipson D, et al. Novel therapies for FSGS: preclinical and clinical studies. Adv Chronic Kidney Dis 2015;22:e1–6.
12. Coppo R. Different targets for treating focal segmental glomerular sclerosis. Contrib Nephrol 2013;181:84–90.
13. Beck LH Jr, Bonegio RG, Lambeau G, et al. M-type phospholipase A2 receptor as target antigen in idiopathic membranous nephropathy. N Engl J Med 2009;361:11–21.
14. Tomas NM, Beck LH Jr, Meyer-Schwesinger C, et al. Thrombospondin type-1 domain-containing 7A in idiopathic membranous nephropathy. New Engl J Med 2014;371:2277–87.
15. Hoxha E, Thiele I, Zahner G, et al. Phospholipase A2 Receptor Autoantibodies and Clinical Outcome in Patients with Primary Membranous Nephropathy. J Am Soc Nephrol 2014;25:1357–66.
16. Waldman M, Austin HA 3rd. Treatment of idiopathic membranous nephropathy. J Am Soc Nephrol 2012;23:1617–30.

17. Fervenza FC, Cosio FG, Erickson SB, et al. Rituximab treatment of idiopathic membranous nephropathy. Kidney Int 2008;73:117–25.
18. Fervenza FC, Abraham RS, Erickson SB, et al. Rituximab therapy in idiopathic membranous nephropathy: a 2-year study. Clin J Am Soc Nephrol 2010;5:2188–98.
19. Remuzzi G, Chiurchiu C, Abbate M, et al. Rituximab for idiopathic membranous nephropathy. Lancet 2002;360:923–4.
20. Pickering MC, D'Agati VD, Nester CM, et al. C3 glomerulopathy: consensus report. Kidney Int 2013;84:1079–89.
21. Fakhouri F, Fremeaux-Bacchi V, Noel LH, et al. C3 glomerulopathy: a new classification. Nature Rev Nephrol 2010;6:494–9.
22. De Vriese AS, Sethi S, Van Praet J, et al. Kidney Disease Caused by Dysregulation of the Complement Alternative Pathway: An Etiologic Approach. J Am Soc Nephrol 2015;26:2917–29.
23. Bertino G, Ardiri A, Proiti M, et al. Chronic hepatitis C: This and the new era of treatment. World J Hepatol 2016;8:92–106.
24. Bruchfeld A, Lindahl K, Reichard O, et al. Pegylated interferon and ribavirin in haemodialysis patients. Nephrol Dial Transplant 2006;21:1444–5.
25. Feng B, Eknoyan G, Guo ZS, et al. Effect of interferon-α-based antiviral therapy on hepatitis C virus-associated glomerulonephritis: a meta-analysis. Nephrol Dial Transplant 2012;27:640–6.
26. Sise ME, Bloom AK, Wisocky J, et al. Treatment of hepatitis C virus-associated mixed cryoglobulinemia with direct-acting antiviral agents. Hepatology *2016*;63:408–17.
27. Sneller MC, Hu Z, Langford CA. A randomized controlled trial of rituximab following failure of antiviral therapy for hepatitis C virus-associated cryoglobulinemic vasculitis. Arthritis Rheum 2012;64:835–42.
28. Saadoun D, Resche-Rigon M, Sene D, et al. Rituximab combined with Peg-interferon-ribavirin in refractory hepatitis C virus-associated cryoglobulinaemia vasculitis. Ann Rheum Dis 2008;67:1431–6.
29. Lan X, Rao TK, Chander PN, et al. Apolipoprotein L1 (APOL1) Variants (Vs) a possible link between Heroin-associated Nephropathy (HAN) and HIV-associated Nephropathy (HIVAN). Front Microbiol 2015;6:571.
30. Booth JW, Hamzah L, Jose S, et al. Clinical characteristics and outcomes of HIV-associated immune complex kidney disease. Nephrol Dial Transplant 2016;31:2099–107.
31. Nobakht E, Cohen SD, Rosenberg AZ, et al. HIV-associated immune complex kidney disease. Nature Rev Nephrol 2016;12:291–300.
32. Cattran DC, Coppo R, Cook HT, et al. The Oxford classification of IgA nephropathy: rationale, clinicopathological correlations, and classification. Kidney Int 2009;76:534–45.
33. Rauen T, Eitner F, Fitzner C, et al. Intensive Supportive Care plus Immunosuppression in IgA Nephropathy. N Engl J Med 2015;373:2225–36.
34. Wyatt RJ, Julian BA. IgA Nephropathy. N Engl J Med 2013;368:2402–14.
35. Karras A, Jayne D. New biologics for glomerular disease on the horizon. Nephron 2014;128:283–91.
36. Berden AE, Ferrario F, Hagen EC, et al. Histopathologic classification of ANCA-associated glomerulonephritis. J Am Soc Nephrol 2010;21:1628–36.
37. Stone JH, Merkel PA, Spiera R, et al. Rituximab versus cyclophosphamide for ANCA-associated vasculitis. N Engl J Med 2010;363:221–32.
38. Jones RB, Tervaert JW, Hauser T, et al. Rituximab versus cyclophosphamide in ANCA-associated renal vasculitis. N Engl J Med 2010;363:211–20.
39. Guillevin L, Pagnoux C, Karras A, et al. Rituximab versus azathioprine for maintenance in ANCA-associated vasculitis. N Engl J Med 2014;371:1771–80.
40. Pedchenko V, Bondar O, Fogo AB, et al. Molecular architecture of the Goodpasture autoantigen in anti-GBM nephritis. N Engl J Med 2010;363:343–54.
41. Greco A, Rizzo MI, De Virgilio A, et al. Goodpasture's syndrome: a clinical update. Autoimmun Rev 2015;14:246–53.
42. Pusey CD. Anti-glomerular basement membrane disease. Kidney Int 2003;64:1535–50.
43. Hahn BH, McMahon MA, Wilkinson A, et al. American College of Rheumatology guidelines for screening, treatment, and management of lupus nephritis. Arthritis care & research 2012;64:797–808.
44. Bertsias GK, Tektonidou M, Amoura Z, et al. Joint European League Against Rheumatism and European Renal Association-European Dialysis and Transplant Association (EULAR/ERA-EDTA) recommendations for the management of adult and paediatric lupus nephritis. Ann Rheum Dis 2012;71:1771–82.
45. Mok CC. Towards new avenues in the management of lupus glomerulonephritis. Nature Rev Rheumatol 2016;12:221–34.
46. Tektonidou MG, Dasgupta A, Ward MM. Risk of End-stage Renal Disease in Patients with Lupus Nephritis, 1970 to 2015 A systematic review and Bayesian meta-analysis. Arthritis Rheumatol 2016;68:1432–41.
47. Mouthon L, Bussone G, Berezne A, et al. Scleroderma renal crisis. J Rheumatol 2014;41:1040–8.
48. Francois H, Mariette X. Renal involvement in primary Sjogren syndrome. Nature Rev Nephrol 2016;12:82–93.

MULTIPLE-CHOICE QUESTIONS

1. A 45-year-old woman presented to the emergency department (ED) with swelling in both ankles, 15-lb weight gain over the past 3 months and shortness of breath. During a physical examination 1 year prior, 1+ protein was noted on dipstick urinalysis, but no further workup was done. There was no family history of renal disease. The patient has blood pressure of 120/80 mm Hg, elevated respiratory rate of 25, and oxygen saturation of 90%. Physical examination was notable for pitting edema in her legs up to the mid-thighs. Lungs were clear. Laboratory testing revealed a hemoglobin level of 14 g/dL, hematocrit of 42%, serum creatinine level of 1.0 mg/dL, blood urea nitrogen (BUN) of 28 mg/dL, albumin level of 1.6 g/dL, serum total cholesterol level of 350 mg/dL (normal, <200 mg/dL), and serum triglyceride level of 800 mg/dL (normal, <150 mg/dL). Serum complement levels were within normal limits. Urinalysis demonstrated 4+ protein on dipstick. Urine microscopy revealed 0–2 nondysmorphic red blood cells per high-power field, oval fat bodies, and fatty casts. A spot protein/creatinine ratio was 8 g/g. The chest radiography result was negative. Computed tomography (CT) angiography revealed multiple pulmonary emboli, and therapeutic anticoagulation was immediately initiated. Renal biopsy was deferred because of the need for urgent anticoagulation. What would have been the most likely biopsy finding if biopsy had been done?
   A. Acute post infectious glomerulonephritis
   B. Alport syndrome
   C. Antineutrophil cytoplasmic antibodies (ANCAs)–associated crescentic glomerulonephritis
   D. Membranous nephropathy (MN)

2. A 20-year-old male with history of short bowel syndrome was admitted to the hospital with elevated creatinine, macroscopic hematuria and proteinuria, and renal insufficiency. He had been on long-term peripheral nutrition via a central venous catheter for 4 years. He was admitted 6 months ago with a line infection and was given antibiotics for 7 days, but the line was not removed. He was noted to have trace proteinuria and hematuria 6 months ago. Upon admission, he was afebrile, and his blood pressure was 140/80 mm Hg. The catheter site was clean, and the remainder of the examination was unremarkable. There was trace edema. Laboratory findings were notable for mild hypoalbuminemia (serum albumin of 3.1 g/dL), serum creatinine of 2.1 mg/dL, hypocomplementemia (C3, 4 mg/dL, C4, 5 mg/dL, and CH50, 10 U/mL). Urinalysis was notable for proteinuria 2+, and urine microscopy revealed 15 acanthocytes per high-power field. The spot urine protein/creatinine ratio was 1.8 g/g. Human immunodeficiency virus (HIV), antinuclear antibody (ANA), and hepatitis B and C screen results were negative. Renal biopsy was performed after admission. Light microscopy (LM) revealed diffuse mesangial and endocapillary hypercellularity, lobulation in the glomeruli, and double contours in the glomerular capillary. Two of 18 glomeruli had segmental crescents. Immunofluorescence microscopy revealed staining for IgG (3+) and C3 (3+) along the capillary and mesangial regions in a granular pattern. Electron microscopy (EM) revealed subendothelial and mesangial electron-dense deposits. What further workup is indicated?
   A. ANCAs testing
   B. Anti–glomerular basement membrane (GBM) antibody
   C. Blood culture
   D. Evaluation for genetic mutations of complement regulatory proteins, including complement factor H (CFH), CFI, and Membrane cofactor protein (MCP).

3. A 60-year-old male with new diagnosis of Hodgkin lymphoma was noted to have 3+ edema on examination by his oncologist. Serum creatinine was 1.1 mg/dL. Serum albumin was 2.5 g/dL. Urinary protein excretion revealed 6 g/24 hr. Urine microscopy showed fatty casts but no cellular elements. Kidney biopsy result was consistent with minimal change disease (MCD). What is the appropriate treatment?
   A. High-dose prednisone
   B. Cyclosporine with prednisone
   C. Mycophenolate mofetil
   D. Chemotherapy

69

Inflammation and Atherothrombosis

Giovanna Liuzzo, Daniela Pedicino, Davide Flego, Filippo Crea

The spectrum of the clinical syndromes caused by coronary atherosclerosis ranges from asymptomatic disease and stable angina (SA) to acute coronary syndromes (ACSs). The latter include unstable angina (UA) and acute myocardial infarction (MI), which is further classified according to electrocardiographic changes as non-ST elevation myocardial infarction (NSTEMI) and ST elevation myocardial infarction (STEMI), and sudden cardiac death.

Although the early outcome of ACS has considerably improved in the past decade, cardiovascular diseases still represent the main cause of morbidity and mortality worldwide, mainly because chronic-smoldering ACS eventually leads to heart failure and sudden cardiac death. The gravity of the syndrome calls for a reappraisal of the mechanisms responsible for coronary instability and for innovative preventive and therapeutic strategies.

The life-threatening complications of coronary artery disease arise as nonlinear events in an otherwise slowly progressive process. This nonlinearity has been attributed to a combination of factors, of which plaque rupture or erosion and superimposed thrombosis are considered to be the most important. Although the pathways leading to plaque rupture or erosion are incompletely understood, our understanding of the importance of inflammation in the process has considerably increased. Atherosclerosis is a chronic inflammatory disease by itself, but an outburst of the inflammatory process within the atherosclerotic plaque of coronary arteries can lead to plaque rupture with thrombosis, resulting in myocardial ischemia and necrosis. The latter is an additional source for local and systemic inflammation. Furthermore, patients with ACS frequently undergo percutaneous coronary intervention (PCI), which may cause iatrogenic myocardial injury, a further source of inflammation. Moreover, plaque instability promotes hyperreactivity to the inflammatory stimulus of myocardial necrosis, which in turn accelerates plaque instability: Inflammation begets inflammation (Fig. 69.1). Thus it is not surprising that a plethora of previous studies have consistently shown a marked elevation of circulating levels of high-sensitivity C-reactive protein (hs-CRP) or of other soluble markers of inflammation at the time of hospital admission. Interestingly, the intensity of this inflammatory surge predicts short- and long-term outcome. Overall, these observations suggest that specific antiinflammatory treatment might improve the outcome of ACS.[1]

ATHEROSCLEROSIS: A CHRONIC INFLAMMATORY DISEASE

Atherosclerosis is a chronic nonresolving inflammatory process that typically occurs at sites of blood flow disturbance (Table 69.1).[2] The atherogenic process develops over the course of decades, beginning in the early teenage years. It is the same worldwide, regardless of gender and race, and proceeds at a faster speed in patients with risk factors such as hypertension, tobacco smoking, diabetes mellitus, obesity, and genetic predisposition. In the early phase, this process is triggered by subendothelial retention of cholesterol-containing plasma lipoproteins and by flow-mediated inflammatory changes in endothelial cells (Table 69.2). At this stage of the disease, endothelial cells show an altered response to vasodilator and vasoconstrictor stimuli, due in part to nitric oxide reduction as well as to increased expression of endothelin. Moreover, the cells express increased levels of adhesion molecules such as intercellular adhesion molecule (ICAM) and vascular cell adhesion molecule (VCAM) and produce higher levels of chemotactic molecules such as interleukin (IL)-8 and monocyte chemotactic protein (MCP), with the consequent increase in leukocytes' recruitment. These mechanisms act together to induce endothelial dysfunction and a chronic self-maintaining inflammatory state. Plaque formation is preceded by a fatty streak, an accumulation of lipid, macrophages/foam cells, and T cells beneath the endothelium. Fatty streaks may disappear or progress to atherosclerotic plaques, asymmetrical focal thickenings of the intima with a core region of foam cells and extracellular lipids surrounded by a cap of smooth-muscle cells and collagen. Atherosclerotic lesions contain monocyte-derived macrophages and T cells interspersed with lipids and debris from dead cells; the lesions are embedded in an extracellular matrix composed of collagen fibers and other constituents produced primarily by vascular smooth muscle cells. T cells, macrophages, and mast cells infiltrate the lesion; they are particularly abundant in the shoulder region of the atheroma, producing proinflammatory cytokines, costimulatory factors for immune activation, eicosanoids, and reactive oxygen and nitrogen species. On the other hand, many immune cells beneath the plaque show antiinflammatory features: Macrophages internalize cholesterol through their scavenger receptors and produce antiinflammatory cytokines, and T cells of the regulatory phenotype produce antiinflammatory and immunosuppressive cytokines. The perturbation of this delicate balance between pro- and antiinflammatory signals is involved both in the nonresolving, chronic inflammation linked to disease progression and in the eventual breakdown of endothelial continuity, with thrombus formation and plaque instability (Fig. 69.1). Fibrous plaques appear in a later stage of the disease, with the chronic accumulation of collagen and calcification. Erosion of the surfaces of some plaques and rupture of a plaque's calcific nodule into the artery lumen may trigger thrombosis.[3] Overall, mechanisms

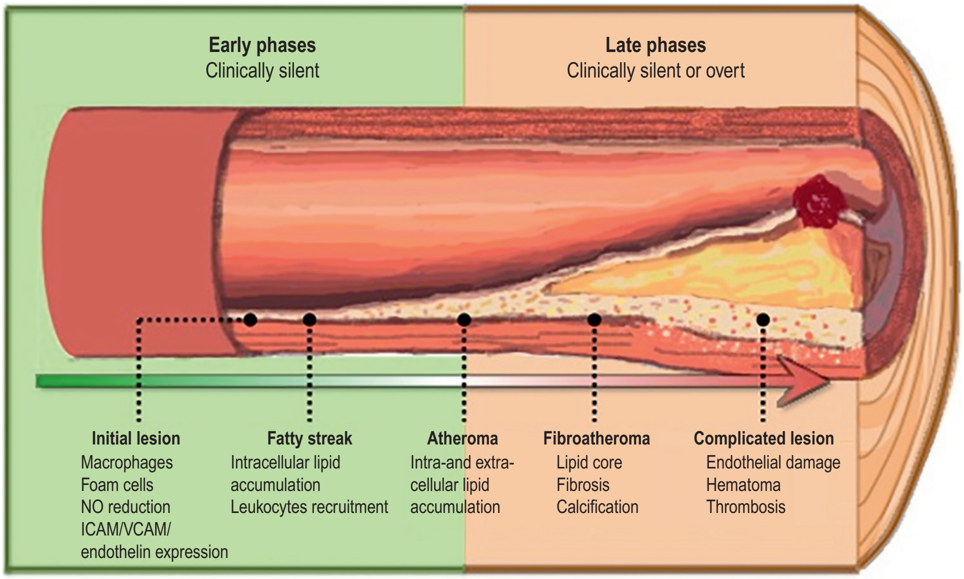


FIG 69.1 **Schematic Representation of the Natural History of Atherosclerotic Disease.** This figure shows the development of the atherogenic process. In the early phase, this process is triggered by subendothelial retention of cholesterol-containing plasma lipoproteins and by flow-mediated inflammatory changes in endothelial cells. Endothelial cells show an altered response to vasodilator and vasoconstrictor stimuli, with nitric oxide (NO) reduction and increased expression of endothelin. Moreover, they express increased levels of adhesion molecules such as intercellular adhesion molecule (ICAM) and vascular cell adhesion molecule (VCAM) and produce higher levels of chemotactic molecules such as interleukin (IL)-8 and monocyte chemotactic protein (MCP), with a consequent increase in leukocyte recruitment. The fatty streak is an accumulation of lipid, macrophages/foam cells, and T cells beneath the endothelium. Fatty streaks may disappear or progress to atherosclerotic plaques, asymmetrical focal thickenings of the intima with a core region of foam cells and extracellular lipids surrounded by a cap of smooth-muscle cells and collagen. Fibrous plaques appear at a later stage of the disease and are characterized by the deposition of collagen and calcification.

TABLE 69.1 Key Steps in Atherogenesis

1. Endothelial activation with increased infiltration of atherogenic lipoproteins at sites of low or oscillating shear stress (branch points and flow dividers).
2. Subendothelial retention and modification of atherogenic lipoproteins (low-density lipoprotein/very-low-density lipoprotein).
3. Endothelial activation with increased mononuclear leukocyte (inflammatory cell) adhesion, chemotaxis, and subendothelial recruitment.
4. Subendothelial inflammatory cell activation with lipid ingestion through monocyte scavenger receptor expression resulting in foam cell formation.
5. Intimal migration and proliferation of medial/adventitial smooth-muscle cells/myofibroblasts in response to growth factors released by activated monocytes with matrix production and formation of the fibrous cap and fibrous plaque.
6. Abluminal plaque growth with positive (outward) arterial adventitial remodeling preserving lumen size in early stages; later, plaque growth or negative remodeling results in luminal narrowing.
7. Neoangiogenesis due to angiogenic stimuli produced by inflammatory cells (macrophages) and other arterial wall cells (vascular endothelial growth factor, interleukin-8).
8. Death of foam cells by necrosis/apoptosis leading to necrotic lipid core formation.
9. Plaque disruption (rupture of fibrous cap or endothelial erosion) due to inflammatory cell–mediated matrix degradation and death of matrix-synthesizing smooth-muscle cells.
10. Exposure of thrombogenic substrate (lipid core–containing tissue factor derived from inflammatory cells) following plaque disruption with arterial thrombosis.

TABLE 69.2 Factors Promoting Endothelial Activation/Dysfunction in Atherosclerosis

- Reduced vasodilator and increased vasoconstrictor responsiveness
- Enhanced oxidant stress with inactivation of nitric oxide
  - Increased expression of endothelin
- Enhanced leukocyte (inflammatory cell) adhesion and recruitment
- Increased adhesion molecule expression (ICAM, VCAM)
  - Increased chemotactic molecule expression (MCP-1, IL-8)
- Increased prothrombotic and reduced fibrinolytic activity
- Increased growth-promoting phenotype
- Dyslipidemia and atherogenic lipoprotein modification
  - Elevated LDL, VLDL, LP(a)
  - LDL modification (oxidation, glycation)
  - Reduced HDL
- Increased angiotensin II and hypertension
- Estrogen deficiency
- Smoking
- Hyperhomocysteinemia
- Advancing age
- Infection?

HDL, high-density lipoprotein; ICAM, intercellular adhesion molecules; IL-8, interleukin-8; LDL, low-density lipoprotein; LP, lipoprotein; MCP-1, monocyte chemotactic protein-1; VCAM, vascular cell adhesion molecules; VLDL, very-low-density lipoprotein.

responsible for coronary instability are multiple and insufficiently understood.

ROLE OF INFLAMMATION IN THE PATHOGENESIS OF ACUTE CORONARY SYNDROMES

Experimental models of atherogenesis have provided a growing body of information about molecular mechanisms of plaque growth; however, transition from coronary stability to instability is less well understood due to lack of animal models reflective of human disease.[1] The abrupt clinical presentation of ACS gives a strong signal of discontinuity in the natural history of atherothrombosis. When primary prevention of atherosclerosis fails, the progression of coronary atherosclerosis can remain clinically silent for years, decades, or even for life, as indicated by the high prevalence of coronary atherosclerosis in subjects dying of noncardiac causes. In contrast, some patients at a certain point in their lives exhibit acute coronary instability, followed by a period of stability that can be short or last for years or decades. These simple clinical observations suggest that the mechanisms responsible for plaque growth and for plaque instability are different and that the causes and mechanisms of plaque instability are multiple. Accordingly, the paradigm that implies a single type of culprit coronary plaque as a cause for instability does not adequately fit the findings of postmortem studies.[1] The concept that inflammatory cell activation plays a key role in the pathogenesis of ACS is now commonly accepted, and it is believed that activation of inflammatory cells in the culprit stenosis is the cause of coronary instability in all patients. However, about 40% of ACS patients have low or very low levels of high-sensitivity C-reactive protein (hs-CRP), a very sensitive marker of inflammation.[4] Coronary angiography fails to demonstrate obstructive atherosclerosis in up to one-third of patients with symptoms suggestive of an ACS along with raised troponin levels and/or ischemic-like ST segment changes, suggesting that functional alterations of epicardial arteries and/or of coronary microcirculation may also play an important pathogenetic role.[5]

We have recently proposed a pathogenetic classification of ACS: (i) patients with plaque fissure and systemic inflammation; (ii) patients with plaque fissure without systemic inflammation; (iii) patients with plaque erosion; and (iv) patients with smooth plaques and functional alterations in the coronary circulation (Fig. 69.2). For the first group, multiple studies have highlighted the prominent role of inflammation in generating plaque instability, with three main features characterizing these ACS patients: (i) the widespread involvement of epicardial arteries, coronary microcirculation, and myocardium; (ii) the activation of innate immunity; and (iii) the activation of adaptive immunity, with a consequent dysbalance between antiinflammatory and proinflammatory immune responses.

In patients in whom plaque rupture and coronary instability occur in the absence of systemic evidence of inflammation, the precise causes of instability are still poorly understood. However, mechanisms such as extreme emotional disturbance, intense physical exertion, and local mechanical stress at the level of the artery wall are likely to play a key pathogenetic role.

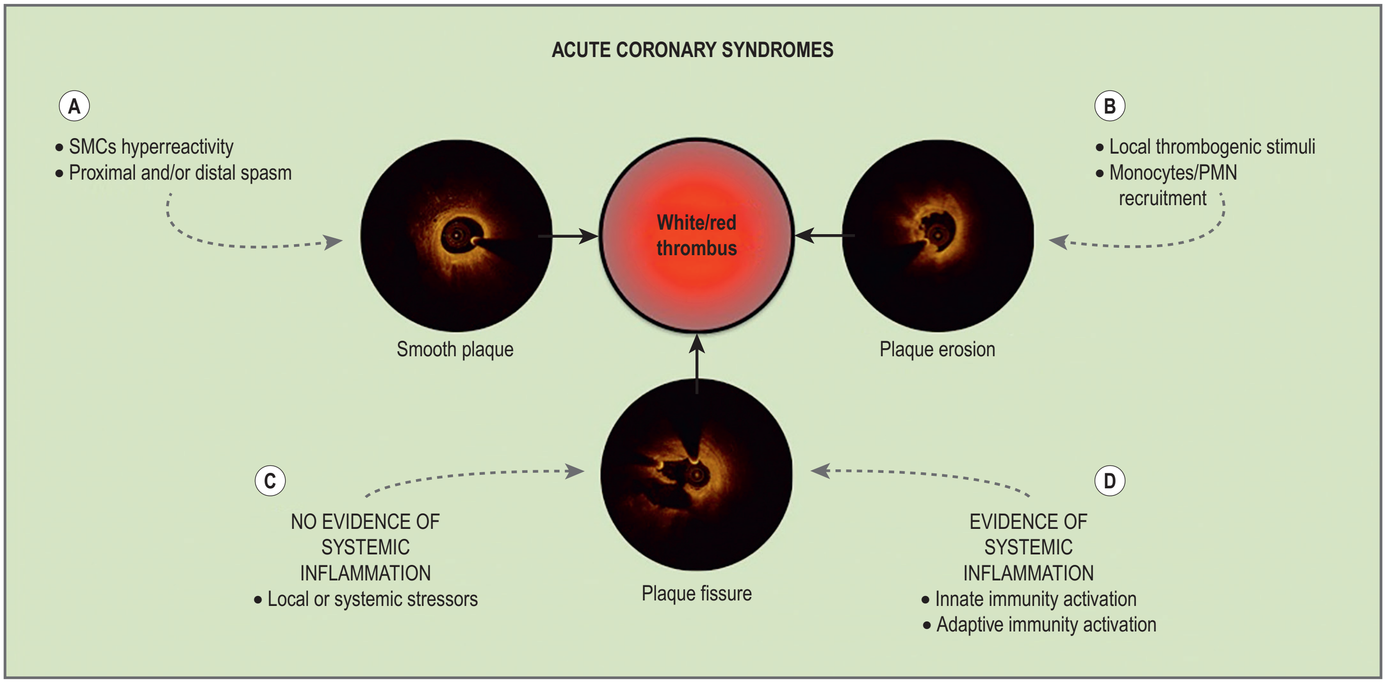

ACUTE CORONARY SYNDROMES
A
• SMCs hyperreactivity
• Proximal and/or distal spasm
B
• Local thrombogenic stimuli
• Monocytes/PMN recruitment
White/red thrombus
Smooth plaque
Plaque erosion
C
NO EVIDENCE OF SYSTEMIC INFLAMMATION
• Local or systemic stressors
Plaque fissure
D
EVIDENCE OF SYSTEMIC INFLAMMATION
• Innate immunity activation
• Adaptive immunity activation

FIG 69.2 Pathogenetic Classification of Acute Coronary Syndromes. In patients with a similar clinical presentation, mechanisms responsible for coronary instability can be significantly different. The figure shows optical coherence tomography (OCT) images of different plaques and the proposed pathogenetic classification of acute coronary syndromes (ACS). This classification is based on the mechanisms responsible for coronary instability in homogeneous groups of patients: (D) patients with plaque fissure and systemic inflammation; (C) patients with plaque fissure without systemic inflammation; (B) patients with plaque erosion; (A) patients with functional defects in coronary circulation. PMN, Ppolymorphonuclear neutrophil; SMCs, smooth muscle cells (SMCs).

The mechanisms of plaque erosion are still largely unknown, although increasing evidence suggests that local neutrophil activation possibly plays a role.

Among the majority of patients who present with ACS in the absence of obstructive atherosclerosis, functional alterations of epicardial coronary arteries or of coronary microcirculation are the likely cause of instability.

PLAQUE FISSURE WITH SYSTEMIC INFLAMMATION

Plaque fissure is caused by the combined action of metalloproteinases, which digest the intercellular matrix and reduce collagen synthesis. This results in the exposure of the highly thrombogenic content of the underlying lipid plaque that triggers occlusive or subocclusive thrombus formation.

Plaque Fissure: Pathological and Clinical Findings

The role played by plaque fissure was initially established in postmortem studies and then confirmed *in vivo* by using angioscopy, intravascular ultrasound, and, more recently, optical coherence tomography (OCT), a light-based imaging technique with high spatial resolution. OCT allows the detection of intracoronary thrombi and the precise measurement of fibrous cap thickness and plaque components such as lipid core, calcium nodules, and microvessels.[6] OCT has been used in clinical studies to obtain more detailed information about the vulnerable plaque and to find specific features to differentiate between erosion-prone and rupture-prone plaques. In particular, the high resolution of OCT has recently clarified that *in vivo* plaque rupture is associated with ACS in about half of patients. It is worth noting that fissures do not necessarily lead to ACS, as asymptomatic plaque fissure is a frequent event leading to plaque progression. Yonetsu et al. using OCT found that about 15% of patients with SA presented with plaque fissure.[7] A larger plaque burden, a thinner fibrous cap, and a smaller lumen size appear to predispose to thrombus formation and ACS.[8]

Among the several inflammatory biomarkers tested in ACS, levels of hs-CRP are typically elevated, in line with CRP's high sensitivity and wide dynamic range.[9] However, it is difficult to establish a cutoff at the present time. Whereas in primary prevention hs-CRP levels ≥2 mg/L (and even ≥1 mg/L) might be recommended, in ACS patients two different cutoffs have been suggested as clinically useful: an admission value ≥10 mg/L and a discharge value ≥3 mg/L. These cutoffs, although confirmed in larger studies, have not yet led to a general consensus.[4,10]

Experimental studies have clarified the molecular mechanisms through which activation of inflammatory cells in the plaque can trigger thrombus formation. Inflammatory responses increase the fragility of the fibrous cap as well as the thrombogenic potential of the plaque. The main mediators of inflammation-induced activation of coagulation are proinflammatory cytokines. Several studies have shown the importance of IL-6 in the initiation of coagulation activation and the role of tumor necrosis factor-α (TNF-α) and IL-1 in the modulation of anticoagulant pathways.

Widespread Coronary Inflammation

In patients with ACS and systemic evidence of inflammation, widespread coronary inflammation is suggested by transcardiac neutrophil activation in the effluent of myocardial regions not perfused by the culprit artery. This event appears unrelated to coronary atherosclerosis or recurrent ischemia, as it is not observed in patients with chronic SA and multivessel coronary disease or in patients with vasospastic angina.[11] Widespread acute coronary inflammation is, therefore, the likely cause of multiple complex stenoses, multiple thrombi, and multiple fissured plaques involving different coronary artery branches observed in clinical studies in ACS, based on angiography and intravascular imaging.[6] The number of disrupted coronary plaques correlates with systemic hs-CRP levels. The notion of widespread coronary inflammation is confirmed by postmortem studies.[1]

Activation of Innate Immunity

The notion that innate immunity plays an important role in ACS is supported by the demonstration of activated monocytes, polymorphonuclear neutrophils (PMNs), eosinophils, and mast cells not only at the site of plaque rupture but also in the whole coronary circulation of patients with ACS (Fig. 69.3).[1] High telomerase activity in PMNs has been reported coming from the culprit coronary plaque of patients with ACS, but not from plaque of patients with SA or in PMNs from the peripheral blood. Telomerase activity is normally absent in differentiated cells such as PMNs. The only predictor of telomerase reactivation in coronary plaques of ACS patients was a short time interval from symptom onset to PMN sampling. Neutrophil apoptosis has been identified as one of the key mechanisms to switch off inflammation. Accordingly, delayed apoptosis of peripheral PMNs in ACS patients has been demonstrated.[12]

Macrophages account for the majority of leukocytes in plaques. Plaque-resident macrophages differentiate from monocytes recruited from circulating blood. However, monocytes represent a heterogeneous cell population, exemplified by the differential expression of CD14 and CD16.[13] Human coronary artery lesions contain macrophage subpopulations with different gene-expression patterns. Patients with coronary atherosclerosis have higher numbers of circulating $CD14^+CD16^+$ monocytes than healthy subjects. Finally, peak levels of $CD14^{hi}CD16^{lo}$ monocytes after acute MI correlate negatively with the recovery of left ventricular ejection fraction 6 months after MI. Noninvasive imaging technologies have now been developed to follow and quantify subpopulations of monocytes and macrophages *in vivo.*

Monocytes accumulated within thrombi, obtained during primary PCIs, specifically overexpress Toll-like receptor 4 (TLR4), together with specific patterns of locally expressed chemokines and cytokines compared with circulating monocytes.[14] TLR are key pattern-recognition receptors expressed by innate immunity cells that recognize a large number of pathogen-associated molecular patterns (PAMPs). Interestingly, Niessner et al. demonstrated that interferon (IFN)-α produced by plasmacytoid dendritic cells in atherosclerotic plaques could enhance TLR4 signaling by sensitizing these cells to lipopolysaccharide and other microbial molecules as well as to (modified) endogenous molecules, all abundantly present in the atherosclerotic lesion microenvironment. These sensitized antigen-presenting cells (APCs) strongly upregulate the production of cytokines such as TNF-α, IL-12, IL-23, and matrix metalloproteinase-9 (MMP-9), thus enhancing plaque instability.[15]

It has been demonstrated that human platelets express functional TLRs capable of recognizing bacterial components. TLR activation directly induces platelet aggregation and increased platelet adhesion to collagen under flow conditions. Moreover, stimulation of TLR, in particular TLR2, induces a significant

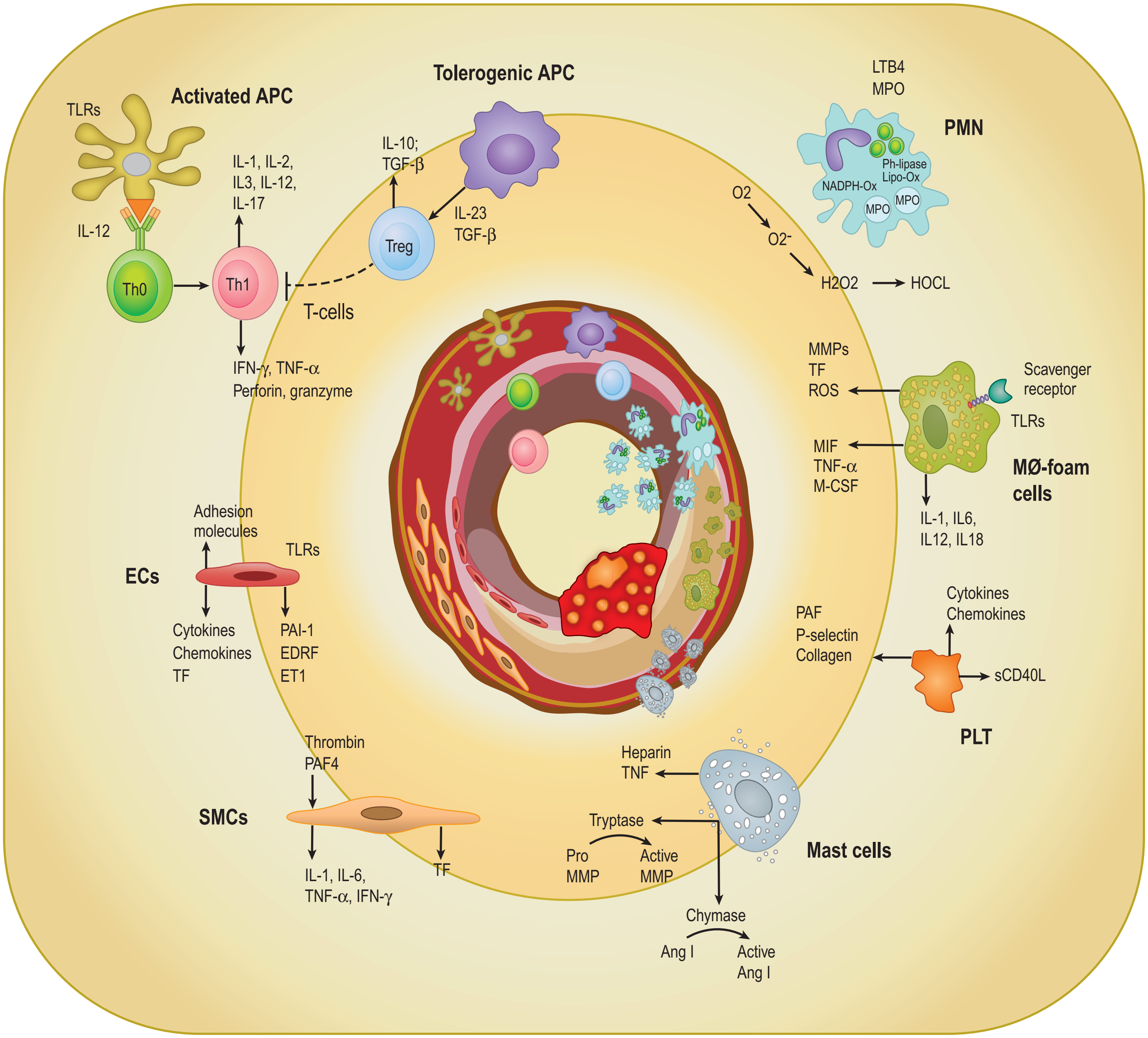

TLRs
Activated APC
IL-12
Th0
Th1
IL-1, IL-2,
IL3, IL-12,
IL-17
IFN-γ, TNF-α
Perforin, granzyme
Tolerogenic APC
IL-10;
TGF-β
Treg
IL-23
TGF-β
T-cells
LTB4
MPO
PMN
Ph-lipase
Lipo-Ox
NADPH-Ox
MPO
MPO
O2
O2-
H2O2
HOCL
MMPs
TF
ROS
Scavenger
receptor
TLRs
MIF
TNF-α
M-CSF
MØ-foam
cells
IL-1, IL6,
IL12, IL18
Adhesion
molecules
TLRs
ECs
Cytokines
Chemokines
TF
PAI-1
EDRF
ET1
PAF
P-selectin
Collagen
Cytokines
Chemokines
sCD40L
PLT
Thrombin
PAF4
SMCs
IL-1, IL-6,
TNF-α, IFN-γ
TF
Heparin
TNF
Tryptase
Pro
MMP
Active
MMP
Mast cells
Chymase
Ang I
Active
Ang I

FIG 69.3 **Histological Features of Fissure-Prone Plaques.** Both innate immunity and adaptive immunity play a key role in the pathogenesis of coronary plaque instability. Multiple types of inflammatory cells are present in atherosclerotic plaques. Macrophages (MØ) and mast cells infiltrate the lesion and are particularly abundant in the shoulder region where the atheroma grows and where the risk of plaque rupture is highest. T-cell infiltrates are always present in atherosclerotic lesions. Such infiltrates are predominantly CD4 T cells, which recognize protein antigens processed and presented by activated antigen-presenting cells (APCs). Regulatory T cells (Tregs) maintain the homeostasis of cell subsets involved in adaptive immunity. In human atherosclerotic lesions, Treg colocalize with interleukin-10 (IL-10) and transforming growth factor (TGF)-β expression. TLR, Toll-like receptor; Ang, angiotensin I; EC, endothelial cell; EDRF, endothelium-derived relaxing factor; ET1, endothelin 1; IFN, interferon; IL12R, IL-12 receptor; LTB-4, leukotriene B-4; M-CSF, macrophage colony–stimulating factor; MMPs, metalloproteinases; MØ, macrophage; MPO, myeloperoxidase; TF, tissue factor; PAF, platelet-activating factor; PAI, plasminogen activator inhibitor; PLT, platelet; PMN, polymorphonuclear cell; ROS, reactive oxygen species; sCD40L, soluble CD40 ligand; SMC, smooth muscle cell.

increase in the interactions of platelets and leukocytes and in the amplification of platelet-derived inflammatory signals. These findings highlight the role of platelets as immunological cells, critically participating in both inflammatory and thrombotic processes. Indeed, platelet TLR2 and related innate immune transcripts have been associated with cardiovascular disease and its risk factors. Expression of TLR2 in megakaryocytes suggests that inflammatory processes, through TLR2 stimulation, can increase megakaryocyte maturation and can modulate megakaryocyte phenotypes, potentially influencing platelet function and thrombosis.

Activation of Adaptive Immunity

In patients with ACS and systemic evidence of inflammation, the higher frequency of activated T cells in comparison to patients with SA suggests that the sudden changes leading to coronary instability might be related to mechanisms involving T-cell immunity. In particular, in ACS patients CD4 T-cell subpopulations are dysregulated, and their abnormalities are associated with worse outcomes, especially in patients with diabetes mellitus.[16-18]

ACS patients have a skewed T-cell differentiation oriented toward aggressive effector phenotypes and defective regulatory T cells, suggesting that the lymphocyte compartment fails to suppress excessive immune responses. Overall, such T-cell abnormalities characterize about half of ACS patients.[16] In this subset of ACS patients, helper T-cell dysregulation might affect the biological outcome of the immune response and contribute to plaque destabilization through multiple damaging pathways (Fig. 69.3).

T cells physiologically drive B-cell hypersomatic mutation, isotype switch, and affinity maturation in the germinal centers of lymphoid organs and also in extranodal sites in autoimmune diseases[19]. Indeed, coronary atherosclerotic plaques obtained by endoluminal directional atherectomy are sites of B lymphocyte proliferation in patients with ACS, strongly suggesting the local presence of an activating or recruiting antigen in these lesions (Table 69.3).

T-helper 1 (Th1) Cells

The importance of Th1 and IFN-γ in atherosclerosis progression and in plaque destabilization is confirmed by numerous studies showing a marked increase in Th1 frequency and increased expression of Th1-related effector molecules such as IFN-γ, signal transducer and activator of transcription 4 (STAT4), and T-bet[20]. Th1 cells are suspected of regulating the fragility of the fibrous cap as well as the thrombogenic potential of the plaque. IFN-γ can contribute to plaque destabilization in several ways: by the recruitment and activation of macrophages in the atherosclerotic lesions; by reducing collagen synthesis; by increasing the production of extracellular matrix–degrading proteins; and by activating APC. Furthermore, increased IFN-γ expression induces a positive feed-forward loop of Th1 induction sustaining a proinflammatory state. When activated in the intima, Th1 cells produce proinflammatory cytokines and enhance the expression of CD40 ligand. Ligation of CD40 on APC by CD40 ligand induces release of extracellular matrix–degrading metalloproteinases and the expression of tissue factor, a key initiator of the coagulation cascade.

TABLE 69.3 **Potential Role of Infection in Atherosclerosis and Thrombosis**

Infectious Organisms Implicated

Viruses
  Herpes viruses (*e.g.*, cytomegalovirus)
Bacteria
  Chlamydia pneumonia
  Helicobacter pylori
  Porphyromonas gingivalis?

Mechanism(s) by Which Infections May Contribute to Atherothrombosis

Direct infection of the vascular wall with endothelial injury, inflammatory cell recruitment, and activation (*Chlamydia pneumoniae, herpes virus, cytomegalovirus*)
Immune-mediated vascular injury through molecular mimicry (*Chlamydia pneumoniae*)
Remote infections with systemic activation of the inflammatory response (*Helicobacter pylori, Porphyromonas gingivalis*)

$CD4^+CD28^{null}$ T Cells

$CD4^+CD28^{null}$ T cells are distinct from classic helper T cells in several aspects. This terminally differentiated subpopulation shows an increased resistance to apoptosis and a wide range of proinflammatory properties.

$CD4^+CD28^{null}$ T cells are present preferentially in unstable ruptured atherosclerotic plaques, and their frequency significantly increases the risk of ACS, particularly in patients with diabetes.[16-18] $CD4^+CD28^{null}$ T cells from ACS patients show several abnormalities such as the ability to mediate endothelial cytotoxicity by triggering the NK stimulatory receptor KIR2DS2, killer Ig-like receptor 2DS2 (KIR2DS2) in the absence of T-cell receptor (TCR) activation and higher levels of costimulatory receptors, *e.g.*, OX40 and 4–1BB.

Th17 Cells

Th17 cells are characterized by the expression of retinoid orphan receptor (ROR)-γt, the master regulator transcription factor responsible for the production of IL-17.

Although the precise role of IL-17 in atherosclerosis and ACS remains controversial,[21] experimental studies in mice have provided direct evidence that IL-17 is predominantly proatherogenic. Indeed, in combination with IFN-γ, IL-17 induces a proinflammatory response in vascular smooth muscle cells. On the other hand, Th17 cells are involved in wound healing and exert powerful fibrogenic activity. These cells activated in the context of the atherosclerotic plaque promote the formation of thick collagen fibers that can resist the mechanical assault exerted by hemodynamic forces. This is due to the capacity of the IL-17 to promote procollagen expression.

Regulatory T Cells

In atherosclerosis, the role of regulatory T cells (Treg) is well appreciated. Indeed, this T-cell subpopulation inhibits atherosclerosis development and progression by suppressing effector T-cell responses. A defective Treg compartment has been demonstrated in the peripheral blood of ACS; these patients show low levels of circulating Treg, a reduced suppressive efficiency of Treg, and an increased Treg susceptibility to apoptosis.[16] In a small study, Treg have been identified as the major T-cell subset in aspirated coronary thrombus adjacent to the culprit lesion in patients with ACS, associated with a restricted TCR diversity in thrombus-resident T cells; this suggests an increased antigen-specific Treg redistribution between the peripheral blood and local sites of inflammation.[22] Nevertheless, ACS patients have a lower induction of Treg after TCR stimulation with anti-CD3 and anti-CD28 antibodies.[23] Notably, the defective TCR-induced Treg generation is partially recovered 1 year after the index event, suggesting a transient phenomenon present during the acute phase of the disease. Thus defects in the Treg compartment might affect the plaque stability by impairing a balanced immune response in atherosclerosis. Indeed, Treg have been implicated in the inhibition of effector T cells, in the suppression of endothelial and macrophage activation, and in the modulation of cholesterol metabolism.

B Cells

The contribution of B cells in autoimmunity is well recognized. The function of B cells has been evaluated in various animal models of atherosclerosis; B1a cells and their secretion of IgM seem to mediate the protective effects of B cells, whereas B2 cells are proatherogenic through modulation of T cell–dependent mechanisms.

Moreover, antibodies against several oxidation-specific epitopes of low-density lipoprotein and other atherosclerosis-related antigens are found in animals and humans, both in the circulation and in atherosclerotic plaques.[24] The mechanisms by which these antibodies act remain unclear, and it is unresolved whether they represent an epiphenomenon of chronic immune stimulation or a disease-specific phenomenon.

TCR Signaling Alteration in ACS

The first step in lymphocyte activation is TCR binding of specific peptides presented by APC. TCR triggering initiates a cascade of phosphorylation events that culminate in the activation and nuclear translocation of transcription factors (Chapter 12). A complex molecular coordination is required, including positive and negative feedback loops necessary to avoid lymphocyte hyperreactivity and the breaking of immune tolerance. The proper tuning of TCR signaling is critical also for T-cell differentiation. Indeed, in addition to the cytokine environment, the induction of different T-helper cell lineages is driven by TCR-mediated signal strength.

CD4 T cells from ACS patients show enhanced response to TCR stimulation and have a lower setting of the T-cell activation threshold, attributable to enhanced amplification of proximal TCR-mediated signals.[23,25] Intriguingly, some of these abnormalities are confined to the acute phase of ACS, whereas others persist also during the stable phase of the disease, suggesting the existence of chronic lymphocyte hyperreactivity.

The molecular TCR signaling alterations observed in ACS include an increase in the positive activation signals, including higher accumulation of CD3 complexes and zeta-chain associated protein kinase of 70 kD (Zap70) in the immunological synapse during antigen presentation; higher early tyrosine phosphorylation after TCR stimulation; and a defective deactivation of the lymphocyte-specific protein tyrosine kinase (LCK). On the other hand, ACS patients exhibit reduced activity of the inhibitory molecular patterns such as reduced phosphorylation of Zap70 at its inhibitory residue Tyr-292, a residue implicated in TCR deactivation; lower expression of PECAM-1, a molecule implicated in downmodulation of T cell activity; and reduced activation of CREB, a transcription factor believed to be particularly important for the generation and maintenance of Treg and for IL-2 and IL-10 production.[25] Finally, ACS patients show enhanced expression of the protein tyrosine phosphatase non–receptor type 22 (PTPN22). This enzyme plays a key role in controlling the intensity of the early TCR signal transduction acting on LCK and on Zap70.[23]

PLAQUE FISSURE WITHOUT SYSTEMIC INFLAMMATION

When plaque fissure occurs without systemic inflammatory activation, other mechanisms, including emotional and physical stress or changes in plaque composition, may play a pathogenic role.[1] The ability of systemic stress to induce plaque fissure might be related to sympathetic nervous system activation and catecholamine release associated with increased heart rate, increased blood pressure, and coronary vasoconstriction. Besides the fissuring of vulnerable plaque, these conditions favor platelet activation, hypercoagulability, and intense coronary microvascular constriction. It has been demonstrated that the highest shear stress is present at the level of the shoulder region of the fibrous cap. Changes in plaque composition may be another possible cause of plaque fissure. Indeed local changes in pH, temperature, cholesterol saturation, and hydration promote cholesterol crystallization in the lipid core associated with quick volume expansion, potentially causing plaque fissure and thrombosis. These mechanisms may be amplified by crystallization of free cholesterol from erythrocyte membranes when intraplaque hemorrhage occurs.

PLAQUE EROSION

Plaque Erosion: Pathological Findings

Plaque erosion is associated with coronary instability in about one-third of patients. Eroded plaques are characterized by thrombus formation at the site of a denuded fibrous atherosclerotic plaque. Endothelial denudation leads to the exposure of the intima that often shows a pathological thickening and consists predominantly of vascular smooth muscle cells and proteoglycans, heavily glycosylated proteins able to form large complexes by binding other proteoglycans, hyaluronan, and collagen. There are no pathognomonic features to identify plaque erosion. An eroded plaque in a histological specimen is recognized when serial sectioning of thrombosed arterial segments fails to reveal plaque rupture. Compared with plaque fissure, which typically has a large lipid pool, plaque erosion shows intact internal and external elastic laminas and a well-developed media (unlike plaque rupture, where the internal lamina is often disrupted and the underlying media is thin and disorganized), rare calcifications, and a lower grade of inflammation.[26] There is still no agreement on the grade of luminal obstruction related to plaque erosion because some groups have demonstrated that eroded plaques are less obstructive than ruptured plaques with thrombosis, and others report the opposite[26] (Fig. 69.4).

The detection of eroded plaques by OCT is based upon exclusion criteria.[6] Similar to histological detection, OCT confirmation of plaque erosion derives from the absence of fibrous cap discontinuation after intraluminal thrombi aspiration. It is not always possible to obtain a proper distinction between fissured and eroded plaque by imaging, and it is even more challenging to distinguish erosion-prone plaques from stable plaques.

Mechanisms of Plaque Erosion

Several studies have highlighted the high frequency of ACS associated with plaque erosion in women and in young persons, mostly after stressful events[1] and possibly reflecting acute changes

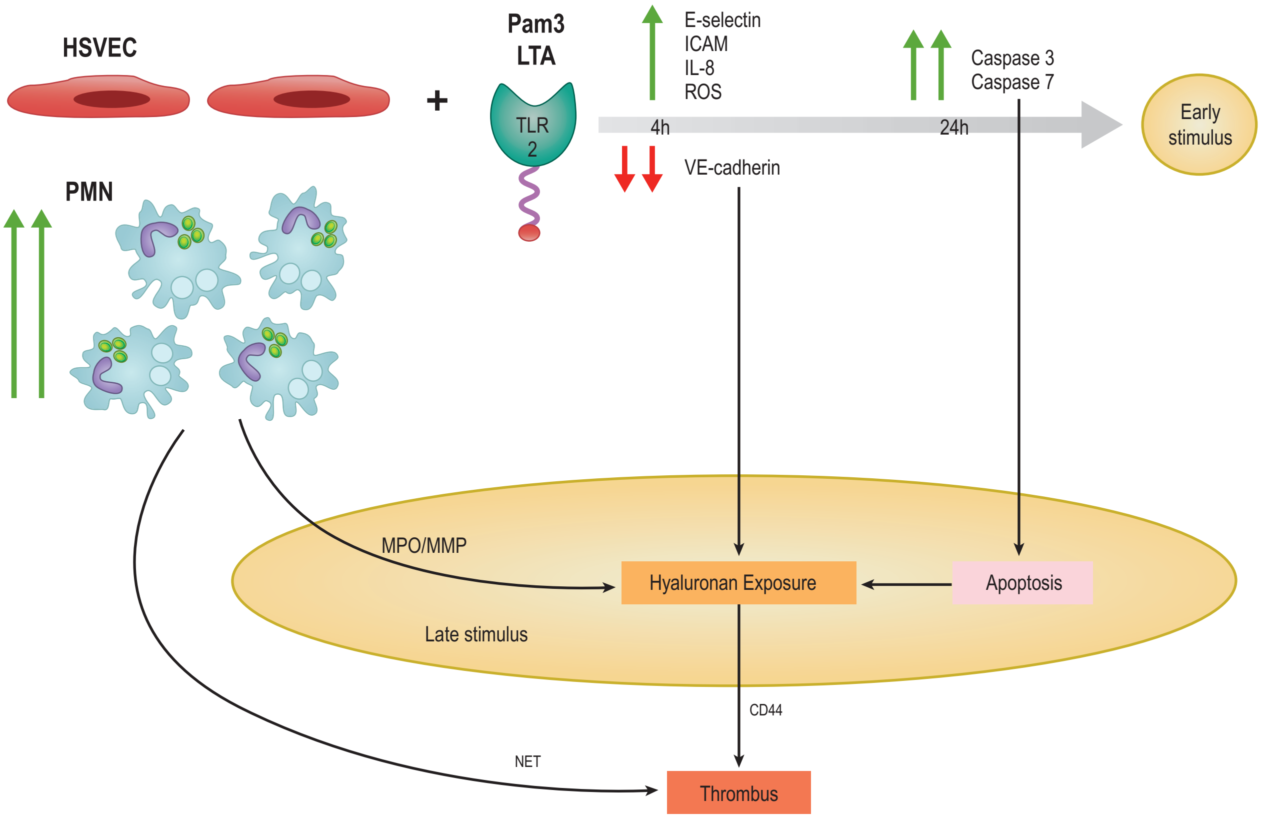


FIG 69.4 **A Hypothetical Model of Plaque Erosion.** Toll-like receptor (TLR) 2 stimulation on endothelial cells (HSVEC [human saphenous vein endothelial cells]) with Pam3 or lipoteichoic acid (LTA) has been reported to induce endothelial activation and apoptosis, as suggested by increased levels of adhesion molecules (E-selectin, ICAM), interleukin-8, and caspases -3 and -7. This process is markedly potentiated but not necessarily initiated by polymorphonuclear neutrophil (PMN) and leads to endothelial damage and extracellular matrix (ECM) exposure. Hyaluronic acid, a component of ECM, is a ligand for TLR2 and for the receptor CD44 and might be involved in the propagation of the inflammatory process. In humans, eroded plaques typically harbor abundant hyaluronan and proteoglycans. Thus hyaluronic acid could be considered a relevant TLR2 agonist. Also, products from infectious microorganisms may operate to promote atherothrombosis through this mechanism. MPO, myeloperoxidase; MMP, metalloproteinase; NET, neutrophil extracellular trap; ICAM, intercellular adhesion molecule; ROS, reactive oxygen species.

in local hemodynamics and atherosclerotic artery. Animal models have demonstrated that local vasoconstriction that precedes myocardial ischemia is occasionally accompanied by intracoronary thrombosis with endothelial lesions resembling human plaque erosion. A role of the innate immune system in this process has also been proposed, specifically a possible involvement of neutrophils. Coculture of neutrophils with endothelial cells induced endothelial injury and triggered endothelial cell apoptosis. Plasma from patients with eroded culprit plaques exhibited elevated systemic myeloperoxidase (MPO) levels, produced primarily by PMNs, compared with levels in patients showing plaque fissure. Moreover, in postmortem coronary specimens, eroded plaque showed a much higher density of MPO-positive cells in luminal thrombi compared with fissured plaque.[6] Endothelial cells overlying atherosclerotic lesions abundantly express the pattern-recognition receptor TLR2.[27] TLR2 has been localized mostly in areas of disturbed blood flow, and animal models have shown that the inhibition of this receptor reduces endothelial dysfunction by low shear stress. In lesions prone to erosion, lipoproteins and proteoglycans could serve as endogenous activators for TLR2.[28] Intense immunostaining for hyaluronan and its receptor, CD44, has been detected along the interface between luminal thrombus and eroded plaque compared with fissured or stable plaque.[28] These data led to the intriguing hypothesis that the accumulation of hyaluronan along the plaque may be involved, through CD44 activation, in the promotion of endothelial discontinuity, thrombus formation, and leukocyte accumulation. The ligation of TLR2 has been shown to result in inflammation and endothelial apoptosis in a process that could be markedly potentiated but not necessarily initiated by polymorphonuclear leukocytes. A “two-step model” has been proposed in which the first hit derives from TLR2 activation and leads to an initial endothelial injury, with subsequent apoptosis and desquamation; the second hit is mediated by neutrophil recruitment, with propagation of the local process and amplification of endothelial damage.[29] In humans, eroded plaques typically harbor abundant hyaluronan and proteoglycans, and TLR2 ligands typically include both endogenous and exogenous gram-positive derived hyaluronan. Thus it has been proposed that hyaluronic acid could be considered a disease-relevant TLR2 agonist, and that endogenous as well as infectious factors may operate to promote atherothrombosis through this mechanism.

Endothelial erosion occurs on plaques scarcely calcified, less prone to expansive remodeling, and only sparsely inflamed. However, neutrophil infiltration that typically follows plaque erosion is considered a hallmark of this distinct morphology of vulnerable plaques. No specific morphological features have been identified to distinguish eroded plaques from stable plaques and rupture-prone plaques, and both the limitation of imaging techniques and the relative lack of animal models have hampered progress in this field.

KEY CONCEPTS

Epidemiology of Ischemic Heart Disease

- Main cause of morbidity and mortality
- Recurrence of ACS eventually leads to the pandemics of heart failure and sudden cardiac death
- Need for a reappraisal of the mechanisms responsible for coronary instability and for innovative preventive and therapeutic strategies

Current Concepts on Atherosclerosis and Chronic Inflammation

- Atherosclerosis is a chronic non-resolving inflammatory disease
- Atherosclerotic plaques can remain stable indefinitely
- The mechanisms responsible for coronary instability are multiple

Pathogenic Classification of Acute Coronary Syndromes

- Plaque fissure with systemic inflammation
- Plaque fissure without systemic inflammation
- Plaque erosion
- Functional alterations of coronary circulation (epicardial or microvascular spasm)

Main Feature of Inflammation and Immunity Associated With Plaque Fissure

- Often associated with systemic inflammation
- Widespread coronary inflammation
- Innate immunity activation (monocytes, neutrophils, Pattern Recognition Receptors)
- Adaptive immunity alterations (unbalances between pro-inflammatory and anti-inflammatory activity)

Main Feature of Inflammation and Immunity Associated With Plaque Erosion

- Neutrophil accumulation
- Hyaluronan accumulation
- TLR2 activation

FUNCTIONAL ALTERATIONS OF CORONARY CIRCULATION

Epicardial coronary vasospasm is a possible cause of ACS when coronary angiography fails to demonstrate the presence of an obstructive atherosclerotic plaque.[1] The mechanisms underlying coronary spasm are multifold and may involve enhanced Rho kinase activity in smooth muscle cells.

Vasoconstriction causing ACS can be located also at the microvascular level. In particular, intense coronary microvascular vasoconstriction plays an important role in the pathogenesis of Takotsubo syndrome, which is diagnosed by ischemic pain at rest, ischemic ST-segment changes, and rise in cardiac enzymes. The syndrome also is characterized by a regional akinesia that more frequently affects distal myocardial regions but also is associated with hypercontractility of the remaining regions.

CLINICAL PERSPECTIVE

Precision medicine does not generally imply the creation of drugs or medical devices that are unique to a single patient. Instead, precision medicine signifies the ability to classify individuals into subpopulations that differ in their susceptibility to a particular disease, in the biology and/or prognosis of those diseases they may develop, or in their response to a specific treatment. Preventive or therapeutic interventions can then be concentrated on those who will benefit, sparing expense and side effects for those who will not. Indeed, coronary thrombosis is the final common pathway leading to ACS. More potent antithrombotic regimens have recently been found to increase the risk of major bleedings, which are associated in turn with a higher risk of mortality. Although the early outcome of ACS has considerably improved in the past decade, cardiovascular diseases still represent the main cause of morbidity and mortality overall. The rate of death, MI, and recurrent ACS at 1-year follow-up in contemporary NSTE-ACS registries is still very high, approaching about 25%.[30] Furthermore, recurrent ACS heralds a progressive impairment of myocardial function leading to chronic heart failure. Importantly, ischemic heart disease is by far the most important cause of heart failure and thus a public health problem of pandemic proportions. Thus the best prevention of heart failure is the prevention of recurrent episodes of coronary instability (Table 69.4).

CLINICAL RELEVANCE

- Different mechanisms and triggers are involved in ACS. The pathogenic classification of ACS might be a useful tool to select the best approach for different patients
- Histological studies have shown that ACS could be initiated by different types of unstable plaque
- Plaque fissure is strictly associated with systemic and coronary inflammation, with an altered cross-talk between innate and adaptive immunity leading to the unbalance between pro- and anti-inflammatory stimuli
- Plaque erosion rely on less clear mechanisms, however an inflammatory trigger might be also involved, with different triggers as compared with plaque fissure, such as hyaluronan exposure and PMN recruitment.

Plaque Fissure With Systemic Inflammation

The experience gained in the diagnosis, risk stratification, and management of chronic inflammatory diseases may provide important clues to the management of atherothrombosis. Extensive studies regarding helper T-cell subset dysregulation in ACS strongly support the notion that adaptive immunity is critically involved in the disease process. Investigating the mechanisms underlying T-cell dysregulation is pivotal to identifying new clinical biomarkers and potential therapeutic targets in the subset of ACS patients in whom a T cell–dependent inflammatory outburst is the likely cause of coronary instability. In this subset of ACS patients, raised levels of hs-CRP or of other soluble markers of inflammation are associated with a worse outcome.[4,10] A specific adaptive immune system signature proved to be more predictive than biomarkers in the identification of homogeneous subsets of patients at higher risk for acute coronary event recurrence.[16] A change of focus from soluble markers of inflammation to markers of adaptive immunity regulation, such as the molecules involved in TCR signaling, might be more rewarding.

TABLE 69.4 Pathogenical Mechanisms, Diagnosis, and Potential Tailored Therapy of Acute Coronary Syndromes

	Mechanism	Diagnosis	Potential Therapeutic Target
Plaque fissure with systemic inflammation	1. Innate immunity Early neutrophil activation Inflammatory macrophages Platelet activation (TLRs) 2. Adaptive immunity $CD4^+CD28^{null}$ T-cell expansion Treg imbalance Th17 cells B cells	Hs-CRP levels T-cell phenotyping Functional imaging	Immunosuppression (cyclosporine, tacrolimus, sirolimus, AZT, imatinib) TCR signaling regulation (synthetic CD31-derived peptide, PTPN22 inhibitor) IL-1β antagonists (anakinra, canakinumab, rilonacept) Treg expansion, anti-CD3-Ab (mice) Vaccines High-dose statins
Plaque fissure without systemic inflammation	1. Sympathetic nervous system activation 2. Changes in plaque composition	Anatomical imaging	Statins PLA_2 inhibitors Enhancement of cholesterol Efflux/impairment of foam cells formation (3-hydroxyanthranilic acid)
Plaque erosion	1. PMN producing MPO 2. Hyaluronan/CD44 interaction 3. TLR2 activation and endothelial injury 4. Acute local hemodynamic changes	MPO levels Functional imaging	Antithrombotic drugs
Functional alterations	1. Epicardial coronary spasm (enhanced Rho kinase activity) 2. Microvascular spasm	Functional tests of vasomotility Rho kinase activity	Nitrates Calcium antagonists Rho kinase inhibitors

AZT, azathioprine; Hs-CRP, high-sensitivity C-reactive protein; IL-1β, interleukin-1β; MMPs, matrix metalloproteinases; MPO, myeloperoxidase; PLA2, phospholipase A2TCR; PMN, polymorphonuclear neutrophil; PTPN22, protein tyrosine phosphatase non–receptor type 22; TCR, T-cell receptor; Th17, T-helper cell-17; TLRs, toll-like receptors.

An unmet need in this patient subset is a specific antiinflammatory treatment.[31,32] Indeed, TCR activation signaling offers several targets for immunosuppressive therapy such as cyclosporine, tacrolimus, sirolimus, azathioprine, and imatinib mesylate. These drugs prevent T-cell activation and reduce hyperreactivity of the adaptive immune system and development of autoimmunity. A synthetic CD31-derived peptide, able to engage a truncated extracellular CD31 fragment expressed by T cells that apparently lack CD31, has an immunosuppressive effect *in vivo* through restoration of the CD31 inhibitory pathway. The development of a novel series of inhibitors of PTPN22 might contribute to a better understanding of the role of this phosphatase in the immune dysregulation of ACS patients.

Another intervention target might be restoring the balance between effector T cells and Treg cells. Adoptive transfer of *ex vivo* expanded Treg might be considered. Of note, in mice anti-CD3 antibody treatment induced rapid regression of established atherosclerosis by reducing CD4 T cells and increasing the proportion of Treg.[33] Additionally, it has been shown that the tryptophan metabolite 3-hydroxyanthranilic acid has immune regulatory properties that can be used to decrease atherosclerosis in mice by regulating T cell–dependent inflammation and lowering plasma lipids. Finally, the identification of antigens that trigger adaptive immunity may open the way for specific vaccinations. Notably, strengthening antiviral immunity in at-risk patients, *e.g.*, by vaccination against influenza, has been associated with reduction in cardiovascular events, demonstrating the value of host protection through health maintenance.[34]

Plaque Fissure Without Systemic Inflammation

Clinical history of extreme emotional disturbance or intense physical exertion might help to identify this subset of patients. Anatomical (more than functional) features of the atherosclerotic plaque are important in determining coronary instability. Thus noninvasive imaging of stress-related plaque fissure might reveal a thinner cap compared with plaque fissure occurring at rest, indicating greater susceptibility to biomechanical forces. The development of micro-OCT that enables imaging of coronary artery microstructure on a scale comparable to that of histopathology might shed some new light on the role played by cholesterol crystallization. Inflammasome activation due to cholesterol crystals has been discussed as a potential mechanism.

Plaque Erosion

In plaque erosion, MPO is an important bystander and might even play a pathogenic role. Molecular imaging techniques can visualize MPO in atherosclerotic plaques in carotid arteries; these methods are more difficult, however, in coronary circulation. Elevated plasma concentration of MPO might represent an important biomarker for the identification of this subset of patients.[6] In patients with plaque erosion, the mechanism of inflammation is probably an intense local thrombogenic stimulus. Thus potent antithrombotic treatment, based perhaps on double antiaggregation and an oral anticoagulant, might be the treatment of choice, but this approach must be tested in prospective studies.

Functional Alterations of Coronary Circulation

Clinical history and 24- to 72-hr ambulatory ECG monitoring are usually sufficient to achieve the diagnosis of vasospastic angina, whereas the use of provocative tests of coronary artery spasm (*e.g.*, intravenous ergonovine, intracoronary ergonovine, or acetylcholine administration) is required in about 10% of patients. Rho kinase activity in circulating neutrophils is increased and might represent a useful biomarker for diagnosis and disease activity assessment. Although nitrates and calcium antagonists are helpful in patients with vasospastic angina, further efforts are needed to identify the molecular alterations responsible for smooth muscle cell hyperreactivity because a sizable proportion of these patients are refractory to standard doses of vasodilator.

Similarly, further efforts are warranted to identify a therapeutic target in Takotsubo syndrome.

CONCLUSION

As observed above, inflammation plays a detrimental role in the setting of ACS not only in determining plaque instability, but also in cardiac remodeling after myocardial necrosis and in PCI-related myocardial injury. The different mechanisms contributing to the inflammatory outbursts observed in ACS might need different forms of antiinflammatory treatment.

In addition to specific interventions targeting adaptive immunity in plaque fissure and systemic markers of inflammation as discussed above, IL-1 blockade might specifically address inflammation due to myocardial necrosis and prove to be useful in ventricular remodeling after myocardial necrosis, as suggested by the results of small pilot studies in STEMI patients.[31,32,34] Finally, IL-6 blockade might prevent inflammation associated with distal embolization in the setting of PCI in NSTEMI, resulting in less myocardial injury.[34]

ON THE HORIZON

Translational research is needed:

- To improve diagnostic criteria of vulnerable plaque
- To find novel predictive markers of plaque instability
- To find specific triggers linked to different types of unstable plaque
- To clarify the role of inflammation in plaque destabilization
- To evaluate alternative treatments and drug targets to prevent and treat plaque instability
    Toll-like receptor (TLR) antagonists
    Repair T-cell signaling abnormalities and CD31 signaling potentiation
    Restore Treg number and activity
    Target myeloperoxidase (MPO) in plaque erosion

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Crea F, Liuzzo G. Pathogenesis of acute coronary syndromes. J Am Coll Cardiol 2013;61:1–11.
2. Vergallo R, Papafaklis MI, Yonetsu T, et al. Endothelial shear stress and coronary plaque characteristics in humans: combined frequency-domain optical coherence tomography and computational fluid dynamics study. Circ Cardiovasc Imaging 2014;7(6):905–11.
3. Libby P, Ridker PM, Hansson GK. Progress and challenges in translating the biology of atherosclerosis. Nature 2011;473:317–25.
4. Liuzzo G, Biasucci LM, Gallimore JR. The prognostic value of C-reactive protein and serum amyloid a protein in severe unstable angina. N Engl J Med 1994;331:417–24.
5. Ong P, Athanasiadis A, Borgulya G, et al. 3-year follow-up of patients with coronary artery spasm as cause of acute coronary syndrome: the CASPAR (coronary artery spasm in patients with acute coronary syndrome) study follow-up. J Am Coll Cardiol 2011;57:147–52.
6. Niccoli G, Liuzzo G, Montone RA, et al. Advances in mechanisms, imaging and management of the unstable plaque. Atherosclerosis 2014;233:467–77.
7. Yonetsu T, Kakuta T, Lee T, et al. In vivo critical fibrous cap thickness for rupture-prone coronary plaques assessed by optical coherence tomography. Eur Heart J 2011;32:1251e9.
8. Bourantas CV, Garcia-Garcia HM, Farooq V, et al. Clinical and angiographic characteristics of patients likely to have vulnerable plaques: analysis from the PROSPECT study. JACC Cardiovasc Imaging 2013;6(12):1263–72.
9. Casas JP, Shah T, Hingorani AD, et al. C-reactive protein and coronary heart disease: a critical review. J Intern Med 2008;264: 295–314.
10. Biasucci LM, Koenig W, Mair J, et al. How to use C-reactive protein in acute coronary care. Eur Heart J 2013;34:3687–90.
11. Buffon A, Biasucci LM, Liuzzo G, et al. Widespread coronary inflammation in unstable angina. N Engl J Med 2002;347:5–12.
12. Narducci ML, Grasselli A, Biasucci LM, et al. High telomerase activity in neutrophils from unstable coronary plaques. J Am Coll Cardiol 2007;50:2369–74.
13. Wyss CA, Neidhart M, Altwegg L, et al. Cellular actors, Toll-like receptors, and local cytokine profile in acute coronary syndromes. Eur Heart J 2010;31:1457–69.
14. Liuzzo G, Montone RA, Gabriele M, et al. Identification of unique adaptive immune system signature in acute coronary syndromes. Int J Cardiol 2013;168:564–7.
15. Libby P, Nahrendorf M, Pittet MJ, et al. Diversity of denizens of the atherosclerotic plaque: not all monocytes are created equal. Circulation 2008;117:3168–70.
16. Niessner A, Shin MS, Pryshchep O, et al. Synergistic proinflammatory effects of the antiviral cytokine interferon-alpha and Toll-like receptor 4 ligands in the atherosclerotic plaque. Circulation 2007;116: 2043–52.
17. Liuzzo G, Biasucci LM, Trotta G, et al. Unusual CD4+CD28null T lymphocytes and recurrence of acute coronary events. J Am Coll Cardiol 2007;50:1450–8.
18. Giubilato S, Liuzzo G, Brugaletta S, et al. Expansion of CD4+CD28null T-lymphocytes in diabetic patients: exploring new pathogenetic mechanisms of increased cardiovascular risk in diabetes mellitus. Eur Heart J 2011;32:1214–26.
19. Mauri C, Bosma A. Immune regulatory function of B cells. Annu Rev Immunol 2012;30:221–41.
20. Liuzzo G, Vallejo AN, Kopecky SL, et al. Molecular fingerprint of interferon-gamma signalling in unstable angina. Circulation 2001;103:1509–14.
21. Liuzzo G, Trotta F, Pedicino D. Interleukin-17 in atherosclerosis and cardiovascular disease: the good, the bad, and the unknown. Eur Heart J 2013;34:556–9.
22. Klingenberg R, Brokopp CE, Grivès A, et al. Clonal restriction and predominance of regulatory T cells in coronary thrombi of patients with acute coronary syndromes. Eur Heart J 2015;36:1041–8.
23. Flego D, Severino A, Trotta F, et al. Increased PTPN22 expression and defective CREB activation impair regulatory T-cell differentiation in non-ST-segment elevation acute coronary syndromes. J Am Coll Cardiol 2015;65:1175–86.
24. Biasucci LM, Liuzzo G, Ciervo A, et al. Antibody response to chlamydial heat shock protein 60 is strongly associated with acute coronary syndromes. Circulation 2003;107:3015–17.
25. Flego D, Severino A, Trotta F, et al. Altered CD31 expression and activity in helper T cells of acute coronary syndrome patients. Basic Res Cardiol 2015;109:448.
26. Virmani R, Kolodgie FD, Burke AP, et al. Lessons from sudden coronary death: a comprehensive morphological classification scheme for atherosclerotic lesions. Arterioscler Thromb Vasc Biol 2000;20: 1262–75.
27. Mullick AE, Soldau K, Kiosses WB, et al. Increased endothelial expression of Toll-like receptor 2 at sites of disturbed blood flow exacerbates early atherogenic events. J Exp Med 2008;205:373–83.
28. Kolodgie FD, Burke AP, Farb A, et al. Differential accumulation of proteoglycans and hyaluronan in culprit lesions: insights into plaque erosion. Arterioscler Thromb Vasc Biol 2002;22:1642e8.
29. Quillard T, Araujo HA, Franck G, et al. TLR2 and neutrophils potentiate endothelial stress, apoptosis and detachment: implications for superficial erosion. Eur Heart J 2015;36:1394–404.

30. Roffi M, Patrono C, Collet JP, et al. ESC Guidelines for the management of acute coronary syndromes in patients presenting without persistent ST-segment elevation: Task Force for the Management of Acute Coronary Syndromes in Patients Presenting without Persistent ST-Segment Elevation of the European Society of Cardiology (ESC). Eur Heart J 2016;37:267–315.
31. Della Bona R, Liuzzo G, Pedicino D, et al. Anti-inflammatory treatment of acute coronary syndromes. Curr Pharm Des 2011;17:4172–89.
32. Ridker PM, Lüscher TF. Anti-inflammatory therapies for cardiovascular disease. Eur Heart J 2014;35:1782–91.
33. Kita T, Yamashita T, Sasaki N, et al. Regression of atherosclerosis with anti-CD3 antibody via augmenting a regulatory T-cell response in mice. Cardiovasc Res 2014;102:107–17.
34. Crea F, Liuzzo G. Anti-inflammatory treatment of acute coronary syndromes: the need for precision medicine. Eur Heart J 2016;37:2414–16.

MULTIPLE-CHOICE QUESTIONS

1. Indicate which of the following represents the correct timeline in the natural history of atherosclerosis:
   - **A.** Subendothelial retention of lipoproteins and flow-mediated inflammatory changes → endothelial dysfunction → self-maintaining inflammatory state → fatty streak → atheroma→ fibroatheroma
   - **B.** Fatty streak → endothelial dysfunction → self-maintaining inflammatory state → subendothelial retention of lipoproteins and flow-mediated inflammatory changes → atheroma→ fibroatheroma
   - **C.** Fibroatheroma → endothelial dysfunction → fatty streak → self-maintaining inflammatory state → subendothelial retention of lipoproteins and flow-mediated inflammatory changes → atheroma
   - **D.** Self-maintaining inflammatory state → atheroma → fibroatheroma → endothelial dysfunction → fatty streak → subendothelial retention of lipoproteins and flow-mediated inflammatory changes
2. Which of the following represents a recognized mechanism involved in plaque instability?
   - **A.** Increase in regulatory T cells (Tregs) around the plaque
   - **B.** Widespread coronary inflammation
   - **C.** Smooth muscle cells hyperreactivity
   - **D.** Metabolic disturbances
3. The mechanisms involved in plaque erosion are still poorly understood; according to recent experimental evidence, which of the following seems to be the most probable trigger involved in endothelial erosion?
   - **A.** Inflammasome activation through cholesterol crystals inside the plaque
   - **B.** Coronary artery dissection
   - **C.** Fibrous cap rupture
   - **D.** Toll-like receptor (TLR) activation, extracellular matrix and hyaluronan exposure, neutrophil recruitment

70

Autoimmune Thyroid Diseases

Anna L. Mitchell, Simon H.S. Pearce

Autoimmune thyroid disorders are common, and, indeed, as a group, they are the most prevalent autoimmune disorders in humans. Despite many common underlying features, such as a marked female preponderance, shared susceptibility alleles, and common autoantigens, Graves disease and autoimmune hypothyroidism have contrasting clinical characteristics. Over recent years, our knowledge about the underlying pathogenesis has increased as a result of advances in human genomics, molecular immunology, and the availability of murine models of disease. Novel therapies based on this increased understanding are now emerging.

KEY CONCEPTS

Classification of Autoimmune Thyroid Disease

Autoimmune hyperthyroidism	Graves disease
Autoimmune thyroiditis	Hashimoto thyroiditis Atrophic thyroiditis Postpartum thyroiditis

GRAVES HYPERTHYROIDISM

Graves disease (GD) is a common autoimmune condition that accounts for the majority of cases of hyperthyroidism in the developed world. Its pathogenesis is unique among the autoimmune endocrinopathies because a key feature is the presence of stimulating autoantibodies directed against the thyrotropin (thyroid-stimulating hormone [TSH]) receptor, which mimic the action of TSH, a native hormone produced in the pituitary, to drive thyroid overactivity. Interestingly, thyroid dysfunction is commonly associated with other extrathyroidal manifestations of GD, the most common being Graves ophthalmopathy (GO).

Epidemiology

GD is one of the most common autoimmune diseases, with a prevalence of approximately 1% in women in the developed world.[1] It is more common in iodine-sufficient countries, where it accounts for 60–90% of cases of hyperthyroidism.[2] GD is seven times more common in women than in men and may affect individuals at any age; however, the peak incidence occurs between the ages of 35 and 40 years.

Etiology

GD is a complex genetic condition, implying that environmental stimuli precipitate disease in genetically predisposed individuals who harbor multiple susceptibility alleles. A large Danish twin study estimated that about 80% of the propensity to develop GD is attributable to genetic factors.[3] Further evidence for the heritability of GD comes from the observation that it clusters within families. Up to one-quarter of individuals with GD have a first-degree relative with the condition or with another autoimmune thyroid disease, such as autoimmune hypothyroidism.[4] Should an individual have a sibling with GD, it is estimated that the relative risk (λ_s) of that individual developing GD is around 10, which is comparable with that of other heritable autoimmune conditions, such as type 1 diabetes, which has a λ_s of 15.

A number of genetic loci have been shown to contribute to GD susceptibility (Fig. 70.1). These genes encode proteins in biological pathways that regulate immune system activity or thyroid biology.[4,5] The major histocompatibility complex (MHC) region on chromosome 6p21 has long been associated with multiple autoimmune conditions. Human leukocyte antigen (HLA) genes found within the MHC region play a vital role in pathogen and self-peptide recognition and therefore have a clear role in immunity and in establishing and maintaining immune tolerance (Chapter 5).

In European populations, the primary association between MHC and GD is with alleles of the class II MHC genes. The HLA-DR3 allele is detected twice as frequently in subjects with GD as in healthy controls (*i.e.,* 50% of GD subjects vs 25% of controls). At the protein level, neutral amino acids alanine or glutamine are substituted for positively charged arginine at position 74 in the HLA-DR peptide-binding pocket, which is thought to alter the binding-pocket configuration, more readily allowing self-peptides to enter the antigen binding site.[5] Importantly, 50% of individuals with GD do not have the *HLA-DR3* allele, implying that there is unlikely to be a single antigenic epitope responsible for GD.

The cytotoxic T lymphocyte antigen-4 (*CTLA4)* gene (chromosome 2q33) encodes a costimulatory molecule expressed on the surface of activated T cells, which plays a pivotal role in downregulating T-cell responses and in checking T-cell activation, emphasizing the contribution of inhibitory signals in setting immune response thresholds (Chapter 18). The *CT60* single nucleotide polymorphism (SNP) downstream from the 3' untranslated region (3'UTR) was found to influence GD susceptibility (odds ratio [OR] 1.5) and has been suggested as a possible etiological variant[6]; however, the functional effects of this variant remain poorly defined and contributions from other variants, such as at codon 17 in the CTLA-4 signal peptide remain likely. About 50% of individuals from healthy European populations carry the autoimmune "susceptible" *CTLA4* haplotype; therefore other important factors are clearly at play. CTLA-4

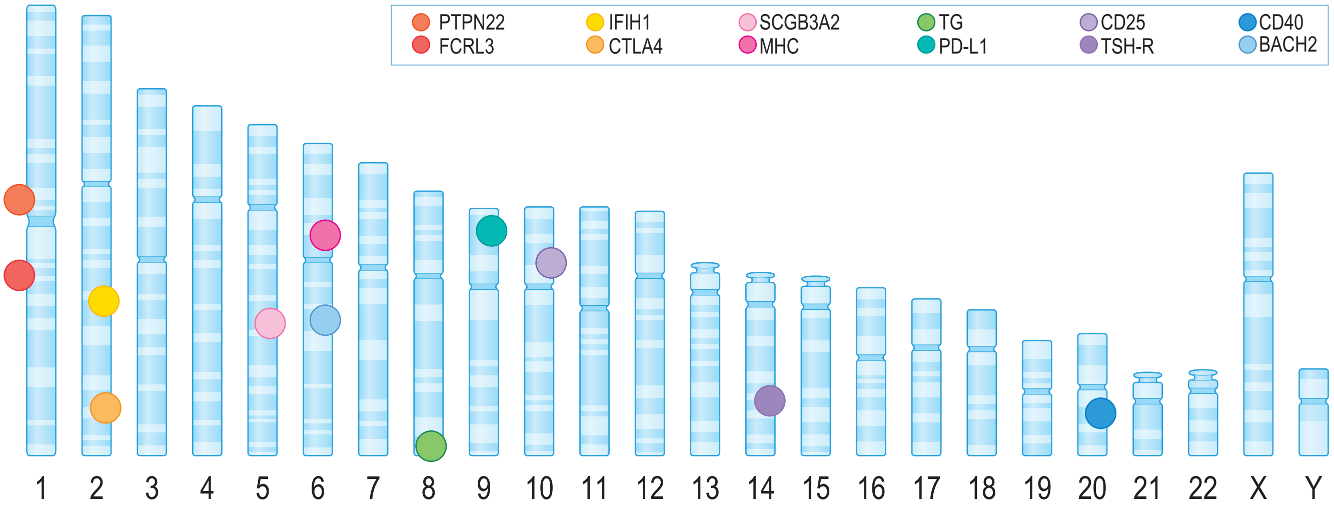


FIG 70.1 Schematic diagram to illustrate the loci that have been associated with Graves disease to date. Each locus is shown on its respective chromosome, with chromosome 1 depicted on the left and chromosome Y on the right.

polymorphisms also contribute to susceptibility to type 1 diabetes, autoimmune adrenal insufficiency, celiac disease, and several other autoimmune conditions.

The protein tyrosine phosphatase, non–receptor type 22 (*PTPN22)* gene on chromosome 1p13 encodes the lymphoid tyrosine phosphatase (LYP) molecule, which, like CTLA-4, is involved in the regulation of T-cell activation. In GD, a coding variant (R620W) was found more frequently in Caucasians with GD than in controls (13 vs 7% respectively, OR 1.8).[7] This variant results in a short-lived protein product, which is believed to alter T-cell receptor signaling compared with the wild-type protein. The PTPN22*620 W variant is very rare in non-Caucasian populations, but different *PTPN22* alleles have been associated with autoimmunity in Asian cohorts.

Products of several other genes involved in immunoregulation have been demonstrated to have allelic variants associated with GD, including CD25, CD40, PD-L1, IFIH1, BACH2, SCGB3A2, and FCRL3. In addition to these variants in immune regulatory pathways, loci specific to GD have been identified on the basis of known thyroid pathophysiology. The gene encoding the TSH receptor (TSH-R) on chromosome 14q31 is an obvious candidate for GD, as the TSH-R is directly stimulated by autoantibodies in affected individuals. Although initial studies reported conflicting results, a definite association between a number of SNPs in intron 1 of the *TSHR* gene and GD has now been confirmed in Caucasians.[5] The mechanism by which these intronic SNPs confer disease susceptibility is unknown. Another GD candidate is the *thyroglobulin* (*TG*) gene (8q24). This encodes the thyroglobulin Tg protein, the precursor for the thyroid hormones triiodothyronine (T_3) and thyroxine (T_4). SNPs in the *TG* gene have been associated with GD in some Caucasian cohorts, and an epistatic interaction between a SNP in *TG* exon 33 and HLA-DR has been suggested. This interesting interaction has raised the possibility that Tg polymorphisms predispose to GD by modulating Tg peptide presentation by antigen-presenting cells on HLA class II molecules to T cells.

Despite years of research into the genetic etiology of GD, 30–40% of the inherited susceptibility has yet to be accounted for. This "hidden" heredity is a common theme in genetically complex traits, but it is likely to be owing to rare genomic variants, polymorphisms in regulatory DNA sequences, including noncoding or microRNAs (miRNAs), and/or epigenetic factors.

GD is one of the few autoimmune conditions for which links to environmental factors have been definitively established.[8] Iodine is one of the most common precipitants of thyroid dysfunction. With regard to GD, more cases are observed in iodine-sufficient areas. A study of individuals in Iceland, where iodine intake is high, and age-matched individuals in East Jutland, Denmark, where iodine intake is low, showed that the incidence of GD was more than double in the higher iodine environment.[2]

Cigarette smoking also influences GD susceptibility and severity, in particular of GO. Meta-analysis of eight studies showed the OR for developing GD was 3.3 for current smokers compared with that for lifelong nonsmokers.[9] The same study also revealed that current smokers are more likely to develop GO compared with nonsmokers (OR 4.4). The mechanism for this interaction remains to be fully elucidated; however, cigarette smoke extract has been shown, *in vitro,* to stimulate adipogenesis and glycosaminoglycans (GAGs) production by orbital fibroblasts, which accumulate in the orbital tissues in GO.[10]

Stress appears to influence both GD disease onset and clinical course. In one study, individuals with GD retrospectively reported more negative life events in the preceding year compared with matched controls.[11] This finding is replicated even when a group of "nonautoimmune" individuals with hyperthyroidism are used as controls, suggesting that stress is truly a precipitant for autoimmune GD.[12]

Changes in immune system function appear to influence the onset of GD. During pregnancy, which is a relatively immunosuppressed state, hyperthyroid GD is often mild and can be managed with smaller doses of antithyroid drugs. In some cases, it remits entirely, allowing the individual to stop medication in the short term. However, in the postpartum period, when the immune system normalizes, there is typically worsening or relapse of GD. A similar phenomenon is seen in individuals who have been significantly immunosuppressed and then have recovered. For example, new-onset GD has been reported in people who have been successfully treated with highly active antiretroviral therapy (HAART) for human immunodeficiency virus (HIV) infection. During treatment, as the immune system recovers, immune

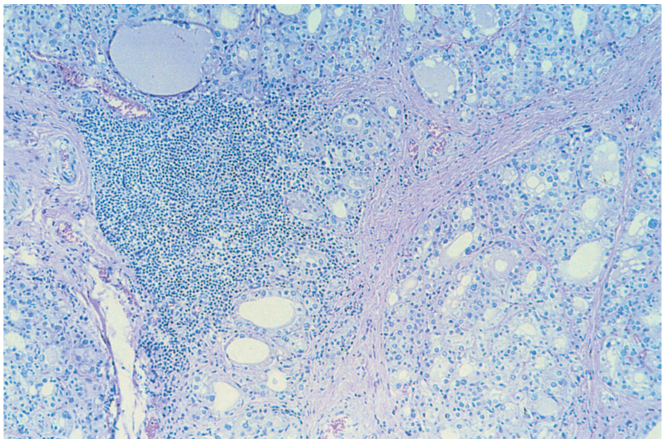


FIG 70.2 Diffuse lymphocyte infiltrate and thyrocyte hyperplasia in a patient with Graves disease.

activation increases. T cells are exposed to thyroid antigens, resulting in an autoimmune response at the time of immune reconstitution.[13] A similar phenomenon has been seen in individuals with multiple sclerosis (MS) treated with alemtuzumab, the lymphocyte-depleting anti-CD52 antibody. An alternative explanation is the weakening of physiological antiinflammatory pathways, which unleashes the immune system. As more novel biological agents become available, this phenomenon is likely to become more common.

KEY CONCEPTS

Environmental Factors Known to Influence Graves Disease Susceptibility

- Smoking
- Iodine
- Stress
- Immune system reconstitution states
    - Post-partum state
    - Successful treatment of human immunodeficiency virus (HIV) with highly active antiretroviral therapy (HAART)
    - T-cell depletion therapy (*e.g.*, alemtuzumab)
- Infections (*e.g.*, hepatitis C)

Immunopathogenesis

Histologically, the thyroid in GD is characterized by a diffuse lymphocytic infiltrate, consisting of both T and B cells, associated with thyrocyte hyperplasia (Fig. 70.2). Although T cells play a major role in inflammatory cell recruitment, cytokine secretion, antigen recognition, and thyrocyte damage, infiltrating B cells also produce antibodies, including those that drive hyperthyroidism. The major autoantigens in GD are the TSH-R, the thyroid peroxidase (TPO) enzyme, and Tg. More than 95% of patients with GD have detectable circulating TSH-R autoantibodies,[14] which are necessary for hyperthyroidism, and approximately 90% have detectable TPO autoantibodies.[15] Antibodies directed against the sodium iodide symporter and the apical iodide transporter, pendrin, have been reported in smaller numbers of patients.[15]

GD is unique among autoimmune conditions in that the TSH-R autoantibodies directly stimulate thyroid gland activity. This is exemplified by neonatal GD, where maternal TSH-R autoantibodies cross the placenta, resulting in transient hyperthyroidism in the newborn. However, although TSH-R autoantibodies are sufficient to result in transient hyperthyroidism in these infants, these autoantibodies are not sufficient, *per se*, to result in the persistent thyroid autoreactivity of true GD. TSH-R antibodies are classically immunoglobulin G1 (IgG1) subclass and target an epitope in the aminoterminal region of the leucine-rich repeat motif in the extracellular domain of the TSH-R (Fig. 70.3).[16] When the autoantibody binds to the TSH-R, this activates intracellular G proteins, which, in turn, induce transcription of genes, such as *TPO* and *TG*, via the cyclic adenosine monophosphate (c-AMP) and phospholipase-C pathways. This results in thyrocyte hyperplasia and increased thyroid hormone synthesis. The TSH-R antibody–induced expression of TPO and Tg, which are also thyroid antigens, may be a mechanism for disease perpetuation. TSH-R autoantibodies can also be “blocking” in nature and prevent receptor activation. This can result in hypothyroidism. These two types of autoantibodies can also coexist, resulting in fluctuating thyroid function. Thus although it is common to equate GD with hyperthyroidism, individuals with GD may occasionally be hypothyroid or euthyroid (presenting with GO).

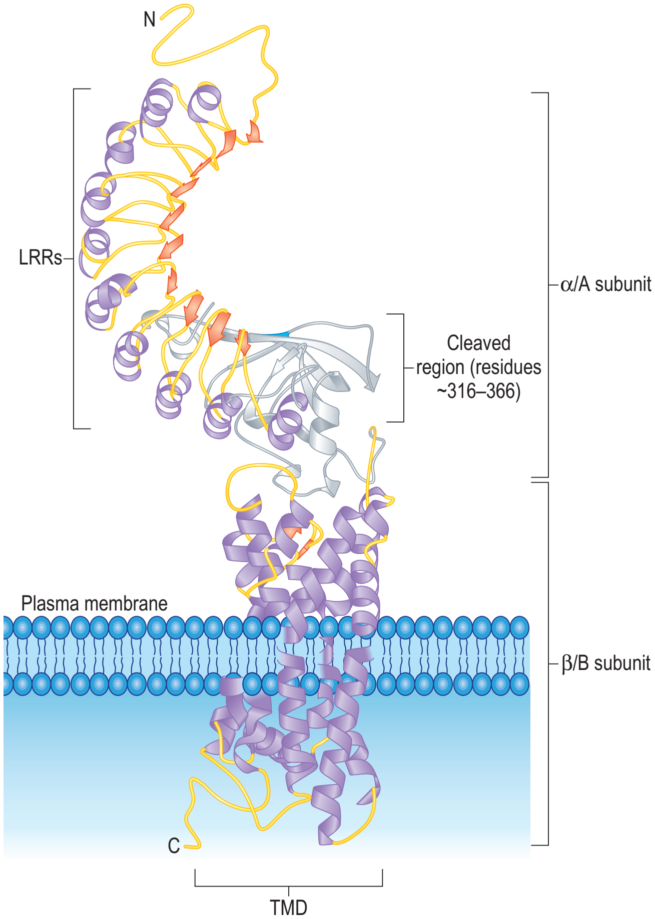


FIG 70.3 Structure of the thyroid-stimulating hormone receptor (TSH-R). (Courtesy of R. Latif; adapted from Davies TF, Ando T, Lin RY, et al. Thyrotropin receptor-associated diseases: from adenomata to Graves’ disease. J Clin Invest 2005;115: 1972–83.)

TPO antibodies can be of IgG subclass 1, 2, or 4 and typically circulate in concentrations 1000-fold higher than those of TSH-R antibodies. They are directed against two structurally complex regions of the TPO molecule, the epitopes involving residues from both the myeloperoxidase-like and the complement control protein-like domains. TPO antibodies may have pathogenic significance in that they can fix complement and target the thyrocyte for cell-mediated cytotoxicity.[17] In contrast to TSH-R antibodies, they do not appear to either stimulate or block the enzymatic activity of TPO.

Clinical Presentation

Hyperthyroid GD can present with manifestations affecting almost any organ system in the body, and as with many endocrine conditions, affected individuals may report a gradual onset of nonspecific symptoms, typically over a period of months. This often leads to a delay in seeking medical attention and in the initial diagnosis being made. The signs and symptoms of GD can be divided into those associated with hyperthyroidism in general and those specific to GD. These are summarized in Table 70.1.

TABLE 70.1 Common and Rarer Clinical Manifestations of Hyperthyroidism

Common	Rare
Neuropsychiatric	
Anxiety	Chorea
Fatigue and exhaustion	Collapse (periodic paralysis)
Fine tremor	Pseudobulbar palsy
Restlessness and fidgeting	Spasticity
Gastrointestinal	
Increased appetite	Hepatosplenomegaly
Loose stools	
Increased frequency of defecation	
Nausea	
Weight loss	
Cardiorespiratory	
Palpitations	Congestive cardiac failure
Shortness of breath on exertion	
Tachycardia (sinus, atrial fibrillation)	
Peripheral vasodilation, flushing	
Systolic hypertension	
Genitourinary	
Menstrual irregularities	
Cutaneous	
Itch	Thyroid acropachy
Heat intolerance	Pretibial myxedema
Hair loss	Onycholysis
Musculoskeletal	
Hyperreflexia	
Proximal muscle weakness	
Ophthalmic	
Lid lag	Optic neuropathy
Lid retraction	
Exophthalmos and proptosis	
Eye dryness	
Chemosis	
Ophthalmoplegia	
Miscellaneous	
Thirst	
Thyroid bruit	

CLINICAL PEARLS

Clinical Signs Specific to Graves Disease

- Graves orbitopathy
- Thyroid bruit
- Thyroid acropachy
- Pretibial myxedema

Investigation and Diagnosis of GD

The diagnosis of GD is a clinical one, supported by laboratory investigations. Imaging is occasionally required if the diagnosis is in doubt. Thyrotoxicosis is diagnosed biochemically on the basis of an elevated serum free T_3 (fT_3) or free T_4 (fT_4) in the presence of a completely suppressed TSH. If thyrotoxicosis is found in a patient with extrathyroidal signs, such as GO or pretibial myxedema, the diagnosis of hyperthyroid GD is clear and further investigations are not required. If no extrathyroidal signs are present, the presence of serum autoantibodies should be sought. TSH-R antibodies are highly sensitive for GD, and assays are now available in many laboratories. The TSH-binding inhibitory immunoglobulin (TBII) assay is also commonly used, which is an indirect means of detecting TSH-stimulating antibodies. TPO antibodies are often measured as a surrogate for thyroid autoimmunity, as the presence of either TSH-R antibodies or TPO antibodies has more than 90% sensitivity for GD.

Imaging is reserved for individuals in whom the diagnosis is not clear. Radionuclide scanning, for example, ^{99}Tc or ^{123}I, is favored over ultrasonography, as the former gives functional information on the activity of the thyroid gland, although Doppler flow studies are making ultrasound evaluation of the thyroid increasingly informative (Fig. 70.4). In GD, there is diffuse uptake in the thyroid gland on radionuclide scanning.

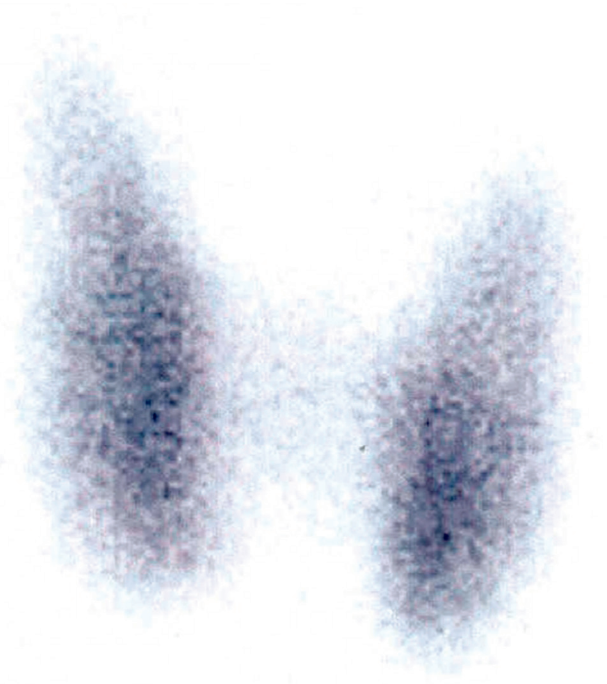


FIG 70.4 ^{99}Tc pertechnetate radionuclide scan image from an individual with Graves disease showing diffuse uptake throughout the thyroid gland.

Management of GD

The management of hyperthyroid GD can be divided into three broad categories: medical management, radioiodine (RAI) treatment, and surgery.[18]

Medical Management—Antithyroid Drugs

The thionamide drugs (carbimazole, its metabolite methimazole, and propylthiouracil) compete with Tg to act as substrates for iodination by TPO. Once iodinated, they are metabolized peripherally, depleting thyroid iodine stores. When the thyroid iodine stores are depleted and the intrathyroidal thionamide concentration is high enough, thyroid hormone synthesis is abrogated. Most individuals become euthyroid following 4–8 weeks of treatment; however, euthyroidism may take longer to achieve in those with poor medication compliance or with a history of recent iodide exposure. Following initial treatment, thionamides may either be administered as a fixed high dose, with levothyroxine supplementation to prevent hypothyroidism (known as a "block and replace" regimen), or at progressively lower doses, titrated to allow adequate thyroid hormone generation. Following 6–18 months of thionamide treatment, about 50% of patients will remain in remission following cessation of therapy. There is no improvement in remission rate in individuals with GD who are treated for longer than 18 months. The mechanism of the thionamide-induced remission in GD remains obscure; however, the lymphocytic infiltrate in the GD thyroid is rapidly abolished by thionamide treatment, and serum TSH-R and TPO autoantibodies also decrease during treatment, suggesting an immunomodulatory effect. It is telling that several induced murine models of thyroiditis can be ameliorated by thionamide treatment, suggesting an immunomodulatory role that is distinct from the antithyroid hormone synthesis action. If the patient's disease relapses following medical treatment, definitive treatment should be considered, as a second course of thionamide treatment rarely induces prolonged remission.

Definitive Treatment

Radioactive iodine (RAI) therapy (^{131}I) and surgical thyroidectomy are both efficacious treatments for GD, effectively removing functional thyroid tissue and rendering the patient hypothyroid and dependent on levothyroxine replacement over the long term. RAI therapy takes advantage of the thyroid's ability to concentrate iodine. RAI is administered orally and concentrates in the thyroid, primarily emitting beta-radiation to produce local thyrocyte DNA damage. These cells then undergo necrotic change, and the thyroid atrophies over the subsequent year. RAI renders more than 80% of those with GD hypothyroid and needing long-term thyroid hormone replacement. RAI is generally well tolerated, and long-term follow-up studies in adults have shown reassuring results with regard to carcinogenicity. RAI is absolutely contraindicated in pregnancy and is best avoided in those with active, inflammatory GO. Total or near-total thyroidectomy is also an effective treatment for GD and is particularly suitable for individuals with a large goiter, those with severe hyperthyroidism who cannot tolerate thionamides, or in those with active GO when RAI is relatively contraindicated. The complications of thyroidectomy include change in voice because of intraoperative damage to the recurrent laryngeal nerve and hypocalcemia (often transient) caused by parathyroid gland damage.

GRAVES OPHTHALMOPATHY

Epidemiology

GO is the most common extrathyroidal manifestation of GD. It is clinically apparent in 25–50% of those presenting with GD, although almost all patients have radiological changes consistent with GO to some degree. The peak age of incidence is in the fifth and seventh decades of life. GO precedes thyroid gland dysfunction in about 20% of patients, arises at the same time in 40%, and occurs after diagnosis of thyroid dysfunction in the remaining 40%. Males, older adults, and smokers are more likely to have severe disease. Ninety-three percent of cases of GO occur in those with hyperthyroid GD. However, 3% and 4% of GO occurs in hypothyroid and euthyroid patients, respectively. A proportion of these patients with "euthyroid Graves" will eventually become hyperthyroid; however, some will remain hypothyroid or euthyroid despite the presence of TSH-R antibodies.

Etiology

GO shares many etiological factors with GD. Smoking is the major environmental risk factor for GO, with smokers or ex-smokers four times more likely to develop GO compared with lifelong nonsmokers. The number of cigarettes smoked per day correlates with the risk of developing GO, suggesting a "dose-dependent" effect. RAI therapy is also known to occasionally cause flare-ups of GO. This is thought to be largely related to uncontrolled post-RAI hypothyroidism. Smokers and those with active GO seem particularly susceptible to this complication, and therefore RAI is relatively contraindicated in patients with active GO.[19]

Immunopathogenesis

The molecular mechanism that links thyroid dysfunction with GO remains incompletely understood. The TSH-R is widely believed to be the primary autoantigen linking the thyroid and the orbit. Orbital fibroblasts (OFs) have been shown to express cell-surface TSH-R, and these are thought to be functional because of the observation that TSH stimulation *in vitro* results in an increase in intracellular cAMP. Mechanistically, the orbital changes that occur in GO are better understood. TSH-R autoantibodies that bind to the TSH-R on OFs (or possibly other OF receptors, such as the insulin-like growth factor receptor; IGF-1R) are believed to activate OFs to secrete cytokines and chemokines, which attract lymphocytes and other inflammatory cells. These infiltrate the orbital tissues, augmenting further the proinflammatory cytokine environment, causing OFs to proliferate and to secrete excessive GAGs.[20] GAGs accumulate in the extraocular muscles, increasing their size. These matrix molecules are also osmotically active, resulting in edema and swelling of surrounding tissues. In addition, the adaptive arm of the immune system is thought to interact via HLA and CD40 molecules expressed by the OFs to increase the presentation of autoantigens, thus perpetuating the cycle. OFs are also thought to differentiate into adipocytes, resulting in increased retroorbital fat deposition. The resulting inflammation gives rise to redness and edema of the orbital tissues, with the cellular proliferation and GAG accumulation producing proptosis (protrusion of the eyeball from the socket), increased intraorbital pressure, and restriction of extraocular eye movements.[20]

Diagnosis and Clinical Presentation

When the clinical signs of GO are associated with thyroid dysfunction and/or circulating thyroid autoantibodies, the diagnosis is

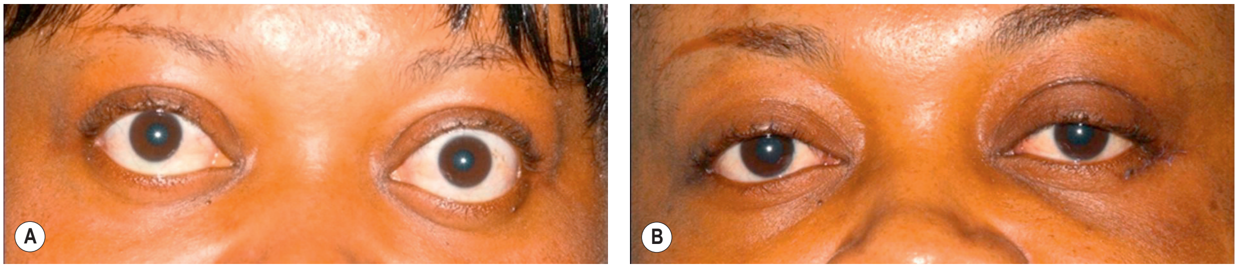


FIG 70.5 (A) Clinical photograph of the eyes of an individual with Graves ophthalmopathy prior to treatment. (B) Clinical photograph of the same patient's eyes following surgical orbital decompression and rehabilitative surgery.

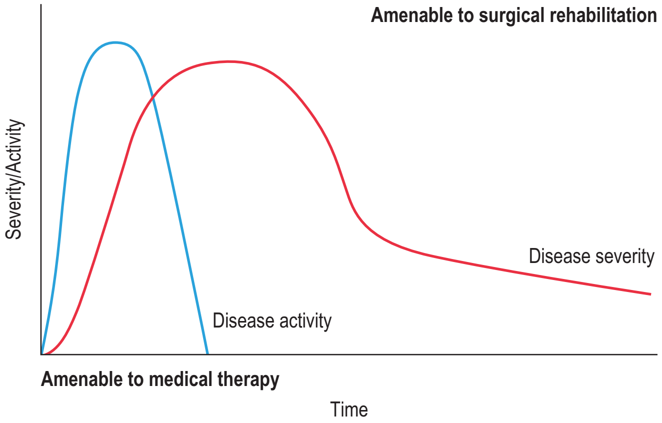


FIG 70.6 Rundle curve illustrating the natural history of Graves ophthalmopathy.

TABLE 70.2 Assessment of Graves Ophthalmopathy (GO)

A: The "NO SPECS" Classification System to Assess the Severity of GO

Class 0—No signs or symptoms
Class 1—Only signs (limited to upper lid retraction and stare, with or without lid lag)
Class 2—Soft tissue involvement (edema of conjunctivae and lids, conjunctival injection, etc.)
Class 3—Proptosis
Class 4—Extraocular muscle involvement (usually with diplopia)
Class 5—Corneal involvement (primarily due to lagophthalmos, the inability to completely close the eyelids)
Class 6—Sight loss (caused by optic nerve involvement)

B: The Clinical Activity Score (CAS) to Assess Activity of GO

A single point is scored for each of the features present. Each feature has equal weighting. A higher score indicates more active disease.

1. Spontaneous orbital pain
2. Gaze evoked orbital pain
3. Eyelid swelling that is attributed to active GO
4. Eyelid erythema
5. Conjunctival redness that is considered to be due to active GO
6. Chemosis
7. Inflammation of caruncle or plica
8. Increase of >2 mm in proptosis
9. Decrease of >8 degrees in uniocular ocular excursion in any one direction
10. Decrease of acuity equivalent to 1 Snellen line

generally straightforward. If proptosis is completely unilateral, or GO features occur without upper lid retraction, then the diagnosis needs to be confirmed by imaging, as the differential diagnosis for unilateral proptosis is a space-occupying retroorbital lesion. The patient with GO may complain of gritty, watery, or uncomfortable eyes, with or without a change in appearance (upper eyelid retraction, soft tissue swelling, redness of the eyes, and proptosis) (Fig. 70.5). Diplopia occurs if eye movements are restricted by stiffness of the extraocular muscles or high intraorbital pressure. Deteriorating visual acuity and color desaturation are sinister symptoms in GO, indicating incipient optic neuropathy.

GO has a predictable and generally monophasic natural history (Fig. 70.6). There is an early phase of increasing disease activity that can be targeted by medical therapy, and it is followed by a plateau phase and then gradual improvement until a stable, inactive phase is reached. GO manifestations can be classified in two ways: on the basis of severity, which indicates the extent of functional, anatomical, and cosmetic features; and on the basis of activity, which denotes the intensity of any acute inflammatory reaction. Severity of GO is assessed using the "NO SPECS" classification system (Table 70.2, panel A) and activity is assessed using the clinical activity score (CAS) (see Table 70.2, panel B). A CAS of 3 or more indicates active GO.[19] These classifications allow determination of which patients require treatment and which therapy is most appropriate.

Treatment

Patients with GD (with and without GO) should be strongly encouraged to stop smoking and offered smoking cessation support. The general health benefits of smoking cessation are numerous; in addition, smokers are more likely to relapse after a course of medical therapy compared with nonsmokers.

In parallel to restoring and maintaining euthyroidism, the successful treatment of GO depends on staging the activity and severity of the disease. In patients with mild, active GO, an observational policy can be employed with symptomatic measures, such as artificial tears and dark glasses. Selenium supplements, at a dose of 100 μg twice daily for 6 months, have been shown to significantly improve quality of life, slow disease progression, and reduce ocular involvement in these patients.[21] In those with moderate to severe, active, and progressive GO, a course of oral or intravenous (IV) glucocorticoids (steroids) is indicated. Orbital radiotherapy is also efficacious in people with active inflammation and diplopia, but since it has a delayed onset of action, it often needs to be used in conjunction with other therapies, such as steroids or orbital decompression surgery. In patients with sight-threatening optic neuropathy, high-dose IV steroids are used, and orbital pressure is relieved by urgent orbital decompression surgery. In patients whose eyelids do not close completely, eye ointments and protective eye pads are essential to protect the eyes against corneal damage and ulceration. Once disease activity has burned out, rehabilitative surgery can greatly improve the function and cosmetic appearance of the eyes. Orbital decompression, strabismus correction, and eyelid surgery are commonly used procedures.[22]

FUTURE DEVELOPMENTS FOR GRAVES HYPERTHYROIDISM AND OPHTHALMOPATHY

Novel approaches to modulating the immune response in GD and GO as a therapeutic strategy are under investigation. In light of the significant side effects associated with steroid therapy for GO, steroid-sparing agents, including methotrexate and mycophenolate mofetil, are of particular interest. Novel biological agents have been evaluated to a limited extent in these conditions and are subject to further ongoing studies.[23] Rituximab, a CD20 monoclonal antibody (mAb) that depletes circulating B cells, appeared to be potentially efficacious in early studies; however, in two randomized controlled trials in individuals with moderate to severe active GO, results have been conflicting, and therefore further studies are now needed. A longer-term goal is the development of an anti–TSH-R antibody or small molecule antagonist that could block binding of stimulatory TSH-R antibodies or inhibit TSH-R signaling, thus ameliorating the cause of hyperthyroidism.

AUTOIMMUNE HYPOTHYROIDISM

The most common cause of autoimmune hypothyroidism (AH) is chronic (or lymphocytic) autoimmune thyroiditis. There are two variants, atrophic and goitrous (Hashimoto thyroiditis).

Epidemiology

In populations living in iodine sufficient areas, AH is common, affecting between 1 and 10% of the population. The prevalence increases with age, with 3–20% of individuals over 75 years of age being hypothyroid. Like GD, AH is more common in women than in men. In a UK community survey, the incidence of hypothyroidism in women was 3.5/1000/year, which increased to 13.7/1000/year in women between 75 and 80 years of age. In men, the incidence was just 0.6/1000/year.[24]

Etiology

AH, like GD, is a complex genetic condition. Familial clustering provides evidence for a genetic etiology, which, in several studies, appears stronger than that for GD. The λs for AH is estimated to be between 10 and 45, suggesting that AH is more heritable compared with GD. In families with autoimmunity, frequently a mixture of individuals affected by AH and GD are seen, suggesting some shared genetic factors. The differing prevalence of AH in different ethnic groups, with AH being more common in Caucasian than in black populations, also supports a genetic background. Knowledge about the genetics of AH is limited.

The MHC class II *HLA-DR3, -DR4,* and *-DR5* alleles have been associated with AH in Caucasians only. Conflicting results have been reported for *HLA-DQ* alleles. One study reported that the *HLA-DQ* alleles *DQA1*0301* and *DQB1*0201* confer susceptibility to AH in Caucasians, with certain *HLA-DQ* alleles (*DQA1*0102* and *DQB1*0602*) reported to confer a protective effect. These data highlight the need to further investigate the MHC region to clarify its role in AH susceptibility.

In common with GD, the *CTLA-4* gene also appears to influence AH susceptibility. Three *CTLA-4* polymorphisms have been associated with AH in a number of populations. An A/G SNP located downstream of the 3’UTR (designated CT60), an A/G polymorphism at codon 17, and a 106-bp microsatellite repeat in the 3’UTR of exon 3. A locus on chromosome 8q24 containing the thyroglobulin gene was linked to AH, and a number of SNPs were subsequently studied in AH individuals with modest reported ORs for association of between 1.32 and 1.56. Other loci implicated in AH susceptibility include the tumor necrosis factor (*TNF-α*) gene, *PTPN22, CYP27B1,* T-cell receptor (TCR) genes, and several immunoglobulin genes and cytokine regulatory genes.

In contrast to GD, environmental factors in AH susceptibility have been challenging to identify. However, the role of iodine is widely accepted, since population studies have reported an increase in the prevalence of thyroid lymphocytic infiltration and autoantibodies following public health salt iodization programs. Infectious agents have also been implicated in susceptibility to AH. Several studies have identified an increased prevalence of IgG and/or IgA antibodies to virulence-associated outer membrane proteins of *Yersinia enterocolitica* in AH patients and in relatives of individuals with AH, suggesting that susceptibility genes for *Yersinia* infection may also confer risk for AH.

The effect of radiation, either “internal” (nuclear “fallout” or from RAI treatment) or “external” (radiotherapy or direct exposure during a nuclear accident), on AH susceptibility has been extensively studied. Following the nuclear reactor accident at Chernobyl, a rise in thyroid autoantibodies was noted 15 years following exposure; however, this was not accompanied by thyroid dysfunction.[25] Long-term follow-up studies of thyroid function in Japanese survivors of the atomic bombing of Nagasaki and Hiroshima have demonstrated a clear link between radiation exposure and thyroid cancer; however, the association with AH remains disputed. One study, at 40 years follow-up, demonstrated a significant relationship between dose of radiation exposure at Nagasaki and AH. However, a further study, at more than 50 years of follow-up, showed that radiation exposure did not correlate with either the occurrence of thyroid autoantibodies or AH.[26]

Immunopathogenesis

The mechanisms by which tolerance to thyroid antigens is lost in the first instance remain obscure. It appears that both a susceptible genetic background and a permissive environment are required before AH develops. Notably, AH is much more frequent in the autoimmune polyendocrinopathy type 1 (APECED) syndrome than GD, suggesting that central thymic T-cell selection, and therefore central tolerance, may be more important in AH than in GD. Histologically, lymphocytic infiltrates can be seen in the thyroid, consisting of both T and B cells (Fig. 70.7). These infiltrates can be diffuse or focal. Scarring and fibrosis may also be seen, with destruction of the normal thyroid architecture and an absence of colloid in thyroid follicles.

Both cell-mediated and humoral immune mechanisms are important in the continuing thyroid damage seen in AH. T cells are known to play a pivotal role in both the initiation and perpetuation of AH. Studies in which researchers induced hypothyroidism in Rag1-deficient transgenic mice that were unable to produce autoantibodies confirm this.[27] T cells respond to antigen-presenting

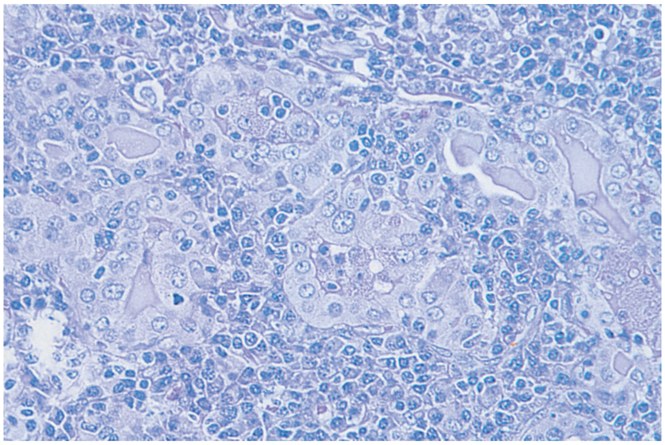


FIG 70.7 Lymphocytic infiltration of the thyroid in a patient with autoimmune hypothyroidism.

cells (APCs) and release cytotoxic and lytic factors, which result in thyrocyte death. Thyroid follicular cells have themselves been demonstrated to express HLA class II molecules, suggesting that they may also have a direct role in antigen presentation.

The humoral immune response is also important in AH. More than 90% of individuals with AH have detectable TPO antibodies. Autoantibodies directed against Tg and, to a lesser degree, the TSH-R are also commonly detected. *In vitro,* TPO antibodies can fix complement and directly induce cell damage.[17] Their presence within thyroid follicles in AH patients suggests they may have the same effect *in vivo,* although the thyrocyte destruction found in the *Rag1*-deficient mouse suggests TPO antibodies are not necessary for AH. Interestingly, the epitopes toward which TPO antibodies are directed in both GD and AH overlap, and there is no disease specificity for the targeted TPO domain.[17] TSH-R autoantibodies found in rare patients with AH are likely to exert a blocking or antagonist effect, thus inducing hypothyroidism. As thyroid hormone secretion falls, increasing thyrocyte stimulation by elevation of serum TSH may induce or augment thyroid autoantigen expression (*e.g.*, TPO, Tg), thereby perpetuating the autoimmune response.

Clinical Presentation

Hypothyroidism can result in changes in almost every organ system in the body (Table 70.3). Initially, the signs and symptoms may be subtle and nonspecific, including tiredness, cold sensitivity, and constipation. Hypothyroidism is frequently diagnosed incidentally following blood tests for another problem. The typical goiter palpable in Hashimoto thyroiditis is moderate in size and firm with a finely granular surface. Individuals often report a gradual increase in size over a number of years. However, rapid growth is unusual. In atrophic AH, the size of the thyroid gland is reduced.

Investigation and Diagnosis

Hypothyroidism is detected biochemically by a raised serum TSH with reduced fT_4. AH is differentiated from other forms of hypothyroidism by the presence of circulating autoantibodies, including TPO and Tg. On ultrasound scanning, the thyroid gland appears finely heterogeneous and hypoechoic, a change that predates serum autoantibody positivity in some. Fine-needle aspiration (FNA) demonstrates lymphocytes and Hürthle cells, but this investigation is generally only indicated if a discrete nodule is present.

TABLE 70.3 Common and Rare Clinical Manifestations of Hypothyroidism

Common	Rare
Neuropsychiatric	
Lethargy	Cerebellar ataxia
Impaired cognitive function	Deafness
Slow speech	Psychosis
Depression	
Gastrointestinal	
Anorexia	Ascites
Weight gain	
Constipation and bloating	
Abnormal liver function tests	
Cardiorespiratory	
Shortness of breath on exertion	Pericardial effusion
Reduced exercise tolerance	
Bradycardia	
Diastolic hypertension	
Cardiomegaly	
Low-voltage electrocardiogram	
Peripheral edema (nonpitting)	
Genitourinary	
Oligomenorrhea, amenorrhea, menorrhagia	
Reduced libido	
Early fetal loss	
Impotence	
Cutaneous	
Cold intolerance	
Skin dryness and thickening	
Malar flush	
Edema of the face, hands, and feet	
Change in face shape	
Pallor	
Nail abnormalities	
Alopecia	
Musculoskeletal	
Bradykinesia	
Joint and muscular pains	
Delayed relaxation of tendon reflexes	
Miscellaneous	
Goiter (in Hashimoto thyroiditis)	
Reduced basal metabolic rate	
Increased sensitivity to exogenous insulin	
Abnormal lipid metabolism	

Reflecting the effects of hypothyroidism on multiple organ systems, many biochemical and hematological abnormalities, such as mild anemia, hyponatremia, or elevated serum creatine kinase, transaminases, lactate dehydrogenase, and low-density lipoprotein (LDL) cholesterol, are also commonly detected in patients with AH.

Management

AH requires lifelong treatment with thyroid hormone replacement therapy. The most commonly prescribed is synthetic thyroxine, levothyroxine (L-T_4), which is widely available and inexpensive.

Except for individuals with known heart disease or the very old, a full, weight-related replacement dose (≈1.6 μg/kg/day) should be started. Once the patient is on a stable dose, thyroid function should be assessed annually to ensure that the patient continues to receive the appropriate dose. Some commonly prescribed medications, such as calcium and iron supplements, and proton pump inhibitors interfere with the absorption of L-T_4, and patients should be advised to take these at least 4 hours before or after their L-T_4 to ensure maximum absorption.

Subclinical Hypothyroidism

Although the need to treat individuals with overt AH is universally accepted, it is unclear whether thyroid hormone replacement is beneficial in individuals with persistent subclinical hypothyroidism (serum TSH raised, but fT_4 and fT_3 within the normal reference range on at least two separate occasions). Progression to overt hypothyroidism from this state occurs in about 2% of individuals per year who are TPO antibody negative, rising to 5% per year if antibodies are present. Persistent subclinical hypothyroidism has been associated with a number of markers of cardiac and vascular dysfunction in observational studies, including left ventricular diastolic dysfunction, increased vascular resistance, and atherosclerosis. A randomized controlled study is now underway to determine whether thyroid hormone replacement can be used to improve cardiovascular outcomes in patients with persistent subclinical hypothyroidism.[28]

CLINICAL PEARLS

Management of Persistent Subclinical Hypothyroidism

- If thyroid-stimulating hormone (TSH) is elevated, immediate treatment should be offered:
  - In pregnancy
  - Preconception, if planning a pregnancy
- If TSH is elevated but <10 mIU/L, a trial of treatment can be offered:
  - To individuals with convincing symptoms of hypothyroidism
- If TSH is >10 mIU/L, treatment should be offered:
  - To individuals under the age of 70 years
  - To individuals over the age of 70 years if there is a clear history of hypothyroid symptoms or there are significant risk factors for cardiovascular disease.

FUTURE DEVELOPMENTS

The genetics of AH remain understudied considering its frequency as the commonest autoimmune disease in humans. Considerable work remains to be done on whether treatment of subclinical hypothyroidism is beneficial. Given the insidious nature of the development of AH, it remains a good target for a preventative immunotherapeutic intervention, if a safe and economic treatment can be found.

OTHER FORMS OF THYROIDITIS

The term *thyroiditis* relates to conditions resulting in inflammation of the thyroid gland. A number of etiologies have been described, including infection, radiation exposure, drugs, and autoimmune factors. A common pattern to the natural history of several thyroiditides is frequently seen, involving an initial thyrotoxic phase of 1–3 months, followed by a rapid drop in serum thyroid hormones and a transient hypothyroid phase, often lasting another 1–4 months. During the thyrotoxic phase, preformed thyroid hormone stores are released from the thyroid follicles, leading to thyrotoxicosis, which may be severe. The hypothyroid phase follows when these preformed stores are exhausted and the thyroid has become depleted of hormones. In about 90% of cases this hypothyroidism is transient, but in some cases, it never resolves.

POSTPARTUM THYROIDITIS

Postpartum thyroiditis (PPT) is a common endocrine condition that manifests within 1 year following pregnancy.[29] It affects between 5% and 10% of women in the general population. PPT is classically a biphasic disorder, consisting of a period of transient thyrotoxicosis (median onset 12–14 weeks post partum) followed by a period of transient hypothyroidism (median onset 18–20 weeks post partum); however, a monophasic (thyrotoxicosis or hypothyroidism alone) or reversed biphasic (hypothyroidism followed by thyrotoxicosis) pattern can also occur. During pregnancy, there is a state of relative immune tolerance, followed by a rebound in immune function following delivery, coinciding with the occurrence of PPT. The presence of thyroid autoantibodies and a lymphocytic infiltrate on thyroid biopsy supports an autoimmune basis for this condition.[29]

Clinically, the thyrotoxic phase of PPT is often mild, resulting in symptoms of fatigue and irritability, which can be misdiagnosed as postnatal depression. If the thyrotoxic episode is short, it may even go unnoticed. Neck pain is not a feature. Women with PPT who are thyrotoxic may benefit from a beta-blocker, such as propranolol, for symptom relief. Antithyroid drugs are not effective, as the thyrotoxicosis is caused by release of preformed thyroid hormones. Following the episode of thyroid dysfunction, 10–20% of women remain permanently hypothyroid.[29] In women who have had PPT and then recovered, an annual assessment of thyroid function is recommended, as their risk of long-term hypothyroidism is considerable.[29] In those women who return to being euthyroid, there is a 75% risk of PPT in subsequent pregnancies and a 50% risk of permanent hypothyroidism at 7 years.[30]

TRANSLATIONAL RESEARCH

ON THE HORIZON

New Approaches to Therapy of Graves Disease

- Novel immunotherapeutic agents for Graves orbitopathy
- Anti–thyroid-stimulating hormone receptor (TSH-R) blocking antibodies for control of hyperthyroid Graves disease
- Small molecule TSH receptor antagonists for management of hyperthyroidism in Graves disease

The major challenge of the next 5–10 years is to take novel immunotherapeutic agents, including "biologicals" that have been developed for rheumatic disorders into the clinical arena for autoimmune thyroid diseases. The primary target of these efforts should be GO, which remains a disfiguring condition, often with substantial functional impairment of vision and associated low quality of life. Two randomized controlled trials of B-lymphocyte–depleting agents have yielded conflicting results, and therefore other powerful agents need to be investigated as this remains a disorder with no wholly satisfactory treatment. Early diagnosis of GO and development of markers that predict progressive or severe disease will also be helpful in identifying patients for early intervention. The development of anti–TSH-R–blocking antibodies as a therapeutic agent for hyperthyroid GD seems likely. These agents might have a role in patients unlikely to gain medical remission from thionamide antithyroid drugs or in those for whom rapid control of hyperthyroidism is desirable. Small-molecule TSH-R antagonists may also come to the clinic, with a similar role.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Tunbridge WM, Evered DC, Hall R, et al. The spectrum of thyroid disease in a community: the Whickham survey. Clin Endocrinol (Oxf) 1977;7:481–93.
2. Laurberg P, Pedersen KM, Vestergaard H, et al. High incidence of multinodular toxic goitre in the elderly population in a low iodine intake area vs. high incidence of Graves' disease in the young in a high iodine intake area: comparative surveys of thyrotoxicosis epidemiology in East-Jutland Denmark and Iceland. J Intern Med 1991;229:415–20.
3. Brix TH, Kyvik KO, Christensen K, et al. Evidence for a major role of heredity in Graves' disease: a population-based study of two Danish twin cohorts. J Clin Endocrinol Metab 2001;86:930–4.
4. Vaidya B, Kendall-Taylor P, Pearce SH. The genetics of autoimmune thyroid disease. J Clin Endocrinol Metab 2002;87:5385–97.
5. Eschler DC, Hasham A, Tomer Y. Cutting edge: the etiology of autoimmune thyroid diseases. Clin Rev Allergy Immunol 2011;41:190–7.
6. Ueda H, Howson JM, Esposito L, et al. Association of the T-cell regulatory gene CTLA4 with susceptibility to autoimmune disease. Nature 2003;423(6939):506–11.
7. Velaga MR, Wilson V, Jennings CE, et al. The codon 620 tryptophan allele of the lymphoid tyrosine phosphatase (LYP) gene is a major determinant of Graves' disease. J Clin Endocrinol Metab 2004;89:5862–5.
8. Brent GA. Environmental exposures and autoimmune thyroid disease. Thyroid 2010;20:755–61.
9. Vestergaard P. Smoking and thyroid disorders—a meta-analysis. Eur J Endocrinol 2002;146:153–61.
10. Cawood TJ, Moriarty P, O'Farrelly C, et al. Smoking and thyroid-associated ophthalmopathy: a novel explanation of the biological link. J Clin Endocrinol Metab 2007;92:59–64.
11. Winsa B, Adami HO, Bergström R, et al. Stressful life events and Graves' disease. Lancet 1991;338:1475–9.
12. Matos-Santos A, Nobre EL, Costa JG, et al. Relationship between the number and impact of stressful life events and the onset of Graves' disease and toxic nodular goitre. Clin Endocrinol (Oxf) 2001;55:15–19.
13. Chen F, Day SL, Metcalfe RA, et al. Characteristics of autoimmune thyroid disease occurring as a late complication of immune reconstitution in patients with advanced human immunodeficiency virus (HIV) disease. Medicine (Baltimore) 2005;84:98–106.
14. Smith BR, Bolton J, Young S, et al. A new assay for thyrotropin receptor autoantibodies. Thyroid 2004;14:830–5.
15. Czarnocka B. Thyroperoxidase, thyroglobulin, Na(+)/I(−) symporter, pendrin in thyroid autoimmunity. Front Biosci 2011;16:783–802.
16. Costagliola S, Bonomi M, Morgenthaler NG, et al. Delineation of the discontinuous-conformational epitope of a monoclonal antibody displaying full in vitro and in vivo thyrotropin activity. Mol Endocrinol 2004;18:3020–34.
17. McLachlan SM, Rapoport B. Thyroid peroxidase as an autoantigen. Thyroid 2007;17:939–48.
18. Hegedus L. Treatment of Graves' hyperthyroidism: evidence-based and emerging modalities. Endocrinol Metab Clin North Am 2009;38:355–71, ix.
19. Perros P, Dayan CM, Dickinson AJ, et al. Management of patients with Graves' orbitopathy: initial assessment, management outside specialised centres and referral pathways. Clin Med (Lond) 2015;15:173–8.
20. Bahn RS. Graves' ophthalmopathy. N Engl J Med 2010;362:726–38.
21. Marcocci C, Kahaly GJ, Krassas GE, et al. Selenium and the course of mild Graves' orbitopathy. NEJM 2011;364:1920–31.
22. Bartalena L, Baldeschi L, Dickinson A, et al. Consensus statement of the European Group on Graves' orbitopathy (EUGOGO) on management of GO. Eur J Endocrinol 2008;158:273–85.
23. Salvi M, Campi I. Medical treatment of Graves' orbitopathy. Horm Metab Res 2015;47:779–88.
24. Vanderpump MP, Tunbridge WM, French JM, et al. The incidence of thyroid disorders in the community: a twenty-year follow-up of the Whickham Survey. Clin Endocrinol (Oxf) 1995;43:55–68.
25. Agate L, Mariotti S, Elisei R, et al. Thyroid autoantibodies and thyroid function in subjects exposed to Chernobyl fallout during childhood: evidence for a transient radiation-induced elevation of serum thyroid antibodies without an increase in thyroid autoimmune disease. J Clin Endocrinol Metab 2008;93:2729–36.
26. Imaizumi M, Usa T, Tominaga T, et al. Radiation dose-response relationships for thyroid nodules and autoimmune thyroid diseases in Hiroshima and Nagasaki atomic bomb survivors 55-58 years after radiation exposure. JAMA 2006;295:1011–22.
27. Quaratino S, Badami E, Pang YY, et al. Degenerate self-reactive human T-cell receptor causes spontaneous autoimmune disease in mice. Nat Med 2004;10:920–6.
28. Pearce SHS, Brabant G, Duntas LH, et al. 2013 ETA guideline: management of subclinical hypothyroidism. Eur Thyroid J 2013;2:215–28.
29. Landek-Salgado MA, Gutenberg A, Lupi I, et al. Pregnancy, postpartum autoimmune thyroiditis, and autoimmune hypophysitis: intimate relationships. Autoimmun Rev 2010;9:153–7.
30. De Groot L, Abalovich M, Alexander EK, et al. Management of thyroid dysfunction during pregnancy and postpartum: an Endocrine Society clinical practice guideline. J Clin Endocrinol Metab 2012;97:2543–65.
31. Davies TF, Ando T, Lin RY, et al. Thyrotropin receptor-associated diseases: from adenomata to Graves' disease. J Clin Invest 2005;115:1972–83.

MULTIPLE-CHOICE QUESTIONS

1. When assessing an individual with thyrotoxicosis, which feature is NOT consistent with a diagnosis of Graves disease?
   - **A.** Increasing severity of symptoms over weeks to months
   - **B.** Grittiness of the eyes and blurred vision on waking in the morning
   - **C.** Neck pain
   - **D.** A family history of autoimmune hypothyroidism and pernicious anemia
   - **E.** Relief of symptoms with beta-blocker medication
2. Which of the listed clinical features is in keeping with a diagnosis of postpartum thyroiditis (PPT)?
   - **A.** Proptosis of the eyes
   - **B.** Pretibial myxedema
   - **C.** Previous history of Graves disease
   - **D.** Goiter with a bruit
   - **E.** Onset of symptoms 4 months after delivery
3. In moderate to severe Graves ophthalmopathy (GO), the following treatments are evidence based and effective except:
   - **A.** Oral selenium supplements
   - **B.** Topical lubricants
   - **C.** Systemic glucocorticoids
   - **D.** Orbital radiotherapy
   - **E.** Orbital decompression surgery

71

Type 1 Diabetes

Leonard C. Harrison

Diabetes is not a single disease but a metabolic syndrome in which different mechanisms lead to deficiency of insulin and/or impaired insulin action and persistent hyperglycemia. The American Diabetes Association classified diabetes into four categories based on etiology rather than age of onset (juvenile-onset versus adult-onset) or requirement for insulin therapy (insulin-dependent versus noninsulin-dependent).[1] The vast majority of cases, approximately 10% and 90%, respectively, are attributed to type 1 and type 2 diabetes. This chapter focuses mainly on type 1 diabetes, which results from an absolute deficiency of insulin secondary to the loss of pancreatic β cells. Type 1 diabetes is classified as 1A (immune-mediated) or 1B (idiopathic), primarily depending on the presence or absence, respectively, of pancreatic islet autoantibodies. However, as discussed below, type 1A diabetes (T1DA) and type 1B diabetes (T1DB) also differ in their natural history and clinical features.

Diabetes is diagnosed on the basis of the following criteria[1]: symptoms in association with a casual plasma glucose ≥200 mg/dL (11.1 mmol/L) OR fasting plasma glucose ≥126 mg/dL (7.0 mmol/L) OR 2-hour plasma glucose ≥200 mg/dL (11.1 mmol/L) in an oral glucose tolerance test (OGTT; 75 g glucose in water). Without symptoms, the diagnosis of diabetes must rest on confirmation of a raised plasma glucose concentration. A further diagnostic criterion introduced by the World Health Organization (WHO) in 2011, in particular for the diagnosis of type 2 diabetes, is a confirmed blood glycated hemoglobin (HbA$_{1c}$) ≥48 mmol/mol (6.5%). Because HbA$_{1c}$ is an integrated measure of glycemia over many weeks, it is not suitable for diagnosing children or in the following circumstances: suspected type 1 diabetes; symptoms of diabetes for <2 months; acute illness; medication that may increase blood glucose (*e.g.,* steroids, antipsychotics); and pregnancy. The criteria for diagnosing gestational diabetes are stricter: fasting plasma glucose ≥101 mg/dL (5.6 mmol/L) or 2-hour plasma glucose in an OGTT ≥140 mg/dL (7.8 mmol/L).

The classic symptoms and signs of T1DA, because of high concentrations of blood glucose, are polyuria, polydipsia, and unexplained weight loss; others include fatigue, increased hunger, impaired visual acuity as a result of changes in the refractive index of vitreous humor, tingling or numbness in the hands or feet resulting from sensory nerve changes, and vaginal irritation caused by *Candida* infection. If diabetes is undiagnosed or untreated, failure to metabolize glucose will result in the breakdown of fat, leading to ketonemia and ketoacidosis, which may be accompanied by nausea and hyperventilation before life-threatening ketoacidotic coma. In children presenting with the classic symptoms, T1D can be diagnosed clinically. However, in Caucasians, the diagnosis can be confirmed and in older individuals and less clear-cut cases may be clearly established by detecting circulating autoantibodies to islet antigens. These autoantibodies to insulin (IAAs), glutamic acid decarboxylase 65 000 mol. wt. isoform (GADA), insulinoma-like antigen-2 (IA-2A), and zinc transporter-8 (ZTA) are markers of β-cell autoimmunity that appear many months to years before symptoms and therefore denote high risk for clinical disease.[2-5] They are present in at least 90% of Caucasian children with T1DA (compared with ≈1% of the background population) but in only ≈50% of Hispanic American and African American children, an increasing number of whom have T2D and, in some cases, T1DB. Negative results for islet autoantibodies in children with diabetes should also alert to the possibility of monogenic maturity-onset diabetes of the young (MODY) and sulfonylurea receptor syndromes.

The hallmark of T1DA is progression to absolute insulin deficiency within several years after diagnosis. The connecting peptide in proinsulin (C-peptide) is secreted in an equimolar ratio to insulin and is used as a surrogate for insulin to evaluate residual β-cell function in the face of exogenous insulin or IAAs, or insulin antibodies induced by exogenous insulin. The plasma C-peptide in response to a mixed meal tolerance test is the gold standard. Although children with T1DA have lost most of their β-cell function at diagnosis, the measurement of plasma C-peptide is not a reliable way of distinguishing T1D, especially at diagnosis. Hyperglycemia impairs β-cell function, and when corrected by rehydration and insulin replacement, it may be followed by a "honeymoon phase" of partial recovery of β-cell function and a decreased requirement for exogenous insulin that may last many months. By several years after diagnosis, most young children display little residual β-cell function; however, in older children and adults, residual C-peptide secretion has been observed for many years and is believed to be associated with better glycemic control.

Classically, T1DA was considered a disease of "juvenile onset" in normal-weight individuals in contrast to type 2 diabetes with onset in middle-aged, overweight individuals. However, this view requires reappraisal.[6] First, 5–10% of adults who present with diabetes diagnosed initially as T2D and are typically overweight, have evidence of low-grade islet autoimmunity manifest by the presence of GADAb, and occasionally IA-2Ab or ZTAb.[7] If we accept that they have a slowly progressive form of autoimmune β-cell destruction, that was termed *latent autoimmune diabetes in adults* (LADA), their number doubles the prevalence of T1DA.

Second, the “obesity epidemic” has impacted on the T1D and T2D stereotypes. “Hybrid” types with β-cell autoimmunity and insulin resistance (“double diabetes”) are becoming increasingly common, characterized by weaker immune and genetic markers (lower avidity autoantibodies predominantly to GAD and lower risk human leukocyte antigen [HLA] alleles). Low-grade innate immune inflammation, associated with insulin resistance in obesity, could promote and uncover latent β-cell autoimmunity in individuals with lower-risk genes for T1DA.

T1DB excludes known specific causes of β-cell dysfunction, such as monogenic diabetes. It encompasses forms of ketosis-prone diabetes initially described in West Africans and African Americans[8] and subsequently in other ethnic groups, but rarely seen in Caucasians. Obesity and relatively well-preserved residual β-cell function were features of the African cases. T1DB also includes fulminant diabetes initially described in Japan, where it may account for 15–20% of new-onset T1D, and subsequently elsewhere in Asia.[9,10] Fulminant diabetes is associated with widespread mononuclear cell infiltration of both the exocrine and endocrine pancreas, and elevated concentrations of serum pancreatic amylase, elastase, and lipase, in individuals with HLA susceptibility genes for T1D.[10] Case reports have associated it with acute viral infection and drug hypersensitivity reactions, but its etiology remains unclear. In Caucasians, T1DB includes T1D, in which islet autoantibodies are undetectable, but the latter may reflect assay insensitivity on a particular occasion. A classification of diabetes that hinges on the presence or absence of islet autoantibodies is clearly not ideal. β cell–specific T cells, not islet autoantibodies, are the effectors of β-cell damage, and reciprocal relationships between autoantibodies and autoreactive T cells to islet antigens have been reported.[11] In addition, the more recent discovery of autoantibodies to ZT8 suggests the possibility that autoantibodies to islet antigens remain to be discovered and might be markers of particular subtypes of T1D. For example, in fulminant diabetes, the clinical picture suggests that if autoimmunity is involved, it targets the exocrine pancreas. Furthermore, in contrast to T1DA, in which there is an extended preclinical history of islet autoimmunity, autoantibodies may not be present at the time when fulminant diabetes presents acutely.

CLINICAL PEARLS

- The type 1 diabetes (T1D) stereotype of the thin juvenile now overlaps with the type 2 diabetes (T2D) stereotype of the obese, insulin-resistant adult.
- In Caucasian children, the diagnosis of T1D can be made on the basis of classic clinical features.
- The presence of islet autoantibodies confirms the diagnosis of type 1A diabetes (T1DA).
- Islet autoantibodies are present in more than 90% of Caucasian children but in only about 50% of Hispanic or African American children, some of whom have T2D.
- Islet autoantibodies can be detected in the first 3 years of life in the majority of Caucasian children who go on to develop T1DA.
- The risk for T1DA increases with the number and titer of islet autoantibodies.

EPIDEMIOLOGY AND NATURAL HISTORY OF T1DA

The incidence of T1DA varies widely across the globe, being highest among Caucasians of northwestern Europe and countries to which they have emigrated (Fig. 71.1). This, in part, reflects the population distribution of specific-risk HLA genes, which account for up to 50% of the lifetime risk for T1DA. However, the incidence of T1DA is rising in many countries on a background of lower-risk HLA alleles. In Western societies, the incidence in childhood has more than doubled since the 1980s and has been rising at ≈3% annually, particularly among younger children.[12] The same trend is also occurring in other countries, such as Kuwait and Saudi Arabia, and some parts of India and China, where Western lifestyles have been adopted but where the prevalence of high-risk HLA genotypes for T1DA is much lower. Just as in T2D, environmental factors may increase the penetrance of risk genes for T1DA. In the case of HLA genes in Caucasians, the increasing incidence of T1DA is accounted for by children with intermediate-risk (DR 4,4 or DR 3,3) or low-risk (DR 4,X or DR 3,X) phenotypes, not the highest-risk HLA phenotypes (DR 3,4; DQ 2,8).[13] These lower-risk phenotypes are also the ones seen in non-Caucasians and adults with T1DA. The greatest number of children with diabetes will soon be found in the most populous regions of the world, that is, India and China. The impact of the modern Western environment on T1DA is discussed further below.

T1DA affects both sexes equally in childhood, with a slight excess of males in early adult life. Of newly diagnosed cases, no more than 10–15% have a family history of T1DA. However, studies of affected families have provided major insights into the genetics and natural history of T1DA. In relatives with T1DA, the rate of progression to clinical diabetes is positively associated with the number and titer of islet autoantibodies,[2,3,5] the number of HLA class 2 risk alleles (DR3, DR4) and specific HLA class I alleles (A24),[14] and the degree of insulin resistance[15] and is negatively correlated with age. The specificity of autoantibodies is also important. IAAs are more often the first sign of islet autoimmunity in children who are followed up from birth and, of all the autoantibodies, has the highest predictive value.[2-5] In Europe, North America, and Australia, birth cohort studies of children who have a relative with T1DA have shown that the development of diabetes by the age of 18 years is almost always associated with the appearance of islet autoantibodies in the first few years of life.[5] Of children with ≥2 islet autoantibodies before the age of 3 years, 57% (95% confidence interval [CI] 51.7–62.3%) and 74.8% (95% CI 69.7–79.9%) progressed to diabetes by 6 and 10 years of age, respectively. With a single islet autoantibody, 14.5% progressed to diabetes by 10 years of age.

Men with early-onset T1DA are more likely to transmit diabetes to their offspring compared with women (Fig. 71.2), but it is not clear whether this results from genomic imprinting of a gene expressed only from a paternally inherited allele or other possible reasons. The risk is also greater in later-born offspring. The peak age of incidence in children is at the cusp of puberty, a time when the body’s requirement for insulin increases along with an increase in insulin resistance. Like the seasonal autumn–winter peak in diagnosis that is attributed to viral infections, this is likely to be “the straw that breaks the camel’s back” on a background of longer-standing autoimmune β-cell dysfunction.

PATHOGENESIS: NATURE AND NURTURE

The model for the staged natural history of T1DA, arising from the seminal studies of Eisenbarth,[16] is illustrated in Fig. 71.3. Although the pattern and rate of β-cell loss is depicted as linear, it is more likely to be episodic, reflecting direct (*e.g.*, virus, chemical toxin) or indirect (*e.g.*, diet, pollutants, microbiome)

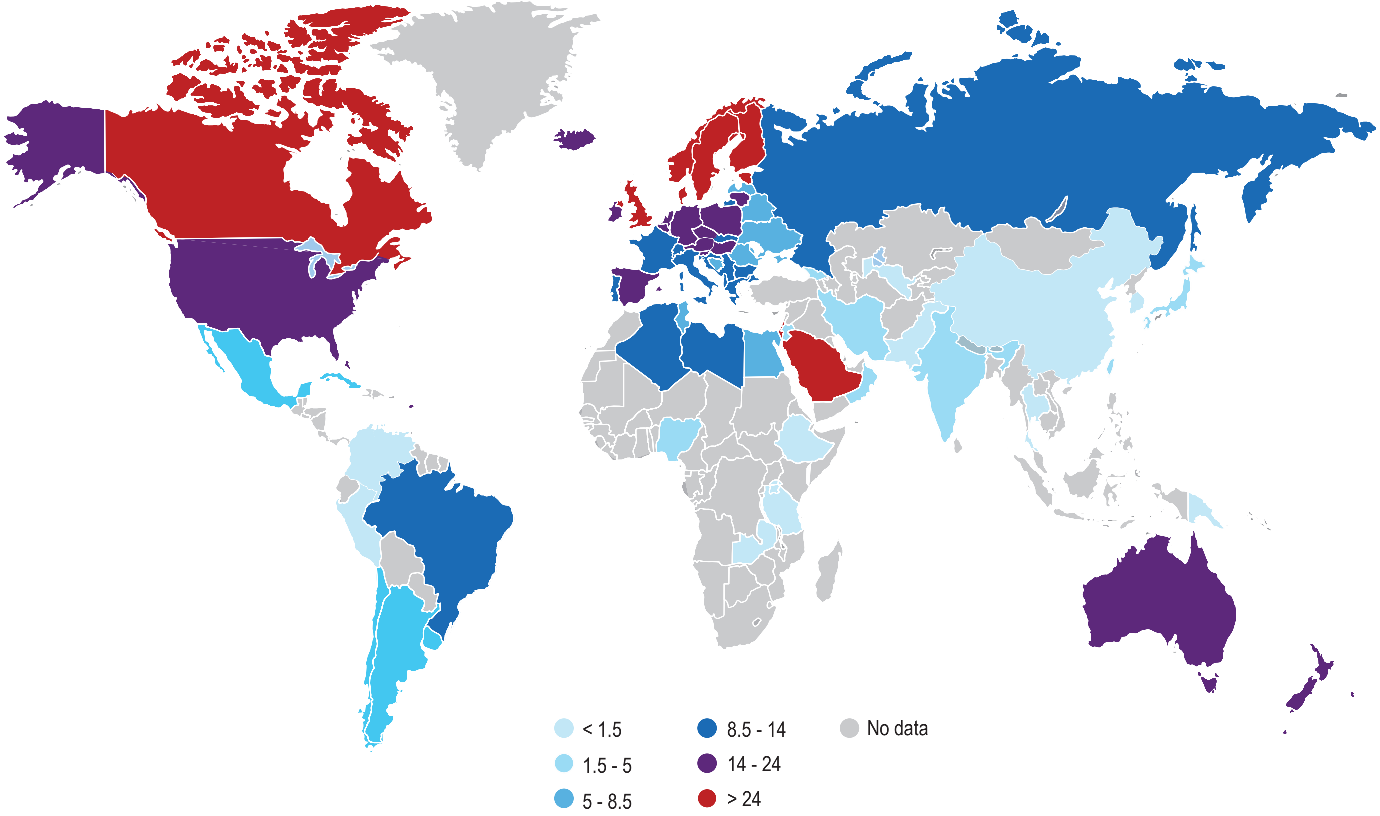

● < 1.5
● 1.5 - 5
● 5 - 8.5
● 8.5 - 14
● 14 - 24
● > 24
● No data

FIG 71.1 Global Incidence of type 1A diabetes (T1DA) in 100 000 children ages 0–14 years per year. (From IDF diabetes atlas, 7th ed. Brussels, Belgium: International Diabetes Federation; 2015.)

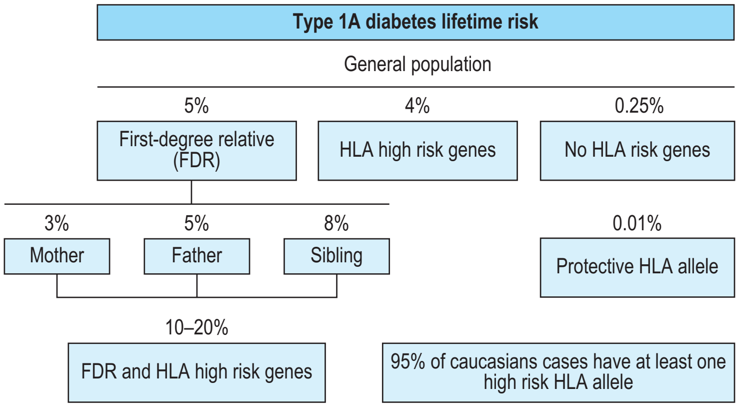

Type 1A diabetes lifetime risk
General population
5%
First-degree relative (FDR)
4%
HLA high risk genes
0.25%
No HLA risk genes
3%
Mother
5%
Father
8%
Sibling
0.01%
Protective HLA allele
10–20%
FDR and HLA high risk genes
95% of caucasians cases have at least one high risk HLA allele

FIG 71.2 Lifetime risk of type 1A diabetes.

interactions between the environment and β cells. In addition, evidence suggests that immune activation and β-cell destruction may be accelerated in the late preclinical stage. The appearance of predictive islet autoantibodies in the first years of life[5] means that the stage for developing T1DA is set very early, even before birth, on a background of genetic susceptibility. These early years must provide clues to environment–gene interactions that lead to immune dysfunction and disease.

It is generally agreed that β-cell destruction in T1DA is mediated by autoreactive T cells, the ultimate effectors being CD8 cytotoxic T cells (Fig. 71.4). The evidence for this is unequivocal in the inbred nonobese diabetic (NOD) mouse model of T1DA, which shares several key features with T1DA in outbred humans (Table 71.1) despite 65 million years of evolutionary distance. The molecular mechanisms of β-cell death, gleaned mostly from the NOD mouse, encompass a combination of both apoptosis induced by activation of extrinsic (*e.g.*, tumor necrosis factor [TNF] receptor or Fas ligation) or intrinsic (*e.g.*, endoplasmic reticulum [ER] stress) pathways and necroptosis induced by cytotoxic CD8 T-cell granule components (granzymes and perforin), reactive oxygen species (ROS), or ischemia.

The evidence from human studies is obviously less compelling, as access to human pathology is limited. Studies of pancreas biopsies and organ-donor pancreas, more recently from the

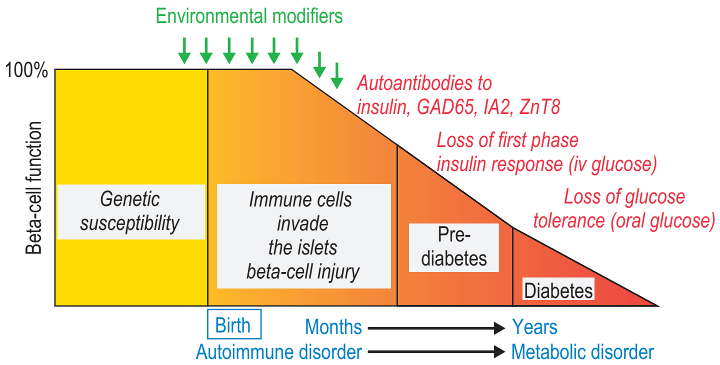


FIG 71.3 Natural history of type 1A diabetes.

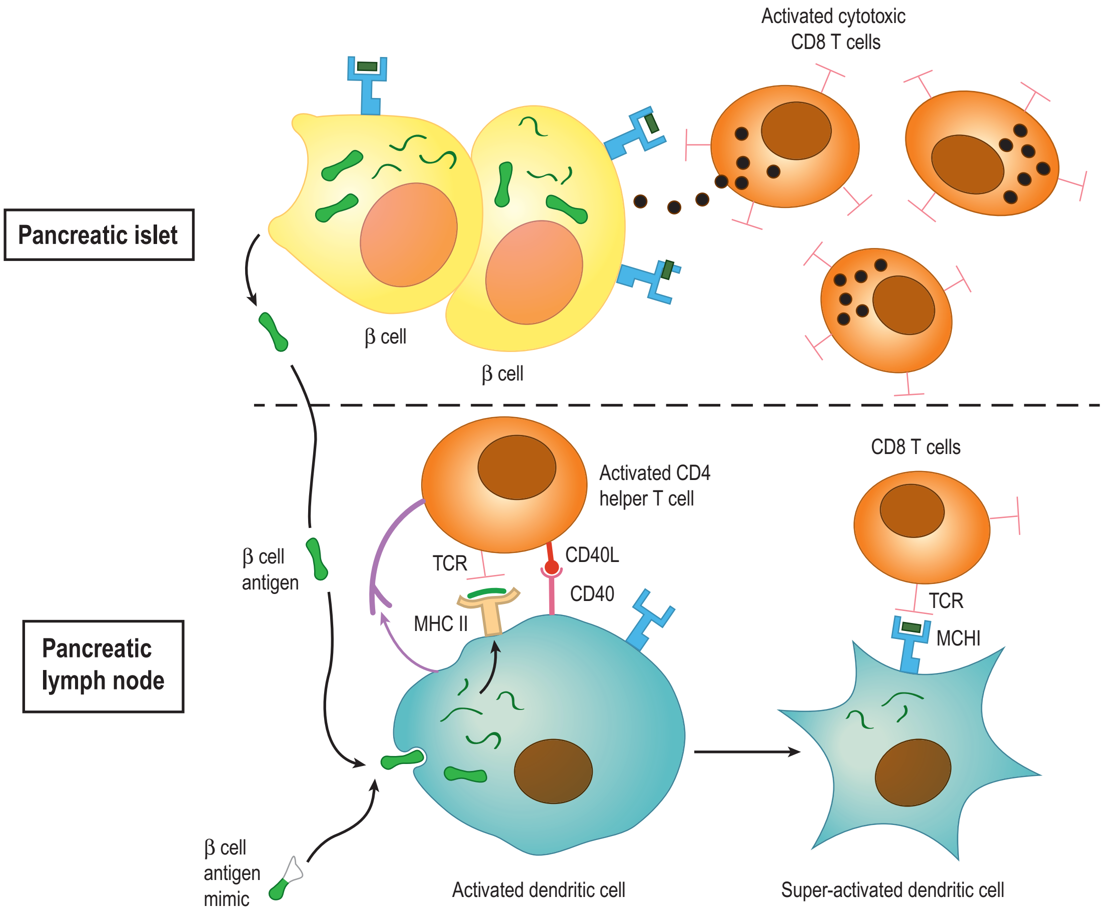

Activated cytotoxic CD8 T cells
Pancreatic islet
β cell
β cell
Pancreatic lymph node
β cell antigen
Activated CD4 helper T cell
TCR
CD40L
CD40
MHC II
CD8 T cells
TCR
MCHI
β cell antigen mimic
Activated dendritic cell
Super-activated dendritic cell

FIG 71.4 Immune mechanisms of β-cell destruction in type 1A diabetes.

Network for Pancreatic Organ Donors with Diabetes (nPOD; www.jdrfnpod.org), have not revealed the florid immune cell infiltration of islets (insulitis) seen in NOD mice but rather patchy insulitis predominantly caused by CD8 T cells and macrophages. A further observation, confirming earlier observations,[17] is hyperexpression of HLA class I by islets, in association with immunoreactive interferon-α (IFN-α), even in atrophic islets lacking insulin, which can be taken as presumptive evidence for persisting virus. A key question is: why are β cells specifically targeted? The answer may lie in insulin itself. Central to the concept of autoimmune disease is the notion that the pathology is driven by loss of immune tolerance to self-antigens. In the case of T1DA, considerable evidence, direct in the NOD mouse, identifies (pro)insulin as a key disease-initiating self-antigen (Table 71.2).

GENES

The concordance for T1D in monozygotic twins, who are almost totally genetically identical, approaches 50%, indicating that both genetic and environmental–epigenetic mechanisms contribute to disease. Large case/control studies and more recently genome-wide

TABLE 71.1 The Nonobese Diabetic (NOD) Mouse as a Model of Human Type 1A Diabetes

Feature	NOD Mouse	Human
Preclinical stage	Yes	Yes
Gender	F > M	M > F after puberty
Genetic susceptibility		
MHC class II fi57 non-Asp	Yes (I-Ag^7)	Yes (HLA-DQ8)
Polygenic non-MHC	Yes	Yes
Environmental influence on gene penetrance	Yes	Yes
Disease transmission via bone marrow	Yes	Yes
Mononuclear cell infiltration of islets (insulitis)	Marked	Moderate
Other organs	Yes	Sometimes
Impaired immune regulation	Yes	Yes
Autoantigens:		
(Pro)insulin	Yes	Yes
Glutamic acid decarboxylase (GAD)	Yes	Yes
Clinical response to autoantigen-specific therapy	Yes	Not yet shown

TABLE 71.2 Evidence for the Key Role of Proinsulin as Primary Autoantigen in Type 1A Diabetes (T1DA)

- β cell–specific (except for thymus)
- Second strongest genetic locus in humans (*IDDM2*) = VNTR 5′ of INS allelism correlates with proinsulin transcription in the thymus
- Early target of autoimmunity in humans and nonobese diabetic (NOD) mice
- Main target of T cells isolated from islets of NOD mice and humans with T1DA

Genetic manipulation in the NOD mouse:

- Transgenic expression of proinsulin (but not glutamic acid decarboxylase [GAD]) in antigen-presenting cells (APCs) prevents insulitis/diabetes
- Transfer of hematopoietic stem cells (HSCs) or myeloid progenitors encoding proinsulin in APC progeny prevents diabetes
- Knockout of proinsulin II (expressed in thymus) accelerates diabetes
- Induction of mucosal tolerance to (pro)insulin prevents diabetes

association (GWA) studies have identified over 40 chromosomal loci associated with risk for T1DA[18,19] (Fig. 71.5). Many are weak associations defined only by single nucleotide polymorphisms (SNPs), and the genes with which the SNPs are in linkage and/or their functional contribution to pathogenesis are not known. It is likely that different combinations of genes/susceptibility loci in different environmental contexts lead by different routes to the final outcome of β-cell destruction and account for phenotypic–clinical heterogeneity in T1D. Nevertheless, what is clear across populations globally is the dominance of the HLA locus (IDDM1) as the single most important genetic contributor to T1D risk (see Table 71.1), accounting for about half the lifetime risk. The highest risk HLA haplotypes in Europeans are DR3 (*DRB1*03:01-DQA1*05:01-DQB1*02:01*) and DR4 (usually *DRB1*04:01* or **04:04* with *DQA1*03:01-DQB1*03:02*). It should also be noted that the HLA DQ6 haplotype *DQA1*01:02-DQB1*06:02,* which is in linkage with DR15 (*DRB1*15:01*), is dominantly protective for T1DA but denotes susceptibility for a subtype of multiple sclerosis and for narcolepsy. The second most important contribution is from the insulin gene (*INS*) locus (IDDM2), which maps to a variable number of tandem repeats (VNTR) 5’ of the coding sequence. Apart from the HLA locus, IDDM2 is still the only locus for which genome-wide association is reflected by linkage, which might be explained by disease heterogeneity. Long (class III) and short (class I) variable number tandem repeat (VNTR) alleles are associated, respectively, with higher and lower transcription of proinsulin messenger RNA (mRNA) in medullary thymic epithelial cells (mTECs) under the control of the autoimmune regulator gene (*AIRE*)[20] and with lower and higher T1DA risk.[21] This leads to the inference that IDDM2 controls the extent of deletion of proinsulin-specific T cells during their intrathymic development and predicts that the long VNTR would be associated with less proinsulin-specific T cells in the periphery; there is evidence for and against this, perhaps because of the difficulty of measuring islet antigen-specific T cells in human blood. The *INS* polymorphism is unique to humans. Mice, instead, have two insulin genes; *INSI* is expressed in the β cell and *INSII* in the thymus, and expression of the latter is impaired in the NOD mouse. In summary, IDDM2 provides a genetic mechanism for autoimmunity targeting proinsulin and the β cell.

Many of the other candidate genes are involved in immune function and are associated with other autoimmune diseases. For example, a nonsynonymous gain-of-function (GOF) polymorphism in *PTPN22* (lymphoid tyrosine phosphatase) results in enhanced T-cell suppression, which may impair negative selection of autoreactive T cells in the thymus and decrease the number and function of natural regulatory T cells (nTregs). Polymorphisms around *IL2RA* (interleukin-2 [IL-2] receptor α chain; CD25) and *IL2* are associated with impaired IL-2 signaling, which in both humans and NOD mice impairs the generation and maintenance of nTregs. Polymorphisms involving the *VDR* (vitamin D receptor) and *CYP27B1* (25-hydroxyvitamin vitamin D3 1-α-hydroxylase) and *IFIH1* (IFN-induced with helicase C domain 1; which enhances the type I IFN response to virus infection) are clues to gene–environment interactions in T1DA.

β CELLS

Why are β cells selectively destroyed by immune inflammation that involves the whole islet, which also contains a minority of glucagon-secreting α cells and other endocrine cells? First, cytotoxic CD8 T cells recognize autoepitope peptides (*e.g.,* from insulin) presented by hyperexpressed HLA class I proteins on β cells. Second, because of their unique metabolic wiring, β cells may contribute to their own death at the hands of the immune system. Evidence for the hypothesis of “assisted suicide” is compelling.[22] Studies of rodent islets, mainly *in vitro,* have indicated that β cells lack efficient antioxidant and free radical scavenging mechanisms and are especially sensitive to mitochondrial oxidative and ER stress in response to cytokines and granzymes,[22] although more recent studies of human islets do not support this view.[23] Third, many of the candidate T1D genes identified in GWA studies are transcribed in β cells (see Fig. 71.4) and encode proteins that interact with the immune system.[23] Finally, in response to inflammation, human islets generate hundreds of RNA splice variants for proteins, which, translated as neoantigens, may not be subject to immune tolerance.

Although it is still unclear whether human β cells are more sensitive than other autoimmune target cells to immune

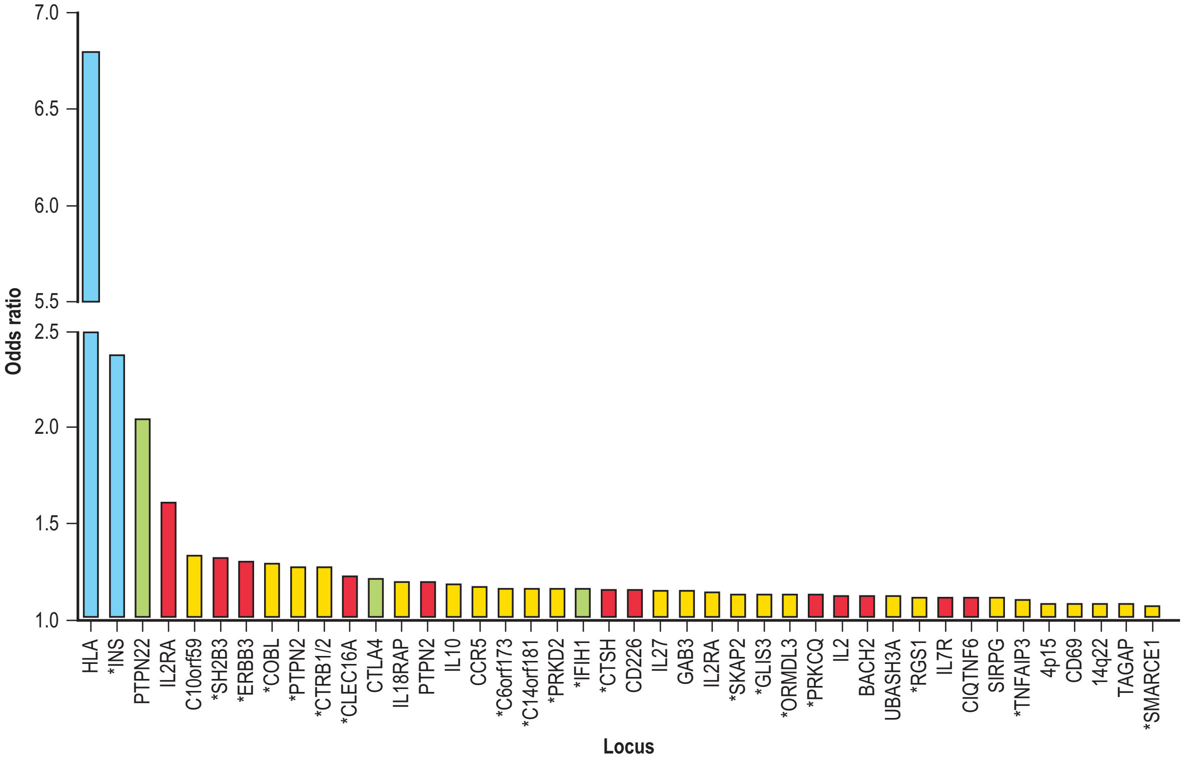


FIG 71.5 Candidate genes in type 1A diabetes (T1DA). The *y*-axis shows the odds ratio for risk alleles at each of the loci on the x-axis.[18] The majority of the known genes are involved in immune responses, and many (marked *) are expressed in islets. Colors indicate era of identification: blue 1970–2000; green 2001–2006; red 2007–2008; yellow 2009–2010.

effectors, what is clear is that having undergone apoptotic or necrotic death, β cells are not restorable in the face of autoimmune memory.

KEY CONCEPTS

Type 1A Diabetes Is an Autoimmune Disease

- Genetic and familial clustering, as well as association with other autoimmune diseases, including autoimmune polyendocrinopathy syndromes (*e.g.,* APS-1 due to mutations in *AIRE*)
- Presence of autoantibodies and T cells reactive with islet cell antigens
- Strong association with specific human leukocyte antigen (HLA) alleles and haplotypes
- Transfer of disease by bone marrow or development of disease in normal pancreas transplanted from identical twin without diabetes to the twin with diabetes
- Neonatal onset with loss of natural regulatory T cells with mutations in *FOXP3* (immune dysfunction/polyendocrinopathy/enteropathy/X-linked [IPEX] syndrome)

ENVIRONMENT—OUTSIDE AND INSIDE

The environment in its various manifestations may impact on β-cell function in several possible ways: by activating innate immune cells (macrophages) in the islets to release cytokines (*e.g.,* IL-1β, TNF) that elicit mitochondrial oxidative ER stress; by driving adaptive immunity to β cells in people with T1D risk genes, either generally or specifically by inducing posttranslational modifications in β-cell proteins that render them immunogenic; and by promoting obesity and insulin resistance and thereby increasing the workload of the β cell.

The environment in Western societies has changed dramatically in many ways during the last century (Fig. 71.6), and some of these changes have been associated epidemiologically or in animal studies with development of T1DA.[6] An index of the modern "exposome" is the ongoing rise in the prevalence of obesity. When children at increased genetic risk for T1DA (with an affected first-degree relative) were monitored from birth, weight gain in the first 2–3 years of life was shown to be a risk factor for islet autoimmunity.[24] When such at-risk children developed islet autoantibodies and were then followed up over time, those with insulin resistance progressed most rapidly to clinical diabetes.[15] Whether insulin resistance is a risk factor for the development of islet autoimmunity is an open question that should be answered by ongoing pregnancy–birth cohort studies.[25] If insulin resistance synergizes with impaired insulin secretion to promote T1DA, then lifestyle factors must be considered in the efforts to forestall or prevent T1DA.

Excessive energy consumption leading to obesity is part of a complex interplay of related factors that promote low-grade inflammation and insulin resistance. These include a poor-quality diet and lack of physical exercise, sunlight (vitamin D), and sleep. The highly processed fast-food "Western" diet lacks diversity of components, lacks plant-derived prebiotics and complex carbohydrates (starches and fiber), is high in saturated fats and in sucrose and fructose sugars, and has added artificial preservatives, emulsifiers, and sweeteners. All of these alter the composition of the gut microbiome and reduce its diversity, which is a feature

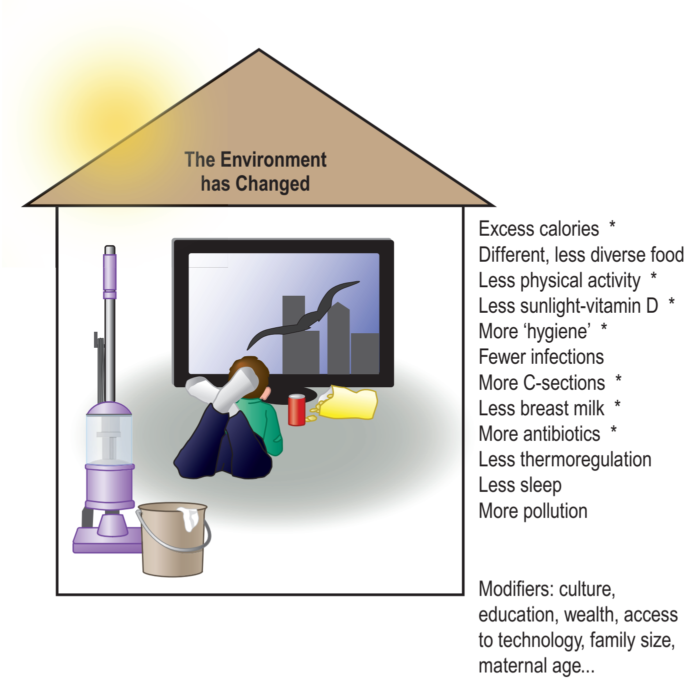


FIG 71.6 Changes in the modern environment. Those that have been associated with an increased risk for type 1A diabetes (T1DA) in humans or nonobese diabetic (NOD) mice are indicated (marked *).

found in children at risk for T1DA.[26] Diets containing a diverse range of plant products (cereals, fresh fruits, and vegetables) provide complex carbohydrates for fermentation by colonic bacteria to short-chain fatty acids (SCFAs), which are antiinflammatory.[27] The dramatic influence of the environment is seen in the comparison between the highly developed Finland and the nearby undeveloped Russian Karelia. Finland has the highest incidence of T1DA in the world, currently 57.6 cases/100 000 population ≤14 years of age (www.diabetesatlas.org). In 2005, there was a sixfold difference in the incidence of T1DA between Finland and Karelia despite overlap in ethnic background and a similar distribution of HLA-DQ risk genotypes.[28] This difference could reflect a relatively lower rate of infections (in Finland), in keeping with the “hygiene hypothesis.”[29]

However, viral infections have long been considered triggers of T1DA, but the evidence remains inconclusive.[30,31] Viral infection is circumstantially associated with the onset of T1DA, but given the long presymptomatic stage of the disease, this may simply represent a nonspecific inflammatory insult that precipitates clinical presentation. Viral mechanisms in T1DA could be direct or indirect (*e.g.,* infection of β cells, infection of the exocrine pancreas with bystander death of β cells, mimicry between T-cell epitopes in a viral protein and β-cell autoantigens, or activation of endogenous retroviruses in β cells by environmental agents). If an exogenous virus was clearly identified, then protective vaccination early in life would be an approach to primary prevention. Interest in viral triggers of T1DA was initiated by the observation that a minority of children with congenital rubella developed insulin-dependent diabetes[32] and were found to have a higher frequency of what was later recognized as the T1D susceptibility haplotype HLA-A1 B8-(DR3-DQ2). Clearly, this may have been an anomaly because congenital rubella has been virtually eliminated by vaccination and yet the incidence of T1D has continued to climb. Much circumstantial evidence exists for enteroviruses in T1DA[31] but has not been consolidated in follow-up studies. Even if a particular enterovirus strain were shown directly to be diabetogenic, vaccination would be elusive because among the many thousands of strain variants, the only one for which a vaccine exists is poliovirus. Mumps virus epidemics have been associated with T1DA onset after 2–4 years, and introduction of a mumps vaccine was associated with a plateau in the rising incidence of T1D in Finland, but this was temporary, and mumps vaccination has clearly not prevented T1D. It must be emphasized that there is no scientific evidence that any form of vaccination itself triggers T1DA.

Rotavirus is the commonest cause of gastroenteritis in young children. The discovery of strong sequence similarities between T-cell epitopes in the VP7 protein of rotavirus and the IA2 and GAD islet antigens in autoantibody-positive children led to speculation that mimicry of rotavirus might contribute to islet autoimmunity. Subsequently, in the Australian BabyDiab Study, rotavirus infection was associated over time with the first appearance of or an increase in islet autoantibodies in children before they developed diabetes.[33] Moreover, rotavirus has been shown to infect β cells in islets from mice, pigs, and monkeys and cause transient involution of the pancreas and hyperglycemia in a Toll-like receptor 3 (TLR3)–dependent manner in mice.[34] Ubiquitous rotavirus infections that drive cross-reactive immunity to islet autoantigens are unlikely to be diabetogenic but could complement and sustain the immune response to direct infection of β cells. As with rubella and mumps, it will be interesting to observe if recently introduced rotavirus vaccines have any effect on the incidence of T1DA.

The microbiome—the trillions of microorganisms (bacteria, fungi, archaea, protozoa, and viruses; Chapter 14) and their millions of genes and proteins that reside within the human mucosae, skin, and secretions—has gained increasing attention as a bellwether of health and disease (“dysbiosis”). Most microbiome analyses are based on DNA extraction and polymerase chain reaction (PCR) amplification of regions of the 16S ribosomal RNA (rRNA) gene that are conserved among bacteria. Although other internal regions of the 16S gene are variable and enable taxonomic classification from the phylum to genus level, 16S sequencing does not have the sensitivity to distinguish individual species and their strains. Direct metagenomic sequencing of all species, as well as analysis of their transcriptomes, epigenomes, and functions, is required to fully comprehend the role of the microbiome in T1DA and other chronic immune–inflammatory diseases. The marked difference in the incidence of T1DA between Finland and neighboring Karelia was found to be associated in Finnish children having decreased bacterial diversity overall, a dominance of the phylum Bacteroidetes, and lack of butyrate-producing bacteria.[26] These changes were seen after the appearance of autoantibodies, suggesting that they followed, rather than preceded, the disease process. However, in a further small study in Finnish children, metagenomic sequencing identified a relative abundance of *Bacteroides dorei,* which peaked around 7–8 months of age with the introduction of solids and preceded the appearance of islet autoantibodies. Gut *Bacteroides* species are abundant in Finnish children, including *B. dorei,* which produces a lipopolysaccharide (LPS) endotoxin that inhibits the immunostimulatory activity of *Escherichia coli* LPS, known to protect against development of diabetes in NOD mice. These findings suggest that an approach to lowering the environmental contribution to T1DA will be intervention with prebiotics, probiotics, or other means to favorably alter the composition of the gut microbiome.

KEY CONCEPTS

Type 1A Diabetes: Epidemiology, Environment, and Genes

- Incidence is increasing as a result of the impact of environment.
- Environment factors enable increasing penetrance of lower-risk human leukocyte antigen (HLA) alleles.
- Environmental factors are multifactorial but generally "pro-inflammatory" and impact via microbiome dysbiosis and epigenetic modifications on polygene expression to cause immune dysregulation in early life.

TREATMENT AND PREVENTION

Treatment of the metabolic syndrome of T1D is focused on optimizing blood glucose control with various modes of insulin delivery to prevent the short- and long-term complications of hyperglycemia. Refinements and advances over the past 30 years in insulin delivery and improved control of blood glucose and blood pressure have greatly improved the lives and prognosis of individuals with T1DA. Nevertheless, despite the introduction of pure, recombinant forms of human insulin, which have varied kinetics of action and can be injected via a syringe or continuously infused via a pump, and blood glucose self-monitoring, which can now be done continuously in real time, those with T1DA have not been rescued from a life sentence, and their blood glucose control remains less than physiological. The "cure" of T1DA requires the transplantation of insulin-secreting cells or their progenitors or the regeneration of β cells *in situ,* in conjunction with approaches to prevent allograft rejection and/or recurrence of autoimmunity (there will be no cure for T1DA without prevention). Currently, in some economically developed societies, allografts of whole pancreas or isolated islets are offered on a limited scale and at considerable cost to individuals with T1DA and life-threatening complications, in particular lack of awareness of hypoglycemia that is unresponsive to conventional treatment. However, shortage of donor tissue, cost, and the need for life-long immune suppressive agents militate against this solution. Longer-term, genetically engineered pig islet xenografts may overcome the tissue supply barrier. Stem cells, in particular autologous induced pluripotent stem cells (iPSCs), remain the great hope, but early proof-of-concept results in rodents await scale-up and translation to humans.

The incidence of spontaneous autoimmune diabetes in inbred NOD mice is decreased by many immune and other (including environmental) interventions introduced early in the disease course, although various interventions are often only effective in a proportion of mice, sometimes only delay disease onset, and sometimes have no effect (and are therefore not reported). Nevertheless, these findings in the NOD mouse suggest that T1DA is potentially preventable in outbred humans. On the basis of the key role of proinsulin in driving β-cell autoimmunity, the NOD mouse has provided "proof-of-concept" for the efficacy of antigen-specific vaccination strategies.[35] In humans, it has been convenient to classify prevention into primary (before the onset of islet autoimmunity), secondary (after the onset of islet autoimmunity), and tertiary (after the onset of clinical disease) prevention. Of course, real prevention is primary prevention—identifying those at genetic risk and intervening to avert islet autoimmunity. Neonatal screening based on HLA genotyping to identify at-risk individuals would be a basis for primary prevention, but even in countries with a high incidence of T1DA, this would have only modest predictive value, and intervention could be justified if it was practical and safe, if not efficacious. The mandate *Primum non nocere* ("first do no harm") dictates that many immune modulating agents are off limits for either primary or secondary prevention in T1DA. Options (*e.g.,* to modification of the diet or microbiome or of forms of vaccination) are greatly restricted. Agents that are effective in other autoimmune diseases are more likely to be effective in the early asymptomatic stage of T1DA, not after clinical presentation when most β cells have been destroyed. Indeed, clinical trials of tertiary prevention with more than 70 different agents since the early 1980s have failed to unequivocally demonstrate sustained preservation of residual β-cell function.[36,37] Interestingly, the number of trials reporting a positive outcome decreased from the 1980s, together with a shift in the primary outcome measure away from clinical remission to the more rigorous measure of residual insulin (stimulated C-peptide) secretion.

Is there a potential solution to the "off limits" dilemma? Classically, T1DA is regarded as a metabolic disorder diagnosed at clinical presentation with symptoms and signs of hyperglycemia. However, based on pathology, T1DA is an immune disease—autoimmune β-cell disorder (ABCD)—and only becomes a metabolic disorder at the end stage of its pathology. Accepting this reality would amount to a paradigm shift that could profoundly change the approach to and outcome of secondary prevention. Consider the situation in other autoimmune diseases. In rheumatoid arthritis, the pathology and clinical features (painful swollen joints) appear concomitantly when effective disease-sparing treatment with newer biological agents is initiated, *not* at the end stage of pathology when the joints are irreversibly deformed. A further benefit of diagnosing of T1DA early in the asymptomatic stage, based on evidence of underlying pathology, is that it would markedly reduce the risk of life-threatening ketoacidosis associated with the classic symptomatic presentation.

In health, immune responses to autoantigens are regulated to prevent development of autoimmune diseases. Autoantigen-specific immunotherapy aims to boost or restore autoantigen-specific immunoregulatory mechanisms. Its parallel, allergen-specific immunotherapy, has been shown in randomized trials to be effective in allergic asthma and rhinitis. Such "negative vaccination" can be achieved in several ways: by administering antigen via a "tolerogenic" route (mucosal, dermal), cell type (resting dendritic cell), mode (with blockade of costimulation molecules), or form (as an "altered peptide ligand"). Mechanisms of antigen-induced tolerance include deletion and/or anergy of effector T cells and induction of regulatory T cells (iTregs). Of clinical importance is the ability of iTregs to exert antigen nonspecific "bystander" suppression. Thus by direct cell contact and/or the release of soluble immunosuppressive factors, such as IL-10, iTregs impair the function of antigen-presenting DCs to elicit effector T-cell responses to the same or other antigens presented locally in the lesion or draining lymph nodes. Bystander suppression does not require that the "tolerizing" antigen is necessarily the primary driver of pathology. Its clinical importance is that it obviates specificity restrictions imposed by polymorphisms in the HLA and human T-cell receptors.

In NOD mice, administration of insulin, proinsulin peptides, or proinsulin DNA via oral or nasorespiratory routes, acting locally on the mucosal immune system, induces Tregs and decreases the incidence of diabetes. Results of randomized secondary prevention trials of insulin or GAD in relatives at risk for T1DA (ClinicalTrials.gov) have been summarized elsewhere.[35-37] In the DPT-1 oral insulin trial, relatives with a 5-year diabetes risk of 25–50% received 7.5 mg human insulin or placebo daily for a median of 4.3 years. There was no effect overall, but posttrial hypothesis testing revealed a significant delay in diabetes onset in participants with significant IAA at entry.[38] This outcome has led to an ongoing follow-up international trial by NIH-TrialNet using the same dose of 7.5 mg daily, which is not optimal because on a body-weight basis this dose is very small compared with that required to induce protective iTregs in the NOD mouse.

Apart from the dose, other variables have not been systematically tested in humans, in part because of the expense and duration of prevention trials. These variables include route of administration, form of the antigen, combinations of antigen with antigen-nonspecific agents, and the nature of induced immune responses. Unfortunately, surrogate markers of a potential therapeutic response have not been included in most trials and, in the case of antigen-specific T cells in blood, are not universally accepted as being sufficiently robust and reproducible. Oral delivery might not be optimal because proteins are degraded after ingestion, and the concentration or form of peptide reaching the upper small intestine may be variable and unpredictable. T-cell responses that were observed after nasorespiratory administration of a peptide have not been observed after oral administration. In a randomized trial of nasal insulin in individuals with recent-onset T1D not initially requiring insulin treatment, participants in the nasal insulin arm had markedly blunted insulin antibody responses after subsequent subcutaneous insulin.[39] This was the first demonstration, in humans, of immune regulation induced by mucosally administered exogenous autoantigen. Although this result provides a rationale for further disease endpoint trials, a Finnish study[40] suggested that this result cannot be extrapolated to endogenous "autoantigenic" insulin. Nasal insulin (1 unit/kg daily) did not alter the rate of progression to diabetes in islet autoantibody-positive children <3 years of age. However, these children were at very high risk, and many appeared to have had borderline β-cell function. As in the oral insulin trial, markers of an insulin bioeffect and evidence of immune tolerance were not reported. In the ongoing Intranasal Insulin Trial II (INIT II) in Australia, New Zealand, and Germany, higher doses of nasal insulin (44 or 440 IU) are being administered weekly for a year to relatives with T1D who have autoantibodies to at least two islet antigens (≈40% risk of diabetes over 5 years). The primary outcome is clinical diabetes; secondary outcomes are measures of metabolic and immune function. In INIT II, the insulin dose is substantially higher than in the Finnish study, the participants are older and have less advanced subclinical disease and therefore are at lower risk. The promise of antigen-specific therapy, therefore, has yet to be realized in humans. If a balance between pathogenic and protective T cells is deterministic for disease development, then antigen-specific therapy is most likely to be effective as a primary preventive strategy. Once disease has been initiated, insulin-specific or other antigen-specific approaches may require complementary therapy with other agents to inactivate or delete the burden of pathogenic effector cells.

ON THE HORIZON

- Closed-loop insulin delivery—blood glucose monitoring devices will become more refined, cheaper, and widely available.
- "Linking the -omes" (exposome–microbiome–metabolome–epigenome) across time in early life will provide new insights into how environment–gene interactions lead to immune dysregulation and type 1A diabetes (T1DA).
- Manipulation of the gut microbiome via scientifically formulated probiotics and prebiotics will be used for primary prevention of T1DA.
- T1DA is primarily an immune disease—autoimmune β-cell disorder (ABCD)—and only secondarily a metabolic disorder. Recognition of these disease components will expand the therapeutic possibilities for secondary prevention.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. The Expert Committee on the Diagnosis and Classification of Diabetes Mellitus. Report of the expert committee on the diagnosis and classification of diabetes mellitus. Diabetes Care 2003;26:S5–20.
2. Verge CF, Gianani R, Kawasaki E, et al. Prediction of type I diabetes in first-degree relatives using a combination of insulin, GAD, and ICA512bdc/IA-2 autoantibodies. Diabetes 1996;45:926–33.
3. Colman PG, Steele C, Couper JJ, et al. Islet autoimmunity in infants with a type I diabetic relative is common but is frequently restricted to one autoantibody. Diabetologia 2000;43:203–9.
4. Bingley PJ, Williams AJ, Gale EA, et al. Optimized autoantibody-based risk assessment in family members. Implications for future intervention trials. Diabetes Care 1999;22:1796–801.
5. Ziegler AG, Rewers M, Simell O, et al. Seroconversion to multiple islet antibodies and risk of progression to diabetes in children. JAMA 2013;309:2473–9.
6. Wentworth JM, Fourlanos S, Harrison LC. Deconstructing the stereotypes of diabetes within the modern diabetogenic environment. Nat Rev Endocrinol 2009;5:483–9.
7. Fourlanos S, Dotta F, Greenbaum CK, et al. Latent autoimmune diabetes in adults (LADA) should be less latent. Diabetologia 2005;48:2206–12.

8. Balasubramanyam A, Nalini R, Hampe CS, et al. Syndromes of ketosis-prone diabetes mellitus. Endocr Rev 2008;29:292–302.
9. Imagawa A, Hanafusa T, Miyagawa J, et al. A novel subtype of type 1 diabetes mellitus characterized by a rapid onset and an absence of diabetes-related antibodies. Osaka IDDM Study Group. N Engl J Med 2000;342:301–7.
10. Imagawa A, Hanafusa T. Pathogenesis of fulminant type 1 diabetes. Rev Diabet Stud 2006;3:169–77.
11. Harrison L, Honeyman M, DeAizpurua H, et al. Inverse relationship between humoral and cellular immunity to glutamic acid decarboxylase in humans at-risk for insulin-dependent diabetes. Lancet 1993;341:1365–9.
12. Gale EAM. The rise of childhood type 1 diabetes in the 20th century. Diabetes 2002;51:3353–61.
13. Fourlanos S, Varney MD, Tait BD, et al. The rising incidence of type 1 diabetes is accounted for by cases with lower risk human leukocyte antigen genotypes. Diabetes Care 2008;31:1546–9.
14. Tait BD, Colman PG, Morahan G, et al. HLA genes associated with autoimmunity and progression to disease in type 1 diabetes. Tissue Antigens 2003;61:146–53.
15. Fourlanos S, Narendran P, Byrnes GB, et al. Insulin resistance is a risk factor for progression to type 1 diabetes. Diabetologia 2004;47:1661–7.
16. Eisenbarth GS. Type I diabetes mellitus. A chronic autoimmune disease. N Engl J Med 1986;314:1360–8.
17. Willcox A, Richardson SJ, Bone AJ, et al. Analysis of islet inflammation in human type 1 diabetes. Clin Exp Immunol 2009;155:173–81.
18. Barrett JC, Clayton DG, Concannon P, et al. Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. Nat Genet 2009;41:703–7.
19. Pociot F, Akolkar B, Concannon P, et al. Genetics of type 1 diabetes: What's next? Diabetes 2010;59:1561–71.
20. Pugliese A, Zeller M, Fernandez A Jr, et al. The insulin gene is transcribed in the human thymus and transcription levels correlated with allelic variation at the INS VNTR-IDDM2 susceptibility locus for type 1 diabetes. Nat Genet 1997;15:293–7.
21. Bennett ST, Lucassen AM, Gough SC, et al. Susceptibility to human type 1 diabetes at IDDM2 is determined by tandem repeat variation at the insulin gene minisatellite locus. Nat Genet 1995;9:284–92.
22. Atkinson MA, Bluestone JA, Eisenbarth GS, et al. How does type 1 diabetes develop?: the notion of homicide or beta-cell suicide revisited. Diabetes 2011;60:1370–9.
23. Eizirik DL, Sammeth M, Bouckenooghe T, et al. The human pancreatic islet transcriptome: expression of candidate genes for type 1 diabetes and the impact of pro-Inflammatory cytokines. PLoS Genet 2012;8: e1002552.
24. Couper JJ, Beresford S, Hirte C, et al. Weight gain in early life predicts risk of islet autoimmunity in children with a first degree relative with type 1 diabetes. Diabetes Care 2009;32:94–9.
25. Penno MAS, Couper JJ, Craig ME, et al. Environmental determinants of islet autoimmunity (ENDIA): a pregnancy to early life cohort study in children at-risk of type 1 diabetes. BMC Pediatr 2013;13:124.
26. Knip M, Siljander H. The role of the intestinal microbiota in type 1 diabetes mellitus. Nat Rev Endocrinol 2016;12:154–67.
27. Thorburn AN, Macia L, Mackay CR. Diet, metabolites, and "Western lifestyle" inflammatory diseases. Immunity 2014;40:833–42.
28. Kondrashova A, Reunanen A, Romanov A, et al. A six-fold gradient in the incidence of type 1 diabetes at the eastern border of Finland. Ann Med 2005;37:67–72.
29. Kondrashova A, Seiskari T, Ilonen J, et al. The 'hygiene hypothesis' and the sharp gradient in the incidence of autoimmune and allergic diseases between Russian Karelia and Finland. APMIS 2013;121:478–93.
30. Honeyman M. How robust is the evidence of viruses in the induction of type 1 diabetes? Curr Opin Immunol 2005;17:616–23.
31. Roivainen M, Klingel K. Virus infections and type 1 diabetes risk. Curr Diab Rep 2010;10:350–6.
32. Menser MA, Forrest JM, Honeyman MC, et al. Letter: diabetes, HL-A antigens, and congenital rubella. Lancet 1974;2:1508–9.
33. Honeyman MC, Coulson BS, Stone NL, et al. Association between rotavirus infection and pancreatic islet autoimmunity in children at-risk for type 1 diabetes. Diabetes 2000;49:1319–24.
34. Honeyman MC, Laine D, Londrigan S, et al. Rotavirus infection induces transient pancreatic involution and hyperglycemia in weanling mice. PLoS ONE 2014;9:e106560.
35. Harrison LC, Wentworth JM, Zhang Y, et al. Antigen-based vaccination and prevention of type 1 diabetes. Curr Diab Rep 2013;13:616–23.
36. Harrison LC, Wentworth JM. Prevention of autoimmune disease: the type 1 diabetes paradigm. In: Rose N, Mackay I, editors. The autoimmune diseases. 5th ed. London, UK: Elsevier Ltd.; 2014. p. 1191–203.
37. Staeva TP, Chatenoud L, Insel R, et al. Recent lessons learned from prevention and recent-onset type 1 diabetes immunotherapy trials. Diabetes 2013;62:9–17.
38. Skyler JS, Krischer JP, Wolfsdorf J, et al. Effects of oral insulin in relatives of patients with type 1 diabetes: the Diabetes Prevention Trial–Type 1. Diabetes Care 2005;28:1068–76.
39. Fourlanos S, Perry C, Gellert SA, et al. Evidence that nasal insulin induces immune tolerance to insulin in adults with autoimmune diabetes. Diabetes 2011;60:1237–45.
40. Nanto-Salonen K, Kupila A, Simell S, et al. Nasal insulin to prevent type 1 diabetes in children with HLA genotypes and autoantibodies conferring increased risk of disease: a double-blind, randomized controlled trial. Lancet 2008;372:1746–55.

MULTIPLE-CHOICE QUESTIONS

1. Type 1 diabetes (T1D):
   - **A.** Is a genetic disease in which environment plays little or no role
   - **B.** Is classified as type 1A or type 1B, depending on evidence for an autoimmune basis (*i.e.*, the presence of absence respectively of islet autoantibodies)
   - **C.** Cannot be diagnosed on the basis of clinical features alone
   - **D.** Is associated with increased diversity of gut bacterial microbiota

2. In T1DA:
   - **A.** The dominant genetic contribution remains to be established.
   - **B.** The incidence has increased along with increased penetrance of lower-risk genetic alleles.
   - **C.** Most newly diagnosed cases have a first- or second-degree relative with the disease.
   - **D.** Most children who develop the disease have islet autoantibodies after 5 years of age.

3. In T1DA:
   - **A.** The absence of islet autoantibodies excludes the diagnosis.
   - **B.** Blood glycated hemoglobin (Hb_{A1c}) can be used to make the diagnosis children.
   - **C.** Immune-modulating agents have been successful in stabilizing or reversing β-cell failure for at least 2 years after the diagnosis.
   - **D.** Strict blood glucose control in pregnancy prevents fetal complications.

72

Immunological Lung Diseases

Joyce S. Lee, Andrew P. Fontenot

The lung serves as an interface between the environment and the sanctuary of the body. The defense systems of the upper airways clear the majority of inhaled particulates. Those that evade the upper airway defenses are combated by the innate and acquired immune responses. Essentially all autoimmune diseases are dependent on the inappropriate activation of autoreactive CD4 T cells as well as autoreactive B cells responsible for pathogenic autoantibodies. Immunological lung diseases develop when the normal mechanisms of immune self-tolerance fail. This chapter deals with the pulmonary manifestations of these disorders.

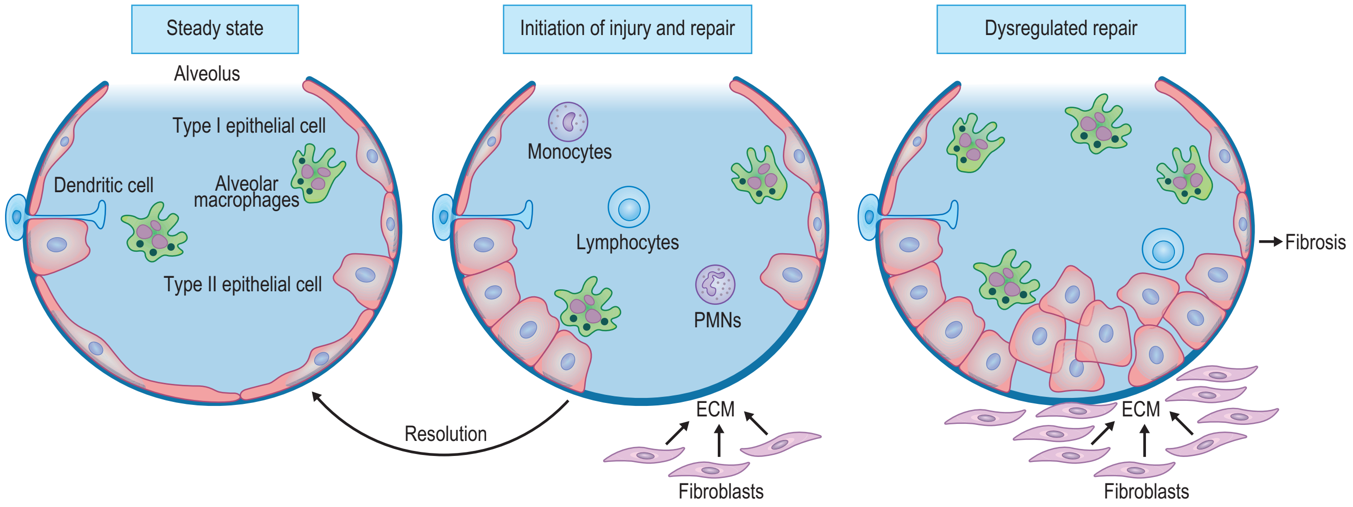
INFLAMMATION IN THE PATHOGENESIS OF INTERSTITIAL LUNG DISEASE

In the normal host, the resident alveolar macrophage is the predominant cell type in fluid from bronchoalveolar lavage (BAL). Resident alveolar macrophages function to ingest and degrade the inhaled antigenic load, to clear pulmonary surfactant, and to tonically suppress the development of inappropriate immune responses. Relatively few lymphocytes are present in the normal lung parenchyma. However, after stimulation by the relevant antigen in the lung-draining lymph nodes, antigen-specific lymphocytes migrate to the lung and participate in the inflammatory response. In addition to lymphocytes, other inflammatory and immune cells, including neutrophils, eosinophils, and other mononuclear cells, accumulate in the lung of patients with immunological lung disease, depending on the underlying disease. Within the normal alveolus, the major cellular constituents are alveolar macrophages and epithelial cells. Resident dendritic cells (DCs) also project dendrites through the tight junctions of the alveolar epithelium to sample the alveolar space. An initial insult typically involves the alveolar epithelial cell. Damage to type I alveolar epithelial cells results in release of chemokines that recruit and activate inflammatory cells, allowing for resolution of inflammation and repair of injured tissue (Fig. 72.1). With either prolonged exposure or failure to adequately clear an inhaled antigen, persistent inflammation results in extracellular matrix (ECM) deposition, culminating in tissue remodeling, progressive collagen deposition, and pulmonary fibrosis. Impaired ventilation and gas exchange occur as a consequence of lung fibrosis, resulting in patient morbidity and mortality. With destruction of type I alveolar epithelial cells, exposure of the underlying basement membrane can cause further inflammation. Proper restoration of the epithelial barrier is critical for resolution of lung inflammation. Type II alveolar epithelial cells can serve as progenitor cells that migrate and differentiate into type I alveolar epithelial cells to reestablish an intact lung epithelium.

Macrophages and lymphocytes have also been localized to areas of pulmonary fibrosis in patients with immunological lung diseases, including idiopathic pulmonary fibrosis (IPF), systemic sclerosis, and rheumatoid arthritis (RA).[1] In the lung, macrophages can be divided, according to their location, into alveolar or interstitial macrophages. Resident alveolar macrophages are yolk-sac–derived cells that are dependent on interleukin-34 (IL-34), colony-stimulating factor 1 (CSF-1), and the transcription factor PU.1 for their development.[2-4] These yolk-sac–derived macrophages are self-renewing,[5] maintained at a stable number throughout life, and are localized on epithelial surfaces and lining fluid of the alveoli and airways. With relatively poor antigen-presenting ability, they function to remove inhaled particles and bacteria. Conversely, interstitial macrophages are derived from hematopoietic stem cells (HSCs)[6] and are located in the tissue spaces between the alveoli. In the context of lung injury, interstitial macrophages greatly increase in number. Although interstitial macrophages have less phagocytic activity than alveolar macrophages, they have increased ability to present antigens to T cells. Upon activation, these macrophages express a variety of cytokines, including tumor necrosis factor-α (TNF-α) and monocyte chemotactic protein-1 (MCP-1). In addition to antigen presentation to T cells, macrophages are important for lung fibrosis and tissue remodeling in immunological lung diseases through secretion of specific growth factors, such as transforming growth factor-β (TGF-β), insulin-like growth factor-1 (IGF-I), platelet-derived growth factor (PDGF), and fibroblast growth factor (FGF).

T cells are also associated with lung fibrosis. For example, T cells have been located in areas of interstitial fibrosis and honeycombing with relatively fewer in areas of normal lung in patients with IPF.[7] In connective tissue disease–associated interstitial lung disease (ILD), T cells are diffusely distributed throughout the lung and within focal lymphoid aggregates.[8] In animal models, depending on their phenotype, T cells can be either profibrotic or antifibrotic. γδ T cells decrease lung inflammation and fibrosis in multiple mouse models.[9] CD8 T cells have been found in high proportion in BAL and surgical lung biopsy samples from patients with IPF, as well as patients with systemic sclerosis, but their role in these immunological lung diseases is not well understood.

Different subsets of CD4 T cells (Chapter 16) have been implicated in the pathogenesis of immunological lung diseases. T-helper type 1 (Th1) cells express interferon-γ (IFN-γ). Although IFN-γ is a potent proinflammatory cytokine, it has antifibrotic effects through inhibiting fibroblast proliferation and collagen expression. Th2 cells are defined by expression of IL-4, IL-5, and



1. Denudation of barrier
2. Cytokine/chemokine expression
3. Recruitment of inflammatory cells
4. Proliferation of type II epithelial cells
5. DC maturation
6. Fibroblasts express ECM

FIG 72.1 Pathogenesis of Interstitial Lung Disease. In the healthy lung, the alveoli maintain an antiinflammatory state to prevent unwanted inflammation. The predominant cells in the healthy alveolus are macrophages and types I and II epithelial cells. Upon injury, there is denudation of the epithelial barrier, resulting in the expression of cytokines/chemokines, recruitment of inflammatory cells, proliferation of type II epithelial cells, dendritic cell (DC) maturation, and expression of extracellular matrix (ECM) by fibroblasts. This inflammatory milieu leads to the resolution of inflammation and repair of lung injury. Conversely, in the presence of either prolonged antigen exposure or an inability to clear antigen, persistent inflammation ensues, resulting in extracellular matrix deposition with subsequent tissue remodeling, fibrosis, and permanent lung dysfunction.

IL-13. In contrast to IFN-γ, Th2 cytokines have been shown to promote lung fibrosis. Therefore the balance between Th1 and Th2 T cells through expression of different cytokines affects the development of pulmonary fibrosis.[10] Th17 cells express IL-17A and IL-17F, which are potent inflammatory cytokines important for the recruitment of neutrophils to areas of inflammation. Th17 cells have been implicated in the development of lung fibrosis in murine models.[11,12] Conversely, regulatory T cells (Tregs) suppress pathogenic T-cell responses that promote inflammation (Chapter 18). In patients with IPF, Tregs may be less able to suppress expression of cytokines by Th1 and Th2 cells, suggesting that Tregs are important in regulating ILD and pulmonary fibrosis.[13]

IDIOPATHIC INTERSTITIAL PNEUMONIAS

Idiopathic interstitial pneumonias (IIPs) are a group of diffuse inflammatory and/or fibrotic lung disorders that include IPF, acute interstitial pneumonitis (AIP), desquamative interstitial pneumonitis (DIP), respiratory bronchiolitis–associated interstitial lung disease (RB-ILD), nonspecific interstitial pneumonia (NSIP), cryptogenic organizing pneumonia (COP), lymphocytic interstitial pneumonia (LIP), and idiopathic pleuroparenchymal fibroelastosis (IPPFE).[14] The diagnosis of IIPs requires the exclusion of connective tissue diseases (CTDs), drug toxicity, and environmental exposures, and a thoracoscopic lung biopsy is often required. In addition, some patients with IIPs have clinical features suggesting an underlying autoimmune disorder. However, these subjects do not meet established criteria for a CTD. A joint European Respiratory Society and American Thoracic Society Task Force recently proposed the term “interstitial pneumonia with autoimmune features” (IPAF) as well as classification criteria for

KEY CONCEPTS

Classification of the Idiopathic Interstitial Pneumonias

- The idiopathic interstitial pneumonias comprise a group of eight histologically distinct disorders.
- The distinction is important since response to treatment and prognosis differs:
    - Idiopathic pulmonary fibrosis (IPF) with histopathology labeled as usual interstitial pneumonitis (UIP): the pathology of UIP can occur in other diseases such as connective tissue diseases
    - Acute interstitial pneumonitis (AIP)
    - Desquamative interstitial pneumonitis (DIP)
    - Respiratory bronchiolitis–associated interstitial lung disease (RB-ILD)
    - Nonspecific interstitial pneumonia (NSIP)
    - Lymphocytic interstitial pneumonia (LIP)
    - Cryptogenic organizing pneumonia (COP)
    - Idiopathic pleuroparenchymal fibroelastosis (IPPFE)
- To differentiate between these diseases, a thoracoscopic lung biopsy is often required.

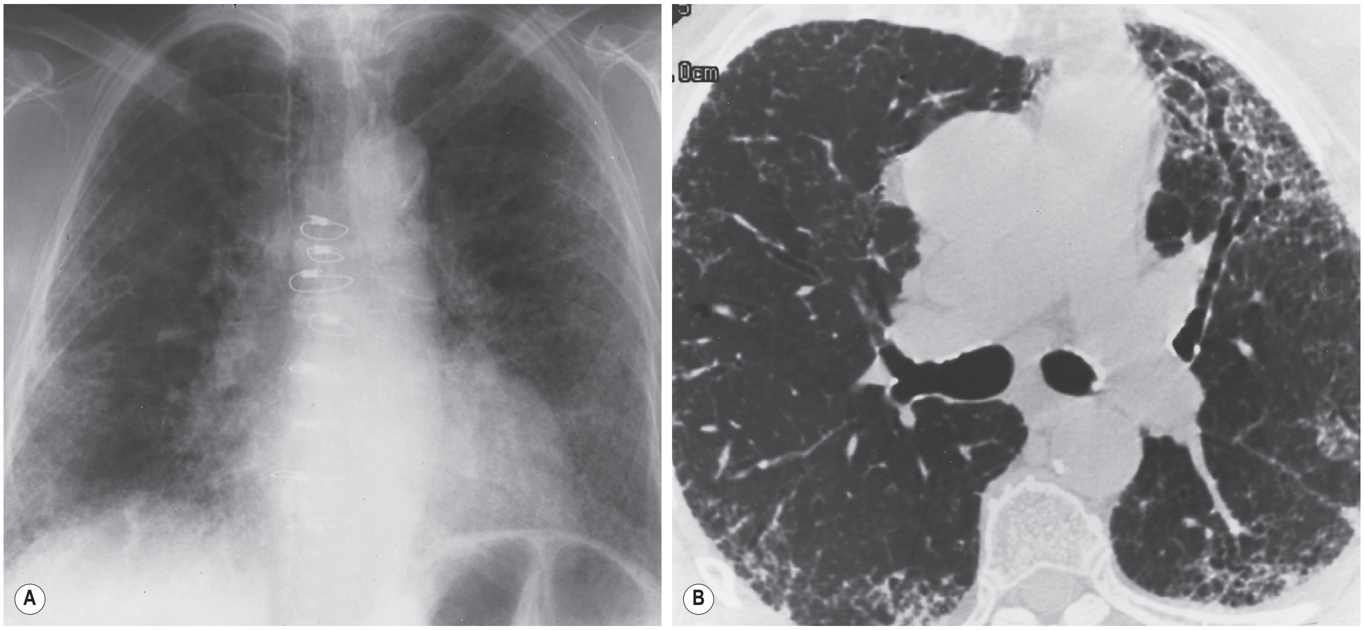


FIG 72.2 Radiographic Manifestations of Idiopathic Pulmonary Fibrosis. (A) Chest radiograph in a patient with idiopathic pulmonary fibrosis showing diffuse, coarse reticular opacities with a lower lung zone predominance. Cystic radiolucencies, consistent with honeycombing, are evident. (B) High-resolution computed tomography shows peripheral reticular opacities, honeycombing, and traction bronchiectasis.

these subjects.[15] Importantly, the distinction between these various disorders is of clinical relevance, since response to treatment and outcome differs.

Idiopathic Pulmonary Fibrosis

IPF, also known as *cryptogenic fibrosing alveolitis,* is the most common diffuse parenchymal lung disease of unknown etiology.[16] It is characterized by progressive clinical deterioration despite available therapy. Although IPF has characteristic clinical, radiographic, and histological appearances, other ILDs, including the CTDs, drug reactions, and environmental exposures, can mimic these findings.

Clinical Manifestations

CLINICAL PEARLS

Idiopathic Pulmonary Fibrosis (IPF)

- IPF is one of the most common causes of diffuse parenchymal lung disease of unknown etiology and is characterized by insidious onset of cough and dyspnea.
- The histopathological pattern of IPF is usual interstitial pneumonitis.
- A confident diagnosis of IPF based on high-resolution computed tomography can only be made in one-third of cases.
- IPF is generally a fatal disorder, characterized by relentless progression and a 5-year survival of 30–50%.
- Pirfenidone and nintedanib appear to slow disease progression in IPF.

The incidence and prevalence of IPF are uncertain, although prevalence rates for men and women of 20.2 per 100 000 and 13.2 per 100 000, respectively, have been reported.[17] Both the incidence and prevalence of IPF increase with age, with most patients presenting between 50 and 70 years of age. Although the clinical features of IPF are variable, most patients present with the insidious onset of exertional dyspnea and a dry, nonproductive cough. Physical examination reveals dry, end-inspiratory crackles with clubbing present in 25–50% of patients.

Chest radiography typically shows diffuse reticular opacities, predominantly in the peripheral lower lung zones. Ground-glass opacities, small cysts (honeycombing), and reduced lung volumes may also be seen (Fig. 72.2A). These radiographic changes often precede the onset of symptoms, and serial chest radiographs usually reveal progressive loss of lung volume. High-resolution computed tomography (HRCT) findings include bibasal peripheral reticular opacities (see Fig. 72.2B). Honeycombing, traction bronchiectasis, and subpleural fibrosis can also be present.

The typical physiological abnormalities in IPF are those of a restrictive lung disease with a low diffusing capacity for carbon monoxide and severe gas exchange abnormalities exacerbated by exercise.

Histopathology

The gross appearance of the lungs in IPF shows a nodular pleural surface, and histopathological examination reveals usual interstitial pneumonitis (UIP).[18] UIP is characterized by nonuniform and variable distribution of the interstitial changes. At low magnification, alternating zones of interstitial fibrosis, inflammation, honeycombing, and normal lung can be seen (Fig. 72.3A). At higher magnification, derangement of alveolar walls with edema, fibrinous exudate, fibroblast proliferation, and fibrosis occur. *Honeycomb change* refers to enlarged airspaces lined by metaplastic bronchial epithelium and surrounded by walls thickened with collagen (see Fig. 72.3B). The earliest finding in UIP is the fibroblast focus, a lesion consisting of distinct clusters of fibroblasts and myofibroblasts in a loose connective tissue matrix within the alveolar wall, with minimal interstitial inflammation or intraalveolar macrophage accumulation (see Fig. 72.3C).[18]

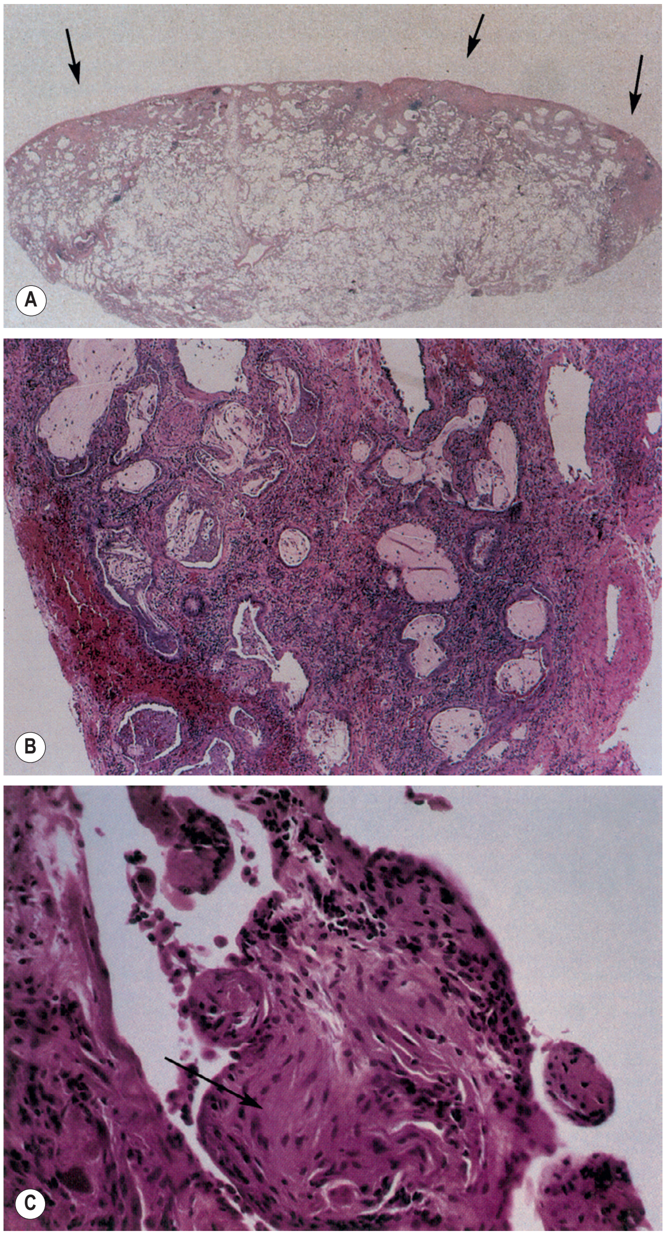


FIG 72.3 **Histopathology of Usual Interstitial Pneumonitis (UIP).** (A) Low-magnification photomicrograph of UIP showing the variegated appearance from one field of view to the next with areas of dense subpleural fibrosis (arrows) separated from other areas of normal lung. (B) High-magnification photomicrograph of UIP showing honeycomb change characterized by enlarged airspaces filled with mucin and separated by fibrosis. (C) Fibroblast focus in UIP is characterized by clusters of spindle-shaped fibroblasts (arrow) in a loose connective tissue matrix within the alveolar wall.

Tables 72.1 and 72.2 compare the clinical and pathological features of UIP, DIP, RB-ILD, and NSIP.

Pathogenesis

The pathogenesis of IPF is poorly understood, but current evidence suggests that fibrosis results from aberrant wound healing in response to repetitive injury and is mediated through an interplay between immunological, genetic, environmental, and viral factors (Fig. 72.4).[16] Some cases of IPF are familial, inherited as an autosomal dominant trait with variable penetrance. Recently, dysregulated expression of a mucin gene, *MUC5B,* has been associated with development of familial IPF.[19] Mutations in the telomerase ribonucleoprotein complex associated with telomere shortening have also been linked with familial IPF.[16] In addition, surfactant protein C mutations have been rarely associated with IPF.

In the normal lung, the interstitium is thin and delicate with few lymphoid cells and fibroblasts. Following the initiation of the inflammatory process, damage to the alveolar epithelial cell occurs, followed by vascular leak, fibroblast activation and proliferation, ECM synthesis, and activation of the innate immune system.[16] The release of danger-associated molecular patterns from dead or dying cells results in macrophage activation. Following activation, alveolar macrophages secrete IL-1, IL-8, tumor necrosis factor-α (TNF-α), PDGF, and IGF-1. This cytokine milieu promotes the activation and recruitment of neutrophils and lymphocytes to the area of alveolitis.

T lymphocytes, which accumulate in the alveolar space and interstitium, express an activated phenotype, including the expression of human leukocyte antigen D-related (HLA-DR) and IL-2 receptor. Following activation, CD4 T cells evolve into three major subsets distinguished by the cytokines produced (Chapters 9 and 16). In IPF, T cells expressing a Th2-type phenotype predominate, producing IL-4, -5, and -13. In addition, the Th17 cytokine, IL-17A, has been linked to the development of bleomycin-induced lung injury and collagen deposition.[11,12] Evidence also suggests that patients with IPF have oligoclonal CD4 T-cell expansions that proliferate in response to antigens present in diseased tissue.[20] Treg function may be impaired in patients with IPF.[13] In addition, immune complexes have been identified in the serum and lungs of patients with IPF.[21]

In addition to their role as scavengers, alveolar macrophages are vital in the repair phase of inflammation. However, the distinguishing feature between a self-resolving inflammatory process and a fibrotic response, as seen in IPF, is the accumulation of collagen. Evidence suggests that the fibrotic process in IPF is a consequence of dysregulation of both collagen synthesis and degradation. Macrophage-derived growth factors, including TGF-β, PDGF, and IGF-1, stimulate fibroblast proliferation and collagen deposition.[22] Adequate resolution of an inflammatory process requires matrix degradation. Matrix metalloproteases (MMPs) produced by macrophages and fibroblasts are involved in matrix degradation, and control of metalloprotease production involves substances known as *tissue inhibitors of metalloproteases* (TIMPs). TIMPs are elevated in the lungs of patients with IPF. In addition, TGF-β can markedly augment TIMP production. Thus there appears to be a loss of balance between the events mediating resolution and those mediating perpetuation of the inflammatory response, setting the stage for lung injury, tissue remodeling, and the development of irreversible pulmonary fibrosis.

Diagnosis

The diagnostic evaluation of a patient with diffuse parenchymal lung disease includes a thorough history and physical examination with particular attention to symptoms and signs that could indicate a CTD, occupational and environmental exposures, or medication and drug use. A careful family history is also important.

TABLE 72.1 Clinical Features of the Idiopathic Interstitial Pneumonias

	UIP	DIP	RB-ILD	AIP	NSIP
Mean age (years)	57	42	36	49	49
Childhood	No	Rare	No	Rare	Occasionally
Onset	Insidious	Insidious	Insidious	Acute	Subacute, insidious
Mortality (mean survival)	68% (5–6 years)	27% (12 years)	0%	62% (1–2 months)	11% (17 months)
Response to steroids	Poor	Good	Good	Poor	Good
Recovery possible	No	Yes	Yes	Yes	Yes

AIP, acute interstitial pneumonitis; DIP, desquamative interstitial pneumonitis; NSIP, nonspecific interstitial pneumonia; RB-ILD, respiratory bronchiolitis–associated interstitial lung disease; UIP, usual interstitial pneumonitis.

Adapted from Katzenstein AL, Myers JL. Idiopathic pulmonary fibrosis: clinical relevance of pathological classification. Am J Respir Crit Care Med 1998; 157: 1301.

TABLE 72.2 Histopathological Features of the Idiopathic Interstitial Pneumonias

	UIP	DIP/RB-ILD	AIP	NSIP
Temporal appearance	Variegated	Uniform	Uniform	Uniform
Interstitial inflammation	Scant	Scant	No	Prominent
Collagen/ fibrosis	Patchy	Diffuse (DIP) Focal (RB-ILD)	No	Diffuse
Fibroblast proliferation	Fibroblast foci prominent	No	Diffuse	Rare
BOOP	No	No	No	Focal
Honeycomb change	Yes	No	No	Rare
Intraalveolar macrophages	Focal	Diffuse (DIP) Focal (RB-ILD)	No	Patchy
Hyaline membranes	No	No	Focal	No

AIP, acute interstitial pneumonitis; BOOP, bronchiolitis obliterans organizing pneumonia; DIP, desquamative interstitial pneumonitis; NSIP, nonspecific interstitial pneumonia; RB-ILD, respiratory bronchiolitis–associated interstitial lung disease; UIP, usual interstitial pneumonitis.

Adapted from Katzenstein AL, Myers JL. Idiopathic pulmonary fibrosis: clinical relevance of pathological classification. Am J Respir Crit Care Med 1998; 157: 1301.

KEY CONCEPTS

Pathogenesis of the Idiopathic Interstitial Pneumonias (IIPs)

- Although the inciting event(s) is unknown in the different diseases, a common result is a dysregulated fibroproliferative response (similar to wound healing), which leads to excessive extracellular matrix (ECM) production and lung remodeling.
- A genetically determined inability to repair and reepithelialize the denuded basement membranes adequately may be a contributing factor and may relate to the familial occurrence of some cases of idiopathic pulmonary fibrosis (IPF).
- The presence of a chronic stimulus (autoantigen), as is seen in the pneumoconioses, may result in a persistent inflammatory and immune response and lead to a failure in the normal healing process.
- The release of transforming growth factor-β (TGF-β) following epithelial injury stimulates collagen synthesis and the prevention of apoptosis of proliferating fibroblasts in the lung and may impair collagen degradation by inhibiting the production of metalloproteases.
- A predominant T-helper type 2 (Th2) response in the lung and the absence of interferon-γ (IFN-γ) favor the development of a fibrotic response.

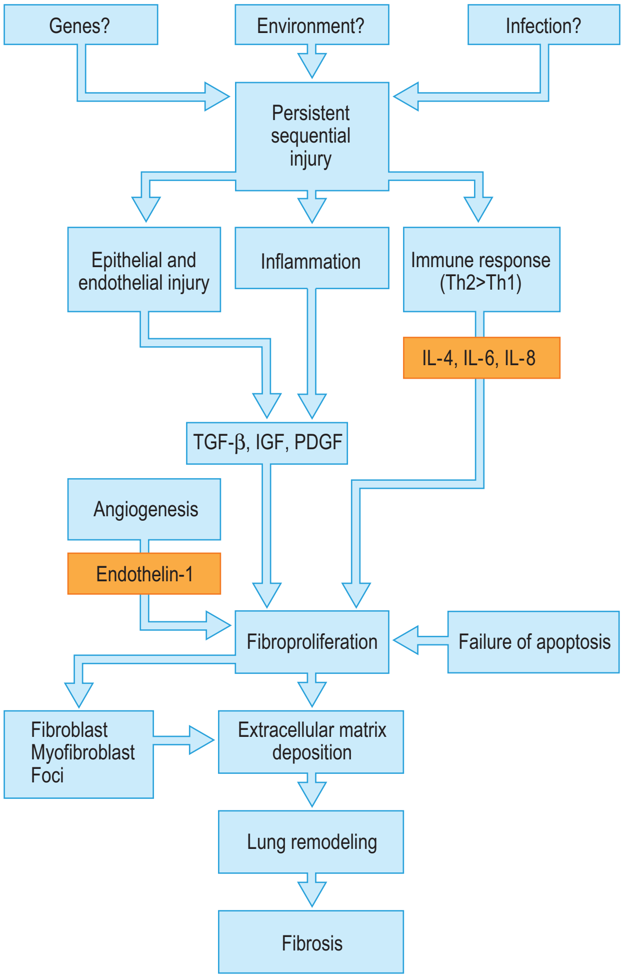


FIG 72.4 Events Hypothesized to Be Involved in the Pathogenesis of Idiopathic Pulmonary Fibrosis. The initiating event(s) leading to persistent lung injury remains poorly understood. The interaction between genetic factors, environmental exposures, and infectious agents leads to epithelial and endothelial injury, resulting in the secretion of macrophage-derived growth factors, including transforming growth factor-β (TGF-β), insulin-like growth factor-1 (IGF-1), and platelet-derived growth factor (PDGF). This cytokine milieu stimulates fibroblast proliferation and collagen deposition. In addition, the resulting T-helper type 2 (Th2) immune response stimulates extracellular matrix production and fibroblast proliferation, resulting in lung remodeling and, eventually, lung fibrosis.

The history and physical findings in IPF are nonspecific. However, extrapulmonary involvement does not occur; the presence of fever, arthralgias, myalgias, or pleuritis should suggest a connective tissue disorder. Antinuclear antibodies (ANAs) and rheumatoid factor are present in 10–20% of patients with IPF, but titers greater than 1 : 320 should suggest an alternative diagnosis.

The majority of patients with IPF have abnormal chest radiography results at the time of presentation. Basal peripheral reticular opacities are the characteristic radiographic findings. A confident diagnosis of IPF from HRCT of the lung requires the presence of patchy, peripheral bibasal reticular abnormalities with honeycombing.[23] The presence of extensive ground-glass opacities on HRCT should suggest an alternative diagnosis, such as DIP, hypersensitivity pneumonitis, bronchiolitis obliterans organizing pneumonia (BOOP), or NSIP.

A surgical lung biopsy is recommended in patients with suspected IPF without a definitive HRCT appearance and who do not have contraindications to the procedure. This is especially important in patients with atypical clinical or radiographic findings, which could suggest the possibility of one of the other histological patterns of the IIPs and an improved prognosis. Biopsy may be omitted in older patients with cardiovascular disease, or those with evidence of extensive honeycomb change. Biopsy through video-assisted thoracoscopy (VATS) is the preferred surgical technique and has been associated with less morbidity and a decreased length of hospital stay compared with open lung biopsy.

Treatment and Outcome

The usual course of IPF is relentless progression without spontaneous remission, commonly with a fatal outcome. The most common cause of death in patients with IPF is progression of the underlying disease with two-thirds of deaths caused by respiratory failure or cardiovascular complications. Other causes of death in IPF include bronchogenic carcinoma, infection, and pulmonary embolism. Recent studies in patients with biopsy-proven IPF indicate a poor prognosis (30–50% 5-year survival).[24] Previously, there was no evidence to support the use of any specific therapy in the management of IPF. However, recent clinical trials have shown decreased decline in forced vital capacity (FVC) after use of either pirfenidone[25] or nintedanib.[26] Thus for the first time, there are two drugs approved by the US Food and Drug Administration (FDA) for the treatment of IPF, with multiple other phase II and III clinical trials scheduled for completion in the near future. Finally, lung transplantation should be considered in patients who have progressive clinical and physiological deterioration and who meet established criteria.

Acute Interstitial Pneumonia

AIP is a fulminant form of IIP. Although it was previously thought to represent an acute phase of UIP, some studies have suggested that it is a distinct entity.[18] However, patients with documented UIP/IPF experiencing acute exacerbations can have the pathology of AIP superimposed on UIP.[27]

Clinical Manifestations

AIP presents with the abrupt onset of dyspnea followed by rapid progression to respiratory failure. The clinical, radiographic, physiological, and histological features are identical to those of the acute respiratory distress syndrome (ARDS) but without any identifiable cause. Most patients are previously healthy individuals over 40 years of age. Men and women are equally affected. A viral prodrome is common, with symptoms that include fever, nonproductive cough, and dyspnea. Laboratory studies are nonspecific. Chest radiography and HRCT show diffuse airspace opacities and ground-glass attenuation, respectively. A similar presentation may occur as the initial manifestation of a CTD.

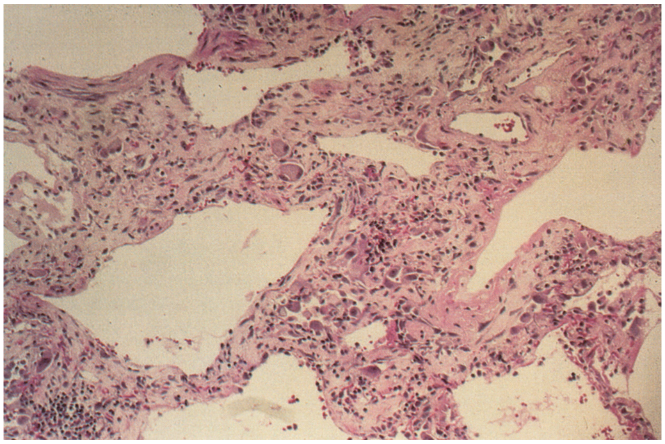


FIG 72.5 Histopathology of Acute Interstitial Pneumonitis. Diffuse thickening of the alveolar septum with an infiltration of mononuclear cells is the characteristic abnormality. The temporal uniformity of this process is also apparent.

Histopathology

AIP is characterized by diffuse interstitial fibrosis that is temporally uniform (Fig. 72.5).[28] The changes are identical to the organizing phases of diffuse alveolar damage, as seen in ARDS. Within the thickened interstitial space, there is active, diffuse fibroblast proliferation similar to the focal fibroblast foci seen in UIP. If this is progressive, honeycomb change occurs. Other features of acute lung injury, which are frequently seen in AIP, are intra-alveolar hyaline membranes.

Diagnosis

The diagnosis of AIP is based on a clinical syndrome of idiopathic ARDS and the presence of organizing diffuse alveolar damage on lung biopsy. Lung biopsy is occasionally performed to establish the diagnosis and exclude other causes of acute ILD.

Treatment and Outcome

No effective therapy exists for patients with AIP. Glucocorticoids are utilized in most cases, but no survival benefit has been shown. Overall, the prognosis of patients with AIP is poor, with mortality rates in the range of 50–88%. Half the patients die within 6 months of disease onset. However, those who survive may have complete recovery of lung function, and AIP rarely recurs in survivors.

Desquamative Interstitial Pneumonitis

DIP represents fewer than 3% of all cases of interstitial lung disease.[29] However, it is a distinct clinicopathological entity that differs substantially from UIP.

Clinical Manifestations

DIP affects individuals in their fourth to fifth decades of life with a male predominance. It predominantly occurs in cigarette

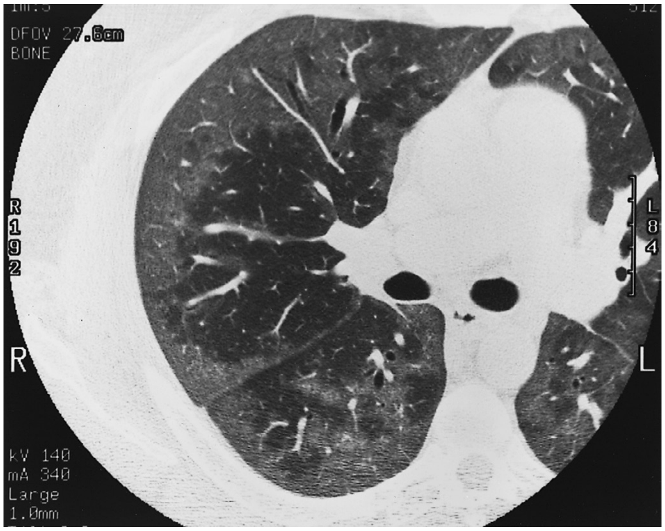


FIG 72.6 Radiographic Manifestations in Desquamative Interstitial Pneumonitis. High-resolution computed tomography scan in a patient with desquamative interstitial pneumonitis shows ground-glass attenuation in the periphery of the upper and lower lung fields.

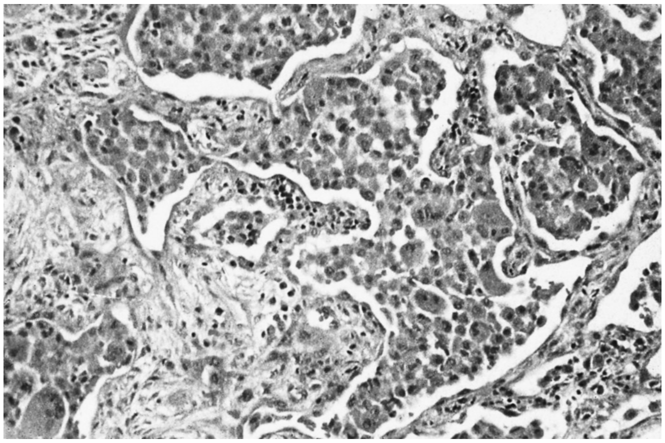


FIG 72.7 Histopathology of Desquamative Interstitial Pneumonitis. A high-magnification photomicrograph of desquamative interstitial pneumonitis shows the uniform, diffuse accumulation of macrophages within the alveolar space with associated thickening of the alveolar septum. These aggregates of macrophages almost completely fill the alveolar spaces.

smokers. Clinically, most individuals present with subacute onset of a dry, nonproductive cough and dyspnea. Clubbing is present in approximately 50% of patients with DIP. Laboratory evaluation is usually nonspecific.

Although the chest radiography results can be normal in up to 20% of symptomatic individuals, it typically shows nonspecific bibasilar ground-glass opacities. Reticulonodular interstitial infiltrates have also been reported. HRCT confirms the presence of ground-glass attenuation in the periphery of the lower lung zones (Fig. 72.6). Pulmonary function testing shows a restrictive defect with hypoxemia and a decrease in diffusion capacity.

Histopathology

DIP is a misnomer. It was initially thought that the intraalveolar cells represented sloughed or desquamated alveolar epithelial cells. However, DIP is pathologically characterized by uniform, diffuse accumulation of macrophages in the alveolar space (Fig. 72.7). At low magnification, the overall appearance is one of uniformity from one field of view to the next as opposed to the variegated appearance of UIP. In addition, there is scant interstitial inflammation with varying degrees of fibrosis of the alveolar septum.

Diagnosis

The diagnosis of DIP requires tissue confirmation of the pathological lesion. This is important since DIP has a better prognosis and response to therapeutic intervention compared with IPF. A DIP-like pattern is frequently seen in other IIPs as well as in pulmonary Langerhans cell histiocytosis, CTDs, and drug reactions. Thus the diagnosis of DIP requires careful correlation of pathological findings with clinical and radiological findings.

Treatment and Outcome

The primary intervention in DIP is smoking cessation. Since this is a rare condition with relatively few published cases, it is unclear whether glucocorticoids alter the natural history of this disease. A mortality rate of 28% with a mean survival of 12 years has been reported compared with a 30–50% 5-year survival in UIP.[24] Of note, 22% patients improved spontaneously, and 60% responded to glucocorticoid therapy. This picture is dramatically different from that of IPF, in which spontaneous improvement rarely, if ever, occurs. However, a significant minority of patients with DIP fail to respond to treatment and progress to respiratory failure secondary to advanced fibrosis.

Respiratory Bronchiolitis–Associated Interstitial Lung Disease

RB-ILD is a distinct clinical entity that occurs in current or former cigarette smokers. It is unclear whether RB-ILD and DIP represent different diseases or different ends of the spectrum of the same disease process.[29] DIP occurs predominantly and RB-ILD occurs exclusively in cigarette smokers, suggesting a common pathogenesis related to cigarette smoke exposure.

Clinical Manifestations

The mean age at presentation with RB-ILD is 36 years. Males are more often affected, and all individuals with RB-ILD are cigarette smokers. Symptoms include a dry, nonproductive cough and dyspnea. Clubbing is absent in RB-ILD, whereas it is frequently present in DIP. Laboratory evaluation is nonspecific.

Chest radiography typically shows diffuse, fine reticular or nodular interstitial opacities with normal lung volumes. Additional findings include bronchial wall thickening and a prominent peribronchovascular interstitium. HRCT may reveal ground-glass opacification and emphysema.

Pulmonary function tests most commonly reveal a mixed restrictive–obstructive pattern with a reduced diffusing capacity and mild hypoxemia. The residual volume may be increased, with no change in other spirometric parameters.

Histopathology

The pathology of RB-ILD is similar to that of DIP. However, in RB-ILD, the intraalveolar macrophages accumulate primarily

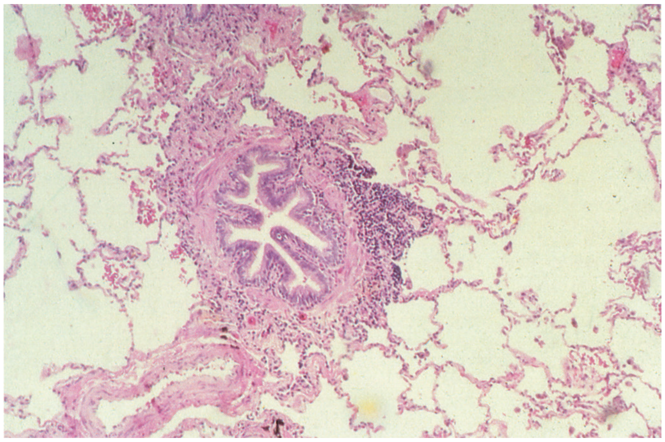


FIG 72.8 Histopathology of Respiratory Bronchiolitis–Interstitial Lung Disease. An ectatic bronchiole with a thickened wall is shown, with a mononuclear infiltrate extending into the immediately surrounding alveoli.

within the peribronchiolar airspaces and are associated with thickening of the alveolar septum in these areas (Fig. 72.8). The differentiation of this lesion from DIP requires sparing of distal airspaces with the lesion confined to the peribronchiolar airspaces in RB-ILD.

Diagnosis

RB-ILD should be suspected in young individuals who have a history of cigarette smoking and complain of cough and dyspnea with chest radiography or HRCT showing nodular and/or reticular interstitial opacities. The diagnosis requires tissue confirmation of the pathological findings noted above.

Treatment and Outcome

The key therapeutic intervention in RB-ILD is cessation of smoking. The use of glucocorticoids has been associated with favorable results. At present, the clinical course and prognosis of patients with RB-ILD are unknown. In most clinical series, patients either improved or stabilized, and mortality is uncommon.[29,30]

Nonspecific Interstitial Pneumonitis

The term NSIP was first used to describe cases of interstitial pneumonia that did not demonstrate a pattern of UIP, AIP, or DIP. Currently, the term NSIP is applied to an IIP or to a similar histological pattern that occurs in CTD, hypersensitivity pneumonitis, infection, or drug-induced lung disease. Thus the diagnosis of NSIP should prompt investigation for a causative agent.[18] In fact, 16% of patients in the original description of NSIP had one of the CTDs.[31]

Clinical Manifestations

Idiopathic NSIP occurs in middle-aged individuals, with a slight female predominance. A dry, nonproductive cough and exertional dyspnea are the most common symptoms, although fever is present in 25% of patients. Symptoms are usually present for 6–10 months before diagnosis. As in other IIPs, the laboratory evaluation is nonspecific.

Chest radiography usually shows bilateral interstitial infiltrates, and sometimes the result can be normal in a symptomatic patient. HRCT characteristically shows bilateral, patchy ground-glass attenuation indistinguishable from DIP or RB-ILD.[23]

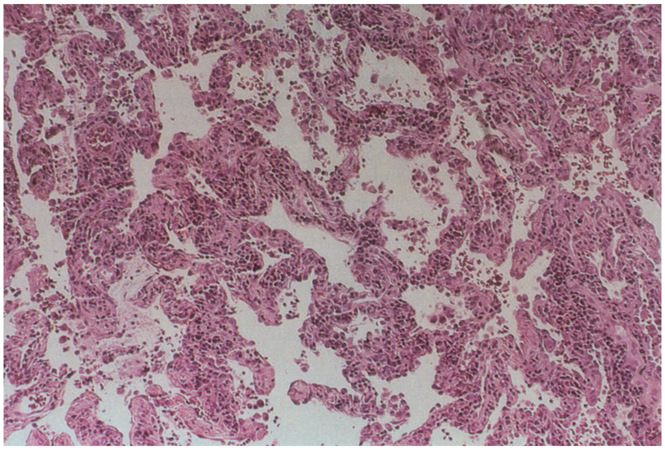


FIG 72.9 Histopathology of Nonspecific Interstitial Pneumonitis. Low-magnification photomicrograph of cellular nonspecific interstitial pneumonitis shows diffuse uniform thickening of the alveolar septum as a result of the presence of a lymphoplasmacytic infiltrate.

Histopathology

NSIP is characterized by varying, but temporally uniform, degrees of fibrosis and inflammation of the alveolar septum, without the histopathological features indicative of UIP, AIP, or DIP (Fig. 72.9). NSIP has been divided into three groups, depending on the presence or absence of interstitial fibrosis: interstitial lymphoplasmacytic inflammation (48% of cases); inflammation and fibrosis (38%); and fibrosis (14%). Although the changes are temporally uniform, they may be patchy with intervening areas of normal lung.

This temporal uniformity is in contrast to the variegated pattern seen in UIP. Fibroblast foci, the earliest lesions seen in UIP, are found in 20% of patients with NSIP, making it difficult to differente fibrotic NSIP from UIP. The key feature in this circumstance is the temporal uniformity of the lesions in NSIP.

Treatment and Outcome

Unlike patients with UIP, individuals with NSIP have a favorable prognosis. In the original description of the disease, 45% of subjects completely recovered, and the condition of another 42% remained stable or improved.[31] Only 11% of patients died, with a mean survival of 16 months. All of the individuals with an aggressive course were in the fibrotic group. Ten-year survival in the cellular group was 90%, compared with 35% in patients with the fibrotic pattern. Despite the worse prognosis of NSIP with a fibrosing pattern, this is still significantly better than the 15% 10-year survival rate for patients with UIP.[32]

Cryptogenic Organizing Pneumonia

COP is a specific clinicopathologic disorder of unknown etiology characterized by excessive proliferation of granulation tissue within the lumen of distal airspaces.[33] The term COP is reserved for cases demonstrating bronchiolitis obliterans organizing pneumonia (BOOP) without an obvious cause, since this histological appearance occurs in a variety of inflammatory lung disorders, including CTDs, malignancy, infections, and those caused by medications.

Clinical Manifestations

The onset of disease is usually in the fifth to sixth decades of life; men and women are equally affected. Most individuals have symptoms for less than 2 months before diagnosis. The initial presentation is usually with a dry, nonproductive cough and flu-like symptoms, including fever, sore throat, and malaise. This is followed by progressive dyspnea, and routine laboratory evaluation is nonspecific.

Chest radiography shows diffuse, often patchy alveolar opacities in the setting of normal lung volumes (Fig. 72.10A). These opacities can be migratory and usually have a peripheral distribution similar to those seen in chronic eosinophilic pneumonia. Rarer radiographic manifestations include linear or nodular interstitial opacities and honeycombing. The presence of a pleural effusion or pleural thickening should suggest an associated CTD. HRCT shows patchy airspace consolidation, especially in the lung periphery with a lower-lung zone predominance (see Fig. 72.10B). Other findings include ground-glass attenuation, small nodular opacities, and bronchial wall thickening.

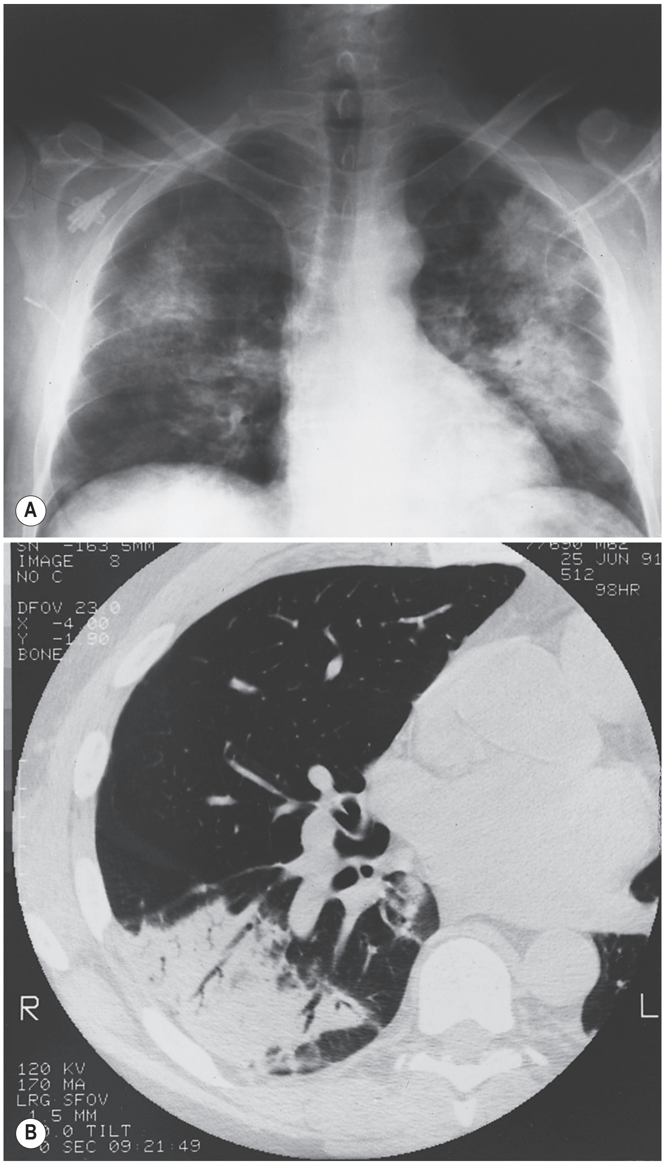


FIG 72.10 **Radiographic Findings in Cryptogenic Organizing Pneumonia.** (A) Chest radiograph in a patient with cryptogenic organizing pneumonia shows bilateral patchy alveolar opacities with a peripheral distribution in the setting of normal lung volumes. (B) Chest computed tomography shows a dense right lower lung consolidation with the presence of air bronchograms.

As in other ILDs, a restrictive ventilatory defect is the most common pulmonary function abnormality. Gas exchange abnormalities are common and are accompanied by decreased diffusing capacity, widening of the alveolar–arterial gradient, and exercise-induced hypoxemia.

Histopathology

The histopathology of COP is characterized by excessive proliferation of granulation tissue in the small airways and alveolar ducts with associated chronic inflammation in the alveolar walls (Fig. 72.11).[33] The intraluminal fibrotic buds (Masson bodies) consist of loose collagen-embedding fibroblasts and myofibroblasts and have a tendency to extend from one alveolus to the next, giving a characteristic "butterfly" pattern. The lesions are patchy in nature and have a uniform temporal appearance at low magnification, with preservation of the underlying lung parenchyma. This pattern has been described as the prototypical healing response of the lung to a variety of insults.

Diagnosis

The presence of BOOP in a lung biopsy does not necessarily represent COP, since COP is a diagnosis of exclusion. Organizing pneumonia is a nonspecific response to many lung injuries and can occur in conjunction with another pathological process or as a component of other primary pulmonary disorders, such as infections, irradiation, CTD, hypersensitivity pneumonitis, granulomatosis with polyangiitis, or chronic eosinophilic pneumonia (Table 72.3).

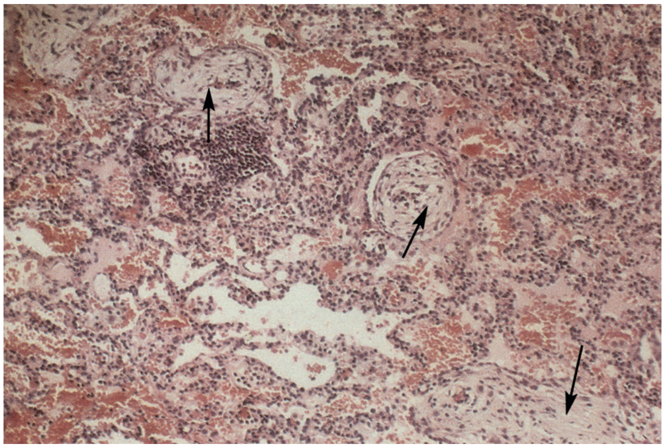


FIG 72.11 **Histopathology of Cryptogenic Organizing Pneumonia.** A photomicrograph of cryptogenic organizing pneumonia shows intraalveolar fibroblast proliferation (arrows) and early collagen production. In addition, thickening of the alveolar septa with a lymphoplasmacytic infiltrate consistent with cellular nonspecific interstitial pneumonitis is present.

TABLE 72.3 Disorders Associated With a Bronchiolitis Obliterans Organizing Pneumonia (BOOP) Pattern

Idiopathic BOOP (cryptogenic organizing pneumonia)
Connective tissue diseases
- Systemic lupus erythematosus
- Rheumatoid arthritis
- Polymyositis/dermatomyositis
- Sjögren syndrome

Hypersensitivity pneumonitis
Chronic eosinophilic pneumonia
Drug-induced
- Gold
- Penicillamine
- Amiodarone
- Bleomycin
- Sulfa drugs

Granulomatosis with polyangiitis (wegener granulomatosis)
Bone marrow transplantation
Lung transplantation/rejection
Inhalational injury
Neoplasms
Lung irradiation
Virus-associated
- Human immunodeficiency virus (HIV)
- Influenza virus
- Adenovirus

TABLE 72.4 Pleuropulmonary Manifestations of Connective Tissue Diseases

	SLE	RA	SSc
Pulmonary hypertension	+	+	+++
Vasculitis	+	±	±
Pleural disease	+++	+++	+
Bronchiolitis obliterans	±	++	+
Aspiration pneumonia	–	–	++
Diaphragmatic dysfunction	++	–	–
Lung nodules	–	++	–
Diffuse alveolar damage	+	±	±
BOOP	±	+	±
UIP	+	++	+
Capillaritis	++	+	±
LIP	+	+	+
NSIP	+	++	++

BOOP, bronchiolitis obliterans organizing pneumonia; LIP, lymphocytic interstitial pneumonia; NSIP, nonspecific interstitial pneumonia; RA, rheumatoid arthritis; SLE, systemic lupus erythematosus; SSc, systemic sclerosis; UIP, usual interstitial pneumonitis.

Treatment and Outcome

Treatment with glucocorticoids usually offers dramatic clinical and radiographic improvement within days to weeks.[33] Complete clinical, physiological, and radiographic recovery occurs in two-thirds of cases. In the remainder, persistent disease progresses to fibrosis. It is common for relapses to occur with glucocorticoid tapering, followed by improvement with reintroduction of treatment; consequently at least 6 months of therapy is recommended. The 5-year survival rate in COP is 73%, compared with 5-year survival rates of 44% in patients with BOOP resulting from other causes (*e.g.,* CTD) or 30% for IPF.

LUNG INVOLVEMENT IN CONNECTIVE TISSUE DISEASES

CTDs are a heterogeneous group of systemic autoimmune diseases that frequently involve the lungs. The pleuropulmonary manifestations of these diseases are diverse, affecting all parts of the respiratory tract (*i.e.,* airways, alveoli, blood vessels, and pleura) (Table 72.4). Although pulmonary complications generally occur in patients with well-established disease, occasionally the lung involvement is the first manifestation of the autoimmune disorder. This section discusses the pleuropulmonary manifestations of systemic lupus erythematosus (SLE), RA, and systemic sclerosis (SSc) (for a discussion of other manifestations in these diseases, see Chapters 51, 52, and 55).

Systemic Lupus Erythematosus

SLE is a disease of unknown etiology characterized by the presence of autoantibodies directed against various nuclear antigens. These autoantibodies and the resultant immune complexes mediate many of the manifestations of SLE (Chapter 51). This disease mainly affects young women (female-to-male ratio >8 : 1) and may involve virtually every organ system. Pleuropulmonary involvement occurs at some point in the disease course in 38–89% of cases.[34] Thus the respiratory system is affected more commonly in SLE than in any other CTD. However, infectious pneumonia remains the commonest cause of pulmonary disease and death in these patients. Thus in patients with SLE presenting with a febrile illness and pulmonary infiltrates, a community-acquired or opportunistic infection must be promptly excluded.

Acute Lupus Pneumonitis

Acute lupus pneumonitis is an uncommon pulmonary manifestation of SLE, occurring in fewer than 5% of cases.[34] The clinical presentation mimics that of an infectious pneumonia with the abrupt onset of fever, cough, and dyspnea. Serum complement levels are often low, and chest radiography typically shows diffuse alveolar opacities. Acute lupus pneumonitis can be accompanied by pericarditis and often pleuritis and pleural effusion.

It can be difficult to distinguish acute lupus pneumonitis from an infectious pneumonia. BAL followed by thoracoscopic lung biopsy is recommended before instituting corticosteroid therapy. The histopathology varies and includes diffuse alveolar damage, BOOP, NSIP, or a combination of these.

There are no controlled trials of therapy for acute lupus pneumonitis. Treatment includes high-dose glucocorticoids (1–2 mg/kg/day) with or without accompanying cytotoxic drugs, such as cyclophosphamide. Mortality rates as high as 50% have been reported. In those patients who fail to respond to treatment, respiratory failure is the usual cause of death.

Diffuse Alveolar Hemorrhage

DAH occurs in fewer than 5% of patients with SLE, and it represents the initial manifestation of disease in 11–20% of those cases.[34] However, most cases develop in individuals with well-established diagnoses of SLE, usually with preexisting lupus nephritis.

The symptoms of DAH mimic those of infectious pneumonia and acute lupus pneumonitis.[34] Hemoptysis is present in 42–66%

of patients at presentation. Therefore the absence of hemoptysis does not exclude the diagnosis, particularly in the setting of a falling hematocrit, diffuse pulmonary infiltrates, and blood-stained BAL fluid. DAH in SLE most often results from pulmonary capillaritis, but it can also be caused by diffuse alveolar damage. Immunofluorescence studies show granular deposits of immunoglobulin G (IgG) and C3 along alveolar walls, interstitium, and capillary endothelial cells.

There are no controlled trials for the treatment of alveolar hemorrhage in SLE. Glucocorticoids, cytotoxic drugs, and plasmapheresis have been used in various combinations. The mortality rate associated with DAH is approximately 50%. Poor prognostic factors include the need for mechanical ventilation, presence of infection, and prior treatment with cyclophosphamide.

Lupus Pleuritis

The pleura are the most common site of respiratory involvement in SLE, with pleurisy and pleural effusions occurring in 50–80% of patients. Lupus pleuritis can be the presenting manifestation of disease, but more commonly, it develops in patients with established SLE. It is often recurrent. The clinical manifestations include chest pain, fever, and dyspnea, and chest radiography typically shows bilateral pleural effusions. The pleural fluid is serous or serosanguineous and exudative in nature. Compared with effusions in RA, the glucose is higher, and the lactate dehydrogenase level is lower. The most helpful measurement is a pleural fluid ANA titer greater than 1:160. Examination of the pleura reveals infiltration with plasma cells and lymphocytes, accompanied by pleural thickening and fibrosis. Treatment with nonsteroidal antiinflammatory drugs (NSAIDs) and/or glucocorticoids is usually effective for relief of pleural discomfort.

Interstitial Lung Disease

The presence of ILD in SLE is uncommon, especially compared with SSc or RA. However, minor interstitial abnormalities can be found on HRCT in approximately one-third of patients with SLE who have normal results of chest radiography and physiological testing. The significance and natural history of these subclinical findings are uncertain. The presence of anti-SSA (Ro) has been noted in approximately 80% of patients who have lupus with interstitial changes. In addition, the prevalence of ILD is increased in a subset of patients with SLE who have sclerodermatous skin changes.

The diagnosis of SLE is usually well established in patients who develop the insidious form of ILD. The disease course is characterized by progressive dyspnea and cough; chest radiography shows reduced lung volumes and reticular interstitial infiltrates. A restrictive lung function pattern with reduced diffusing capacity and exercise-induced hypoxemia are typical. The histopathology of chronic interstitial disease in SLE resembles NSIP, although cases of BOOP, LIP, and UIP have been described. Response to therapy depends on the underlying histopathology, with the UIP-like form being least responsive.

Pulmonary Vascular Disease

Although previously thought to be unusual, the development of pulmonary hypertension has been increasingly noted in SLE, with an incidence ranging from 0.5–14%. Pulmonary hypertension in SLE has been associated with the presence of Raynaud syndrome, serositis, digital vasculitis, and antiphospholipid antibodies.[35] Dyspnea and fatigue, despite normal results on chest radiography, is the most common presentation. Most of the patients with SLE who have pulmonary hypertension are female, with 3- and 5-year survival rates of 45% and 17%, respectively, representing a worse prognosis compared with that for patients with idiopathic pulmonary hypertension. The vascular changes of SLE-associated pulmonary hypertension are similar to those seen in idiopathic pulmonary hypertension with intimal hyperplasia, smooth muscle hypertrophy, and medial thickening. Several pathological mechanisms have been proposed for the development of pulmonary hypertension, including vasoconstriction, in addition to vasculitis and thrombosis associated with antiphospholipid and anticardiolipin antibodies. Serum endothelin levels are elevated in patients with SLE-associated pulmonary hypertension and correlate with pulmonary arterial pressures.

As the pulmonary hypertension advances, the central pulmonary arteries enlarge. Pulmonary function testing shows an isolated decrease in the diffusing capacity for carbon monoxide. Patients with SLE-associated pulmonary hypertension may respond to immunosuppressive therapy. In a small study, five of 12 patients with SLE responded to monthly intravenous bolus doses of cyclophosphamide in addition to systemic glucocorticoids. A positive response was indicated by sustained hemodynamic improvement after at least 1 year of treatment without the need for additional pulmonary hypertension–specific therapies. Patients who responded to immunosuppression could be maintained on azathioprine or mycophenolate mofetil to avoid potential adverse effects of cyclophosphamide. Patients with SLE who were treated with bosentan did not have clinical worsening and showed an improvement in 6-minute walk distance.[36] Overall, the long-term survival of patients with SLE-associated pulmonary hypertension is poor,[37] and the optimal treatment regimen for SLE-associated pulmonary hypertension remains unknown.

Respiratory Muscle Dysfunction

The shrinking lung syndrome is caused by diaphragmatic weakness as well as weakness of other respiratory muscles. This entity accounts for the findings of dyspnea without evidence of interstitial infiltrates or pulmonary vascular disease. It occurs in 25% of patients with SLE. Chest radiography typically shows elevated diaphragms and basilar atelectasis. The pathogenesis of respiratory muscle weakness is unknown, but it is not associated with generalized muscle weakness. Glucocorticoids are frequently ineffective in the treatment of this syndrome. Improvement has been noted with inhaled β-agonist and theophylline therapy. Despite a variable response to therapy, it is unusual for this manifestation of SLE to be progressive.

Rheumatoid Arthritis

RA is an autoimmune disease associated with autoantibodies directed against citrullinated antigens and characterized by the presence of a symmetrical, inflammatory polyarthritis (Chapter 52). It occurs more frequently in women, with a female-to-male ratio of 2:1. Disease onset is most commonly in the fourth to fifth decades of life. The pleuropulmonary complications of RA occur more commonly in individuals with subcutaneous nodules, high titers of rheumatoid factor, and more severe chronic articular involvement. Although RA itself is more common in women, the pleuropulmonary manifestations have a higher incidence in men. The pleuropulmonary complications of RA are numerous, but the treatment-related lung toxicity and pulmonary infections are difficult to differentiate from the primary pleuropulmonary manifestations of the disease.

Pleuritis and Pleural Effusions

As in SLE, pleural abnormalities are one of the most common pulmonary complications of RA. Pleural effusion is clinically evident in approximately 5% of patients, and this can occur before the development of arthritis. Pleural disease is often discovered as an incidental finding on routine chest radiography, but nonspecific chest pain, dyspnea, and fever are not unusual. The effusion can be unilateral or bilateral and can coexist with ILD.

Typically, the effusion is an exudate, with a glucose level less than 30 mg/mL in 70–80% of cases. The mechanism underlying the low pleural fluid glucose is impaired membrane transport of glucose. A low pleural fluid pH is thought to occur secondary to impaired carbon dioxide exit from the pleural space. If the effusion is chronic, the cholesterol concentration can be increased, and the pleural fluid can have a milky appearance (pseudochylothorax). Cytological examination reveals multinucleated giant cells, spindle-shaped macrophages, and necrotic debris.

Most rheumatoid effusions are small and asymptomatic, thus requiring no treatment. They resolve over several months without complications. The use of glucocorticoids for active articular disease hastens the resolution of the pleural process.

CLINICAL PEARLS

Lung Involvement in Rheumatoid Arthritis (RA)

- RA is more common in women, but pleuropulmonary complications occur more frequently in men.
- Factors associated with pleuropulmonary complications of RA include more severe articular involvement, subcutaneous nodules, and high levels of rheumatoid factor.
- Pleural effusions are the most common complication, characterized by an exudate and a low glucose and low pH.
- The differentiation of rheumatoid nodules from malignant lesions can be difficult.
- The rapid growth of a nodule should prompt aggressive investigation for a malignant cause.

Rheumatoid Nodules

Rheumatoid or necrobiotic nodules are the only pleuropulmonary manifestation specific for RA. These nodules are most commonly seen in men with active articular disease, high rheumatoid factor titers, and subcutaneous nodules. Most individuals are asymptomatic and are diagnosed on routine chest radiography. Radiographically, these nodules can be singular or multiple with an upper to midlung zone predominance. Cavitation occurs in approximately 50% of cases. HRCT indicates a higher frequency of nodules than previously thought. Rarely, subpleural necrobiotic nodules can erode into the pleural space, resulting in a pneumothorax with a complicating bronchopleural fistula. It can be difficult to differentiate these nodules from malignant lesions, and open-lung biopsy is frequently necessary. Evidence of rapid growth on chest radiography should prompt an aggressive diagnostic evaluation.

Caplan syndrome refers to the rapid development of pulmonary nodules predominantly in the upper lung zone. This syndrome was originally described in Welsh coalminers with RA. Histologically, these nodules are identical to necrobiotic nodules. Other occupational dust exposures have also been associated with this syndrome.

Airway Disease

Airflow limitation is a common finding in patients with RA, being present in approximately one-third of patients. The mechanism(s) responsible for airway disease is poorly understood. The interplay of cigarette smoking and RA may play a role.

A life-threatening complication of RA is upper airway obstruction, resulting from synovitis of the cricoarytenoid joint. Common presenting complaints include a sore throat, hoarseness, and fullness in the throat. It can progress to inspiratory stridor and upper airway obstruction. This complication occurs more commonly in women, particularly in those with advanced RA. Seventy-five percent of patients were found to have cricoarytenoid abnormalities when screening with direct or indirect laryngoscopy and computed tomography (CT) was utilized. The treatment of cricoarytenoid arthritis includes antiinflammatory medications.

Bronchiolitis obliterans is a progressive form of obstructive lung disease that is being increasingly recognized as a complication of RA.[38] This entity was thought to develop secondary to the use of penicillamine in the treatment of RA, but most cases occur in the absence of this therapy. The histopathological lesion of bronchiolitis obliterans is constrictive bronchiolitis, which is characterized by concentric submucosal and peribronchiolar fibrosis resulting in extrinsic compression and obliteration of the bronchiolar lumen. The typical clinical presentation is insidious onset of cough and dyspnea, with a normal or hyperinflated chest on radiography. This complication occurs more commonly in women than in men. Pulmonary function studies show airflow limitation with hyperinflation and a reduced diffusing capacity. Expiratory HRCT shows multiple areas of air trapping (mosaic pattern). Some individuals respond to high-dose glucocorticoids and cytotoxic drugs, but in most patients, bronchiolitis obliterans progresses to respiratory failure.

Bronchiectasis occurs at an increased frequency in RA, usually in individuals with long-standing articular disease. Productive cough and dyspnea are the most common respiratory symptoms. In most patients, bronchiectasis is not clinically significant. Recurrent pneumonia and respiratory failure are potentially fatal complications of this problem.

Interstitial Lung Disease

Although ILD is a common complication of RA, the incidence is difficult to determine, since different methods of detection have been employed and dissimilar populations of patients have been studied. However, clinically significant ILD occurs in approximately 14% of patients.[39] The development of ILD in relation to the onset of arthritis is variable. Most often, the ILD develops subsequent to arthritis, but in approximately 20% of patients, the lung disease precedes the onset of arthritis and is associated with cigarette smoking, presence of the shared HLA-DR4 epitope, and RA-specific anticitrullinated protein antibodies.

The most common histopathologies identified in this patient population are UIP, LIP, NSIP and BOOP. The clinical manifestations of ILD in RA resemble those seen in idiopathic disease and include a dry, nonproductive cough and dyspnea on exertion. Chest radiography and HRCT show increased reticular markings with a predilection for the peripheral lower lung zones. Often, pleural abnormalities accompany the interstitial changes. With advanced disease, progression to honeycomb lung occurs. LIP usually occurs in cases of RA complicated by Sjögren syndrome;

the presence of keratoconjunctivitis sicca and xerostomia in a patient with RA and ILD should suggest this histological type.

In general, ILD in RA appears more indolent than in the IIPs. Thus because of uncertain treatment benefits and possible adverse effects, the decision to institute therapy should be based on clinical, radiographic, and physiological deterioration.

Drug-Induced Lung Disease

Methotrexate and gold are the two main anti-RA drugs capable of causing lung injury. Methotrexate administered weekly (10–20 mg/week) is associated with the development of interstitial changes in 1–5% of patients with RA. No correlation with age, sex, disease duration, or cumulative dose has been identified. The clinical presentation is subacute, with fever, cough, and dyspnea occurring 1–5 months after initiation of the drug. Chest radiography shows mixed interstitial–alveolar infiltrates. Nonspecific laboratory abnormalities include leukocytosis, sometimes with mild eosinophilia, and an elevated erythrocyte sedimentation rate (ESR). In most cases, BAL reveals a lymphocytosis. Histologically, cellular NSIP is seen with areas of organizing pneumonia. Noncaseating, granulomatous inflammation similar to that seen in hypersensitivity pneumonitis may also be present. The primary treatment of methotrexate-induced pneumonitis is withdrawal of methotrexate, as well as appropriate supportive care.

Gold-induced pneumonitis occurs in fewer than 1% of patients with RA who are treated with gold. Dyspnea and cough usually begin after 4–6 weeks of therapy; eosinophilia occurs in a minority of patients. Chest radiography typically reveals mixed alveolar–interstitial opacities with a predilection for the upper lung zone. The histology is similar to that seen in patients with RA-associated ILD. Thus the differentiation of gold-induced pneumonitis from RA-associated ILD can only be established when discontinuation of the medication results in remission.

Systemic Sclerosis (Scleroderma)

SSc is characterized by excessive deposition of ECM in the skin and internal organs, and vascular involvement (Chapter 55). The degree of visceral organ involvement determines morbidity and mortality. Pulmonary involvement occurs in 70–100% of patients with SSc. There is no correlation with the degree of extrapulmonary disease. ILD is the most common pulmonary manifestation of SSc. Of note, with the improved mortality associated with renal involvement in SSc, lung disease has become the most important cause of morbidity and mortality.

Interstitial Lung Disease

The incidence of ILD in SSc depends on the method of detection. Autopsy studies have reported an ILD incidence of 60–100% of cases, whereas studies based on chest radiography have noted interstitial changes in 14–66% of cases. Cough and dyspnea on exertion are the most common symptoms. Physical examination reveals bibasilar rales. Radiographic findings include basal reticulonodular infiltrates, enlargement of pulmonary arteries, and progressive volume loss. Pulmonary function testing reveals restrictive lung disease, preservation of flow rates, and decreased diffusing capacity. A disproportionate decrease in diffusing capacity compared with lung volume changes should suggest pulmonary hypertension, especially in individuals with limited scleroderma (calcinosis, Raynaud phenomenon [RP], esophageal dysmotility, sclerodactyly, telangiectasia [CREST] syndrome). The predominant histopathologic abnormality is NSIP. Rarely, LIP may complicate cases of SSc associated with Sjögren syndrome. Although the 5-year survival for patients with SSc and ILD is 38–45%, it is better than that of patients with IPF.

Pulmonary Vascular Disease

Pulmonary hypertension is a frequent complication of SSc, occurring in approximately 30% of patients with diffuse scleroderma and in 10–50% of those with limited scleroderma (Chapter 55). It is a major cause of morbidity and mortality in systemic sclerosis and has also become part of the diagnostic criteria for the disease.[40] Pulmonary hypertension can either be associated with interstitial fibrosis or result from involvement of small and medium-sized arteries and arterioles with smooth-muscle hyperplasia, medial hypertrophy, and intimal proliferation (plexogenic). Direct involvement of the pulmonary circulation is more common with limited scleroderma, whereas pulmonary hypertension in patients with diffuse scleroderma is more likely associated with ILD.

The clinical presentation is characterized by the insidious onset of fatigue and dyspnea on exertion. Physical examination and chest radiography show signs typical of pulmonary hypertension, whereas a decreased diffusing capacity is seen on pulmonary function testing. Risk factors for developing SSc-associated pulmonary hypertension include limited skin involvement, duration of disease greater than 10 years, onset of SSc at older age, and severity and duration of RP.

The pathogenesis of SSc-associated pulmonary hypertension is poorly understood. Vascular changes occur at an early stage in SSc and include apoptosis, endothelial cell activation with increased expression of cell adhesion molecules, inflammatory cell recruitment, intimal proliferation, and adventitial fibrosis leading to vessel obliteration. Endothelial injury is reflected by increased levels of soluble cell adhesion molecules, disturbances of angiogenesis with increased levels of circulating vascular endothelial growth factor (VEGF), and the presence of angiostatic factors. To what extent dysregulated angiogenesis in SSc-associated pulmonary hypertension is driven by an inflammatory process or other as yet unidentified mechanisms remains unclear.

Treatment of SSc-associated pulmonary hypertension has been disappointing, with no therapy showing a significant survival benefit. Calcium channel blockers are not usually indicated for patients with SSc-associated pulmonary hypertension, although often used at lower doses for RP. Continuous intravenous epoprostenol improves exercise capacity and hemodynamics.[41] Randomized clinical trials with phosphodiesterase inhibitors, including sildenafil, showed a modest effect on exercise capacity, hemodynamic parameters, and functional class after 12 weeks of treatment. Carefully selected patients may be considered for heart–lung transplantation but are often excluded because of the risk of postoperative complications arising from SSc-related gastroesophageal reflux disease (GERD) and renal dysfunction.

CONCLUSIONS

ILDs comprise a diverse group of disorders ranging from idiopathic etiologies to those related to an underlying autoimmune condition. There is likely a complex interplay between the innate and adaptive arms of the immune system and the profibrotic pathways that lead to the development of these various disease states. Future work should focus on better understanding this relationship and, more importantly, translating these findings to the clinical setting. The role of standard immunosuppressive approaches with the ongoing discovery of novel approaches to

ON THE HORIZON

- Our understanding of the idiopathic interstitial pneumonias (IIPs) and autoimmune-related interstitial lung diseases (ILDs) has evolved over time. With increased emphasis being placed on diagnosis, because of the implications for treatment and prognosis, we are more accurately phenotyping the diseases. This has led to a better understanding of genetic associations, biological pathways, as well as more accurate identification of risk factors for disease development and progression.
- However, there are still many areas in need of further understanding and research. In the future, determining whether risk factor identification and modification can lead to primary and/or secondary prevention strategies in the treatment of immunological lung diseases will be essential.
- Testing if the novel antifibrotic therapies approved for idiopathic pulmonary fibrosis (IPF) are more broadly applicable to other forms of ILDs, including autoimmune-related ILD, will be important.
- Other therapies that target the immune system also need to be more carefully studied in diseases outside of IPF.

targeting the immune system as well as novel antifibrotic therapies will need to be determined.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Luzina IG, Todd NW, Iacono AT, et al. Roles of T lymphocytes in pulmonary fibrosis. J Leukoc Biol 2008;83:237–44.
2. Ginhoux F, et al. Fate mapping analysis reveals that adult microglia derive from primitive macrophages. Science 2010;330:841–5.
3. Ginhoux F, Jung S. Monocytes and macrophages: developmental pathways and tissue homeostasis. Nat Rev Immunol 2014;14:392–404.
4. Hashimoto D, et al. Tissue-resident macrophages self-maintain locally throughout adult life with minimal contribution from circulating monocytes. Immunity 2013;38:792–804.
5. Yona S, et al. Fate mapping reveals origins and dynamics of monocytes and tissue macrophages under homeostasis. Immunity 2013;38:79–91.
6. Schulz C, et al. A lineage of myeloid cells independent of Myb and hematopoietic stem cells. Science 2012;336:86–90.
7. Parra ER, Kairalla RA, Ribeiro de Carvalho CR, et al. Inflammatory cell phenotyping of the pulmonary interstitium in idiopathic interstitial pneumonia. Respiration 2007;74:159–69.
8. Wells AU, et al. Fibrosing alveolitis in systemic sclerosis: increase in memory T-cells in lung interstitium. Eur Respir J 1995;8:266–71.
9. Simonian PL, et al. Gammadelta T cells protect against lung fibrosis via IL-22. J Exp Med 2010;207:2239–53.
10. Wynn TA. Integrating mechanisms of pulmonary fibrosis. J Exp Med 2011;208:1339–50.
11. Sonnenberg GF, et al. Pathological versus protective functions of IL-22 in airway inflammation are regulated by IL-17A. J Exp Med 2010;207:1293–305.
12. Wilson MS, et al. Bleomycin and IL-1beta-mediated pulmonary fibrosis is IL-17A dependent. J Exp Med 2010;207:535–52.
13. Kotsianidis I, et al. Global impairment of CD4+CD25+FOXP3+ regulatory T cells in idiopathic pulmonary fibrosis. Am J Respir Crit Care Med 2009;179:1121–30.
14. Travis WD, et al. An official American Thoracic Society/European Respiratory Society statement: update of the international multidisciplinary classification of the idiopathic interstitial pneumonias. Am J Respir Crit Care Med 2013;188:733–48.
15. Fischer A, et al. An official European Respiratory Society/American Thoracic Society research statement: interstitial pneumonia with autoimmune features. Eur Respir J 2015;46:976–87.
16. Ahluwalia N, Shea BS, Tager AM. New therapeutic targets in idiopathic pulmonary fibrosis. Aiming to rein in runaway wound-healing responses. Am J Respir Crit Care Med 2014;190:867–78.
17. Coultas DB, Zumwalt RE, Black WC, et al. The epidemiology of interstitial lung diseases. Am J Respir Crit Care Med 1994;150:967–72.
18. American Thoracic S, European Respiratory S. American Thoracic Society/European Respiratory Society International Multidisciplinary Consensus Classification of the Idiopathic Interstitial Pneumonias. This joint statement of the American Thoracic Society (ATS), and the European Respiratory Society (ERS) was adopted by the ATS board of directors, June 2001 and by the ERS Executive Committee, June 2001. Am J Respir Crit Care Med 2002;165:277–304.
19. Seibold MA, et al. A common MUC5B promoter polymorphism and pulmonary fibrosis. N Engl J Med 2011;364:1503–12.
20. Feghali-Bostwick CA, et al. Cellular and humoral autoreactivity in idiopathic pulmonary fibrosis. J Immunol 2007;179:2592–9.
21. Dreisin RB, Schwarz MI, Theofilopoulos AN, et al. Circulating immune complexes in the idiopathic interstitial pneumonias. N Engl J Med 1978;298:353–7.
22. Lee CG, et al. Interleukin-13 induces tissue fibrosis by selectively stimulating and activating transforming growth factor beta(1). J Exp Med 2001;194:809–21.
23. Lynch DA, et al. High-resolution computed tomography in idiopathic pulmonary fibrosis: diagnosis and prognosis. Am J Respir Crit Care Med 2005;172:488–93.
24. Ley B, Collard HR, King TE Jr. Clinical course and prediction of survival in idiopathic pulmonary fibrosis. Am J Respir Crit Care Med 2011;183:431–40.
25. King TE Jr, et al. A phase 3 trial of pirfenidone in patients with idiopathic pulmonary fibrosis. N Engl J Med 2014;370:2083–92.
26. Richeldi L, et al. Efficacy and safety of nintedanib in idiopathic pulmonary fibrosis. N Engl J Med 2014;370:2071–82.
27. Hyzy R, Huang S, Myers J, et al. Acute exacerbation of idiopathic pulmonary fibrosis. Chest 2007;132:1652–8.
28. Katzenstein AL, Myers JL, Mazur MT. Acute interstitial pneumonia. A clinicopathologic, ultrastructural, and cell kinetic study. Am J Surg Pathol 1986;10:256–67.
29. Ryu JH, et al. Desquamative interstitial pneumonia and respiratory bronchiolitis-associated interstitial lung disease. Chest 2005;127:178–84.
30. Portnoy J, et al. Respiratory bronchiolitis-interstitial lung disease: long-term outcome. Chest 2007;131:664–71.
31. Katzenstein AL, Fiorelli RF. Nonspecific interstitial pneumonia/fibrosis. Histologic features and clinical significance. Am J Surg Pathol 1994;18:136–47.
32. Travis WD, Matsui K, Moss J, et al. Idiopathic nonspecific interstitial pneumonia: prognostic significance of cellular and fibrosing patterns: survival comparison with usual interstitial pneumonia and desquamative interstitial pneumonia. Am J Surg Pathol 2000;24:19–33.
33. Cordier JF. Cryptogenic organising pneumonia. Eur Respir J 2006;28:422–46.
34. Kamen DL, Strange C. Pulmonary manifestations of systemic lupus erythematosus. Clin Chest Med 2010;31:479–88.
35. Hassoun PM. Pulmonary arterial hypertension complicating connective tissue diseases. Semin Respir Crit Care Med 2009;30:429–39.
36. Mok MY, et al. Bosentan use in systemic lupus erythematosus patients with pulmonary arterial hypertension. Lupus 2007;16:279–85.
37. Qian J, et al. Survival and prognostic factors of systemic lupus erythematosus-associated pulmonary arterial hypertension: a PRISMA-compliant systematic review and meta-analysis. Autoimmun Rev 2016;15:250–7.
38. Schwarz MI, Lynch DA, Tuder R. Bronchiolitis obliterans: the lone manifestation of rheumatoid arthritis? Eur Respir J 1994;7:817–20.
39. Gabbay E, et al. Interstitial lung disease in recent onset rheumatoid arthritis. Am J Respir Crit Care Med 1997;156:528–35.
40. van den Hoogen F, et al. 2013 classification criteria for systemic sclerosis: an American college of rheumatology/European league against rheumatism collaborative initiative. Ann Rheum Dis 2013;72:1747–55.
41. Highland KB. Recent advances in scleroderma-associated pulmonary hypertension. Curr Opin Rheumatol 2014;26:637–45.

MULTIPLE-CHOICE QUESTIONS

1. The most common pulmonary manifestations of systemic lupus erythematosus (SLE) is:
   A. Interstitial lung disease (ILD)
   B. Diffuse alveolar hemorrhage (DAH)
   C. Pleural effusion
   D. Pulmonary hypertension

2. Idiopathic pulmonary fibrosis (IPF) is characterized by:
   A. The pathological lesion of usual interstitial pneumonitis
   B. A relentless progression and a poor prognosis despite therapeutic intervention
   C. The presence of interstitial opacities in a basilar and peripheral distribution
   D. All of the above

3. Which of the following statement regarding systemic sclerosis (SSc) is correct?
   A. Renal involvement and renal crisis remain the predominant cause of mortality in SSc.
   B. Pulmonary hypertension is a newly added diagnostic criteria for SSc.
   C. The most common lung pathological lesion in SSc is lymphocytic interstitial pneumonitis (LIP).
   D. Calcium channel blockers are indicated for patients with SSc-associated pulmonary hypertension.

73

Sarcoidosis

Edward S. Chen, David R. Moller

Sarcoidosis is a systemic inflammatory disorder characterized by the presence of noncaseating granulomas and lymphocyte inflammation.[1,2] The disease most frequently involves lungs and thoracic lymph nodes, although any organ system can be involved (Table 73.1). The manifestations and clinical course of sarcoidosis vary greatly.[3] An estimated 30–60% of patients with sarcoidosis are asymptomatic, usually with isolated bilateral hilar adenopathy, and clearly some of these patients remain unidentified. Symptomatic manifestations most frequently involve the respiratory system, with symptoms of cough, dyspnea, and chest discomfort. Löfgren syndrome is a distinct presentation of acute sarcoidosis, with polyarthritis, bilateral hilar adenopathy, erythema nodosum, and often uveitis. This form of sarcoidosis usually resolves completely. Other clinical manifestations depend on the range and extent of extrapulmonary involvement. More than one-third of patients will experience chronic active disease, which is associated with considerable risk of long-term morbidity resulting from progressive organ dysfunction and complications of immunosuppressive treatment. Given the lack of validated biomarkers of disease activity, the management of sarcoidosis involves serial assessment of symptoms, radiographic imaging, and objective measurement of organ function.

EPIDEMIOLOGY

Sarcoidosis is found worldwide, although there are striking differences in the prevalence of the disease in different geographical areas and racial groups.[4] In Europe and North America, the estimated prevalence ranges from 10 to 64 cases per 100 000, with higher rates in African-American populations. Although it can occur at any age, over 80% of cases are diagnosed between 20 and 50 years of age, with a further peak after age 50 years.[4,5] Globally, there is a slight female preponderance. Sarcoidosis may present differently in men and women,[6] and it presents differently in older people, complicating case identification.[7] Recent studies have supported prior observations of worse outcomes in African American women compared with other demographic groups in the United States.[8,9]

The frequency of clinical manifestations and disease severity varies among different groups. Löfgren syndrome has a particularly high frequency in the Scandinavian countries and in Ireland but occurs in <5% of African American patients with sarcoidosis. Lupus pernio, a disfiguring nodular facial condition associated with chronic sarcoidosis, is more frequent in patients of African descent. Cardiac involvement occurs in >50% of Japanese patients with sarcoidosis, but in only 10–25% of European and North American patients. Retrospective studies suggest that sarcoidosis is the direct cause of death in 1–6% of cases, usually from pulmonary, cardiac, or neurological involvement.[10]

GENETICS

There is substantial evidence for a genetic predisposition to sarcoidosis, including higher concordance among monozygotic twins than among dizygotic twins and familial clustering in 5–16% of patients. A US-based multicenter study compared 706 newly diagnosed cases of biopsy-proven sarcoidosis to age-, sex-, and race-matched relatives and found an approximate fivefold increased relative risk in siblings and parents of individuals with sarcoidosis.[11]

Major histocompatibility complex (MHC) class II alleles are major contributors to disease susceptibility across different ethnic populations.[12] Human leukocyte antigen–D related 3 (HLA-DR3) is associated with increased sarcoidosis risk in Scandinavian and European populations, whereas HLA-DR1 and HLA-DR4 alleles are associated with disease protection. A significant association was found with HLA-DRB1*11:01 in African Americans and Caucasians, whereas HLA-DRB1*15:01 was a risk factor only in Caucasians. Residues forming pocket 4 of HLA-DR and pocket 9 of HLA-DQ seem to influence sarcoidosis risk, which suggests specific antigenic peptides are involved in causing sarcoidosis. HLA class I alleles (B*07, A*03) also confer risk for sarcoidosis, independently from HLA class II alleles. Together, these HLA associations with sarcoidosis may reflect the effects of an 8.1 ancestral haplotype (*A1, B8, DR3, DQ2*) that is linked to several other immune-related diseases; larger studies will be necessary to determine the independent risks associated with each allele.

HLA class I and II alleles are associated with clinical course. HLA DRB1*03:01 and DQB1*02:01 are associated with favorable outcomes, whereas DR14 and DR15 are associated with severe, chronic disease in European and Japanese populations, as are the closely linked alleles DRB1*1501:01 and DQB1*06:02.

Studies of non-HLA polymorphisms have yielded few clues to the genetic basis of sarcoidosis. A meta-analysis found polymorphism of the gene for tumor necrosis factor-α (TNF-α) associated with a 1.5-fold increased risk of developing sarcoidosis.[13] Associations were reported between sarcoidosis and the CC chemokine receptors CCR2 and CCR5 and the receptor for

TABLE 73.1 Major Clinical Features of Systemic Sarcoidosis

Organ System (Approx. % Involvement)	Major Clinical Features
Pulmonary (90)	Bilateral hilar adenopathy, restrictive and obstructive disease, fibrocystic disease, bronchiectasis, mycetomas
Upper airway (5–10)	Hoarseness, laryngeal or tracheal obstruction, nasal congestion, sinusitis, saddle nose deformity
Ocular (25)	Anterior and posterior uveitis, chorioretinitis, conjunctivitis, optic neuritis, glaucoma, lacrimal gland enlargement
Skin (20–30)	Cutaneous and subcutaneous nodules and plaques, erythema nodosum, lupus pernio, hypopigmented macules
Hepatic (10)	Hepatomegaly, pruritus, jaundice, cirrhosis
Cardiac (10–15)	Arrhythmias, heart block, cardiomyopathy, sudden death
Central nervous system (5–10)	Cranial neuropathies (*e.g.,* Bell palsy), aseptic meningitis, brain mass, seizures, obstructing hydrocephalus, myelopathy, polyneuropathy, mononeuritis multiplex, small-fiber neuropathy
Salivary and parotid gland (<10)	Salivary and parotid gland enlargement, sicca syndrome
Hematologic (30–50)	Lymphadenopathy, splenomegaly, hypersplenism, anemia, lymphopenia, thrombocytopenia
Joints/ musculoskeletal (10–20)	Polyarthritis, bone cysts, Achilles tendonitis, dactylitis, heel pain, myopathy
Endocrine (<10)	Hypercalciuria (more common), hypercalcemia, hypopituitarism, diabetes insipidus
Renal (<5)	Renal calculi, nephrocalcinosis, renal failure, epididymitis, testicular mass

advanced **g**lycation **e**nd products (RAGE), but wider confirmation is lacking. Although strongly associated with early-onset sarcoidosis in children, NOD2/CARD15/IBD1 has not been linked with sarcoidosis in adults.[14] No association has been found between sarcoidosis and serum angiotensin-converting enzyme (ACE), vitamin D–converting enzyme, or the *NRAMP1* gene (linked to tuberculosis).

Genome-wide association studies (GWAS) in both familial and sporadic cases of sarcoidosis have identified chromosomal regions both inside and outside the MHC region, which contribute to sarcoidosis susceptibility. The butyrophilin-like 2 (*BTNL2*) gene within the MHC locus is associated with sarcoidosis risk in Caucasians and, to a lesser extent, African Americans.[15] Despite its proximity to the MHC locus, the *BTNL2 rs206530* allele can confer increased risk for chronic active sarcoidosis in patients without Löfgren syndrome.[16,17] *BTNL2* belongs to the immunoglobulin (Ig) gene superfamily and is costimulatory for T-cell activation, providing a hypothetical link with sarcoidosis. The gene encoding annexin A11 has also been associated with sarcoidosis, potentially affecting granuloma formation through its involvement in apoptosis and cellular proliferation.[18] A gene interaction network analysis suggested that T-helper 1 (Th1) and Th17 signaling pathways are involved in sarcoid immunopathogenesis.[19]

ENVIRONMENTAL FACTORS

Environmental factors have been implicated in the etiology of sarcoidosis with reports of time–space clustering and a higher incidence of disease in the springtime months. Health care workers, military personnel, and firefighters may also have a higher incidence of sarcoidosis.

The largest study of sarcoidosis etiology found weak positive associations (odds ratio [OR] ≈1.5) with workplace exposure to insecticides, mold, mildew, and musty odors.[3] Sarcoidosis was not associated with exposure to heavy metals, wood dusts, or rural residence, but smoking appeared to be protective. Studies of rescue workers and others exposed to the heavy dust burden from the World Trade Center terror attack found an increased incidence of a sarcoidosis-like granulomatous pulmonary disease.[20] A significant proportion of these patients had extrapulmonary manifestations, including chronic nonerosive polyarthritis, suggesting that they fall within the spectrum of systemic sarcoidosis.[21]

Role of Infectious Agents

Since the initial descriptions of sarcoidosis, investigators have looked for evidence of microbial infections in sarcoidosis because of its clinical similarities to other infectious diseases, most notably tuberculosis. Polymerase chain reaction (PCR) studies have shown microbial DNA in sarcoidosis tissues consistent with prior exposure to mycobacteria. Overall mycobacterial DNA was detected 10–20-fold more often in sarcoidosis versus control tissue.[22] The Kveim reaction also suggests a transmissible agent—intradermal inoculation with an extract from sarcoidosis tissue induces epithelioid granulomas indistinguishable from those in sarcoidosis biopsies.[23] Several groups have now demonstrated lung and blood CD4+ and CD8+ T-cell responses to multiple mycobacterial antigens in sarcoidosis, including mKatG, *Mycobacterium tuberculosis* ESAT-6, Ag85, superoxide dismutase, and heat shock proteins.[24,25] These studies suggest that mycobacterial organisms trigger sarcoidosis in a subgroup of patients.

Studies from Japan report the presence of *Propionibacterium acnes* or *P. granulosum* genomes in 80–98% of sarcoidosis tissues from Japan and Europe but also in up to 60% of control tissues.[26] In animal models, *P. acnes* proteins have been shown to induce Th1-driven granulomatous inflammation, partly via Toll-like receptor 9 (TLR9) stimulation. Given that these are commensal organisms on normal skin and upper airways, their pathogenic role remains undefined.

Although direct demonstration of an infectious etiology remains unproven, most investigators favor the hypothesis that certain classes of microbial organisms trigger sarcoidosis in those with genetic susceptibility.

AUTOIMMUNITY

Patients with sarcoidosis often have low-titer autoantibodies that may be part of the generalized hypergammaglobulinemia observed in many patients as a result of T-cell activation.[27] No disease-specific autoantibodies have yet been identified, so it remains unknown whether autoimmunity plays a role in sarcoidosis.

IMMUNOLOGY/PATHOGENESIS

KEY CONCEPTS

Immunopathogenesis of Sarcoidosis

- Noncaseating epithelioid granulomas
- Genetic susceptibility primarily involving major histocompatibility complex (MHC) genes
- Oligoclonal expansion of T-cell receptor specific T cells consistent with antigen-driven inflammation
- T-helper 1 (Th1) polarization at sites of disease
  - Decrease in regulatory T-cell (Treg) function contributes to Th1-mediated inflammation
  - Potential role of interferon (IFN)-γ–expressing Th17.1 cells
  - Potential role of classical Th17 effector cells
- Etiology involves microbial triggers
  - Mycobacterial or propionibacterial organisms most commonly implicated
- Role of innate immunity in macrophage activation and granuloma formation
- Progressive accumulation of serum amyloid A (SAA) within granulomas may drive chronic disease
  - Nidus for granuloma formation
  - Feed-forward amplification of local Th1 responses in part through Toll-like receptor 2 (TLR2)

The histological hallmark of sarcoidosis is the presence of discrete, compact noncaseating granulomas (Fig. 73.1). The dominant cell in the central core is the epithelioid cell, thought to be a differentiated mononuclear phagocyte. CD4+ T cells and mature macrophages are interspersed throughout the epithelioid core, whereas CD4+ and CD8+ T cells are seen around the periphery of the granuloma. Multinucleated giant cells are scattered throughout the inflammatory locus. In the lung, granulomas tend to form along areas that are rich in lymphatic vessels, such as the bronchovascular, bronchial submucosal, and interlobular septal regions.[28]

Prior to the 1980s, sarcoidosis was thought to be a disease of immune depression with blood lymphopenia and cutaneous anergy. Subsequently, bronchoalveolar lavage (BAL) studies in pulmonary sarcoidosis showed that there were, in fact, enhanced cell-mediated immune processes at sites of granulomatous inflammation with activated CD4 BAL lymphocytes. These BAL T cells are $CD4^+$ with a $CD45RO^+$ "memory" phenotype expressing high levels of activation molecules DR and VLA-1 (very late antigen-1, CD49a). Direct evidence that sarcoidosis is an antigen-driven disorder comes from studies of T-cell receptor (TCR) gene expression. These studies document the expansion of oligoclonal populations of T cells expressing specific α/β or γ/δ TCR in BAL, blood, or skin (Kveim biopsy). The most compelling data were derived from HLA-DRB1*03:01-positive Swedish patients who have greatly increased numbers of $AV2S3^+$ Th1 cells in the lung. The expanded oligoclonal $\alpha\beta^+$ T-cell subsets in sarcoidosis often contain shared amino acid motifs in the CDR3 region of their $V\beta$ or $V\alpha/J\alpha$ genes, consistent with an antigen-driven T-cell response. The antigenic specificities of these expanded T-cell clones remain to be identified.

Th1 Polarization in Sarcoidosis

T cells at sites of disease show a highly polarized Th1 cytokine profile at the time of diagnosis (Fig. 73.2).[29] BAL studies have demonstrated expression of interferon-γ (IFN-γ) and tumor necrosis factor (TNF) in pulmonary sarcoidosis. Consistent with a polarized type 1 process, Th1-promoting cytokines interleukin-12 (IL-12), -18, -2, -15, and -27 are upregulated in sarcoidosis-affected tissues. Chemokines and chemokine receptors that are typically associated with Th1 responses, such as CXCR3, CCR5, MCP-1, RANTES, MIP-1, MIG, and osteopontin (early T-lymphocyte activation protein), are upregulated in sarcoidosis, as are the Th1-differentiation transcription factor signal transducer and activator of transcription 1 (STAT1) and its phosphorylated form. Th1-associated gene expression signatures in sarcoidosis tissues have linked CXCL9 with chronic disease activity.[30]

Other evidence of the importance of polarized Th1 immunity in sarcoidosis comes from clinical experience with biological therapies with Th1-promoting effects, such as IFN-α, IFN-γ, and IL-2, which have been associated with new or recrudescent

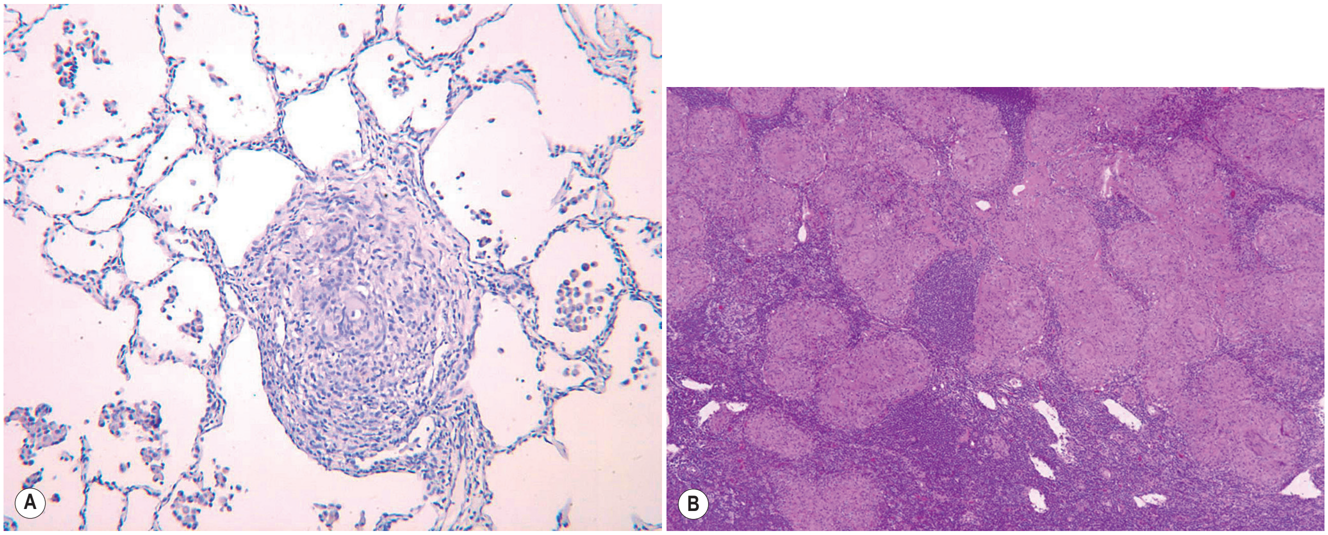


FIG 73.1 (A) Open lung biopsy showing typical noncaseating epithelioid granuloma, giant cells, and lymphocytic infiltrates in lung parenchyma in sarcoidosis. (B) Lymph node biopsy showing extensive replacement with well-defined epithelioid granulomas in a patient with sarcoidosis.

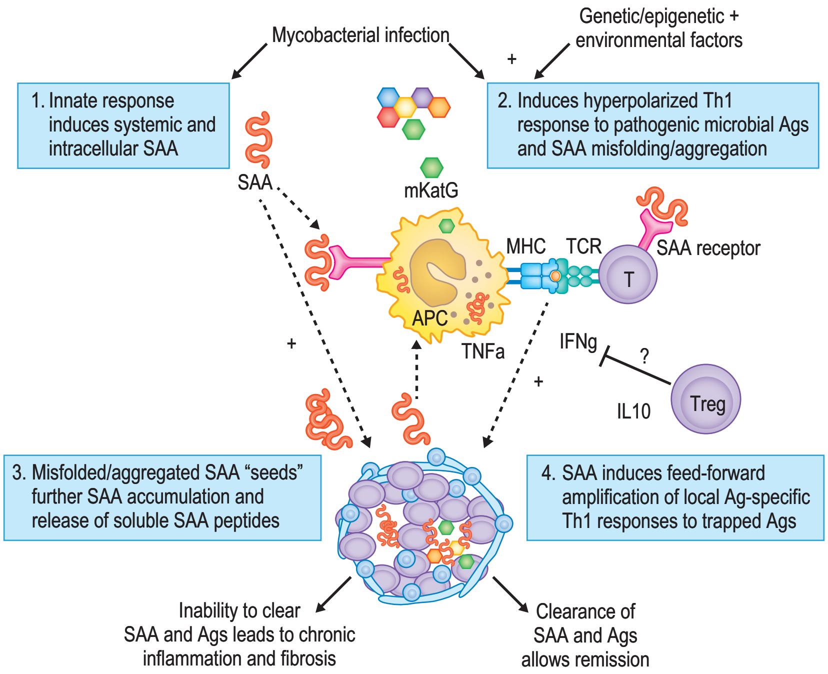

Mycobacterial infection
Genetic/epigenetic + environmental factors
+
1. Innate response induces systemic and intracellular SAA
2. Induces hyperpolarized Th1 response to pathogenic microbial Ags and SAA misfolding/aggregation
SAA
mKatG
MHC
TCR
SAA receptor
T
APC
TNFa
IFNg
?
IL10
Treg
+
+
3. Misfolded/aggregated SAA "seeds" further SAA accumulation and release of soluble SAA peptides
4. SAA induces feed-forward amplification of local Ag-specific Th1 responses to trapped Ags
Inability to clear SAA and Ags leads to chronic inflammation and fibrosis
Clearance of SAA and Ags allows remission

FIG 73.2 Conceptual Model of the Immunopathogenesis of Sarcoidosis. Granuloma formation in sarcoidosis results from stimulation by poorly soluble antigens that evoke a hyperimmune T-helper 1 (Th1) response with stimulation of interferon (IFN)-γ, along with tumor necrosis factor (TNF), interleukin (IL)-12, IL-10, and other cytokines from mononuclear phagocytes and dendritic cells. As a consequence of this response, misfolded serum amyloid A (SAA) aggregates in an amyloid-like process to provide a persistent poorly soluble nidus and a template for further SAA aggregation within sarcoidosis granulomas. SAA and SAA peptides released from the granulomas stoke a feed-forward stimulation of macrophages and T cells that amplifies polarized Th1 responses to local tissue antigens. This course continues to progress unless there is removal of stimulating antigen(s) and clearance of SAA, allowing remission of Th1-driven granuloma formation. The model depicts mycobacterial organisms as the etiological trigger, although other microbes or environmental agents might trigger a similar pathogenic pathway.

sarcoidosis. Conversely, low or undetectable levels of IL-4, IL-5, and Th2-associated chemokines and chemokine receptors are found in the sarcoidosis lung.[31] Sarcoidosis has also been described in patients with 5q-myelodysplasia with deletion of Th2 genes.

Th17 effector responses may drive the granulomatous response in sarcoidosis, although they are a minority of CD4 effector T cells. Most IFN-γ–producing cells in BAL from patients with sarcoidosis bear a Th17 phenotype termed *Th17.1*, although their transcriptional profile remains to be studied.[32] BAL cells from patients with Löfgren syndrome strongly express transcriptional regulators of Th1 (T-bet$^+$) and Th17 (RORγT$^+$).[33] Further work is needed to assess whether and how a Th1/Th17 permissive environment affects clinical outcomes in sarcoidosis.

CD4+ regulatory T cells (Tregs) show reduced numbers and function in sarcoidosis.[34,35] This suggests that Treg deficiency contributes to the imbalanced Th1 (and Th17) responsiveness in these patients. CD4$^+$CD25$^{\text{bright}}$FOXP3$^+$ T cells (natural Tregs) accumulate at the periphery of sarcoidosis granulomas, in BAL fluid, and in the peripheral blood of patients with active disease but are functionally deficient, with reduced ability to inhibit granuloma formation *in vitro*.[34] Nebulized vasoactive intestinal peptide (VIP) has been shown to significantly reduce TNF production by lung cells in sarcoidosis and increased the frequency of lung CD4$^+$CD127$^-$CD25$^+$ Tregs, possibly by converting naïve CD4$^+$CD25$^-$ T cells to Tregs, suggesting that inhaled VIP might be a treatment for sarcoidosis.[36] Longitudinal studies will be necessary to confirm the relationship between T-cell subsets and the clinical course of sarcoidosis. Other studies find a reduced number of immunoregulatory CD1d-restricted Vα24$^+$ natural killer T (NKT) cells in the blood of patients with nonremitting sarcoidosis, suggesting that lack of these cells may play a role in chronicity.

Innate Immunity in Sarcoidosis

Far fewer studies have looked at innate immune mechanisms, but evidence is mounting that TLRs and other innate immune receptors may be involved in sarcoidosis. For example, enhanced responses to TLR2 stimulation, including induction of TNF, have been observed in the lung and blood cells of patients with sarcoidosis.[37] TLR2 interacts with innate ligands as a heterodimer with other TLR subunits. Analysis of the TLR10–TLR1–TLR6 gene cluster on chromosome 4 has suggested that the absence of the common haplotype is associated with increased risk for developing chronic active sarcoidosis.[38]

Patients with sarcoidosis also show increased responsiveness to ligands for the TLR-2/1 heterodimer and decreased responses TLR-2/6 ligands.[37] Together, these findings support a potential role for aberrant TLR2 responsiveness in sarcoidosis.

A major conceptual challenge in sarcoidosis is understanding how pulmonary fibrosis occurs in an environment dominated by IFN-γ, which is known to inhibit collagen synthesis. Alternatively activated macrophages (M2) have been histologically associated with myofibrosis in biopsy specimens of muscular sarcoidosis.[39] Although an M2 macrophage phenotype is often associated with a Th2 cytokine environment (which has never been identified in sarcoidosis, even in late fibrotic disease), a fibrosis-promoting M2c macrophage phenotype could be induced by IL-10 or CCL18, which are present in sarcoidosis tissues.[40,41] M2 macrophages can promote fibrocyte recruitment through the expression of transforming growth factor (TGF), CCL18, and CXCL12; M2 macrophages are also capable of differentiation to fibrocyte-like cells that express collagen. It remains unclear whether M2-like macrophages result from mechanisms inherent to sarcoidosis or from generalized wound-healing responses.

Serum Amyloid A Aggregation Hypothesis

Although sarcoidosis is associated with prior exposure to mycobacterial or other microbial organisms, there is no clinical or pathological evidence of active infection or reactivation of latent microbial infection in sarcoidosis. This holds true in patients after years of corticosteroid, immunosuppressive, or anti-TNF therapy.

The physicochemical properties of the Kveim reagent closely resemble those of amyloid or prion proteins. Serum amyloid A (SAA) can be detected in epithelioid macrophages in sarcoidosis. This appears to be disease specific and several orders of magnitude greater than in other granulomatous disorders.[42,51] SAA stimulates expression of TNF, Th1-related cytokines, and IL-10 via TLR2 in the lung cells of patients with sarcoidosis. Sarcoidosis may therefore involve the induction, misfolding, and aggregation of SAA as a result of an aberrant hyperpolarized Th1 response to specific microbial triggers (see Fig. 73.2). Misfolded SAA may promote progressive self-aggregation in an amyloid-like process, with SAA and its peptides amplifying Th1 responses to pathogenic tissue antigens present within sarcoidosis granulomas. This mechanism could explain the cardinal clinical feature of chronic monophasic progressive inflammation in untreated sarcoidosis, in the absence of any tissue infection.[1]

PATIENT EVALUATION AND DIFFERENTIAL DIAGNOSIS

Pulmonary Sarcoidosis

The most common symptoms of pulmonary sarcoidosis are cough, progressive shortness of breath, and ill-defined chest discomfort of variable severity (see Table 73.1). Chronic sputum production and hemoptysis are more frequent in advanced fibrocystic disease. Typically, there are few physical findings. Lung crackles are heard in <20% of patients, and clubbing is rare. Pulmonary hypertension and cor pulmonale are important but underrecognized complications, associated with higher mortality rates.

Chest radiography (CXR) results are abnormal in over 90% of sarcoidosis patients. By convention, the chest radiograph is divided into the following stages:

0: Normal CXR
I: Bilateral hilar adenopathy
II: Bilateral hilar adenopathy + interstitial infiltrates
III: Interstitial infiltrates only (nonfibrotic)
IV: Fibrotic interstitial lung disease

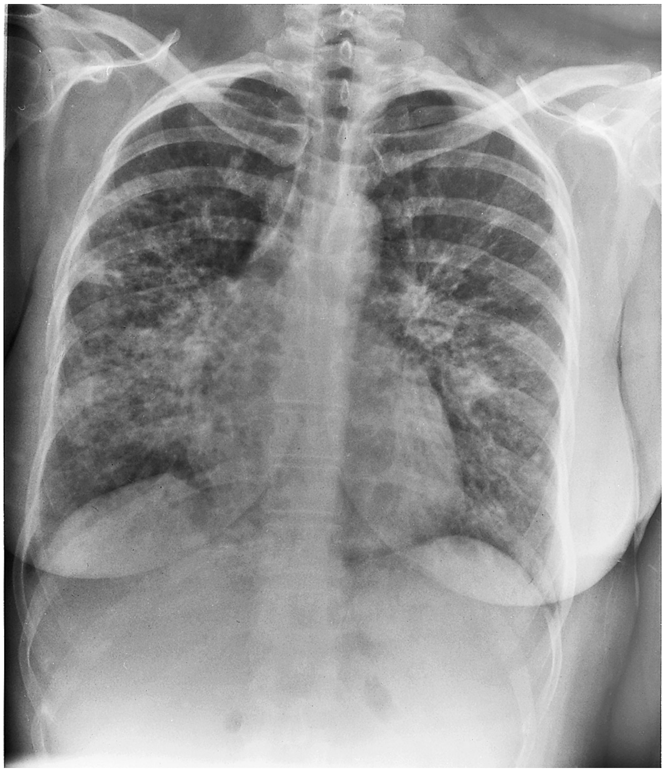


FIG. 73.3 Stage II chest radiograph with bilateral hilar adenopathy and reticulonodular infiltrates.

Stage I is initially seen in 40–50% of cases. Typically, the hilar adenopathy is discrete, symmetrical, and often accompanied by right paratracheal adenopathy. Stage II CXR, often with mid- or upper-zone infiltrates, is seen in 20–30% of presenting cases (Fig. 73.3). Stage III is seen in 10–20% of cases at presentation. Those with fibrotic changes are frequently placed in a separate subgroup, stage IV, in recognition of the poor outcome of this group of patients (Fig. 73.4). More unusual patterns of pulmonary sarcoidosis include large, well-defined, nodular infiltrates; miliary disease; or a pattern of patchy air-space consolidation with air bronchograms, termed “alveolar sarcoidosis.” Pleural effusions and pneumothorax are rare.

Computed tomography (CT) of the chest is more sensitive than plain radiography in demonstrating enlarged lymph nodes and pulmonary infiltrates. Chest CT typically demonstrates nodular infiltrates that follow central bronchovascular structures (Fig. 73.5). Studies correlating the value of the CXR stage against objective assessments of chest CT are lacking. However, in childhood sarcoidosis, longitudinal scoring of the findings on high-resolution CT (HRCT) of the chest correlates with changes in pulmonary function.[43]

Pulmonary function tests do not correlate well with CXR stage. In patients with stage I CXR, pulmonary function tests are normal in about 80% of cases or show just isolated reduction in carbon monoxide diffusing capacity (DL_{co}). When pulmonary infiltrates are present on CXR, 40–70% of cases show restrictive impairment with reduced forced vital capacity (FVC) and forced expiratory volume per second (FEV_1), and/or reduced DL_{co}. Obstructive impairment is found in 30–50% of patients with abnormal CXR results, mainly in advanced fibrocystic sarcoidosis. Bronchial hyperreactivity is present in 10–30% of patients.

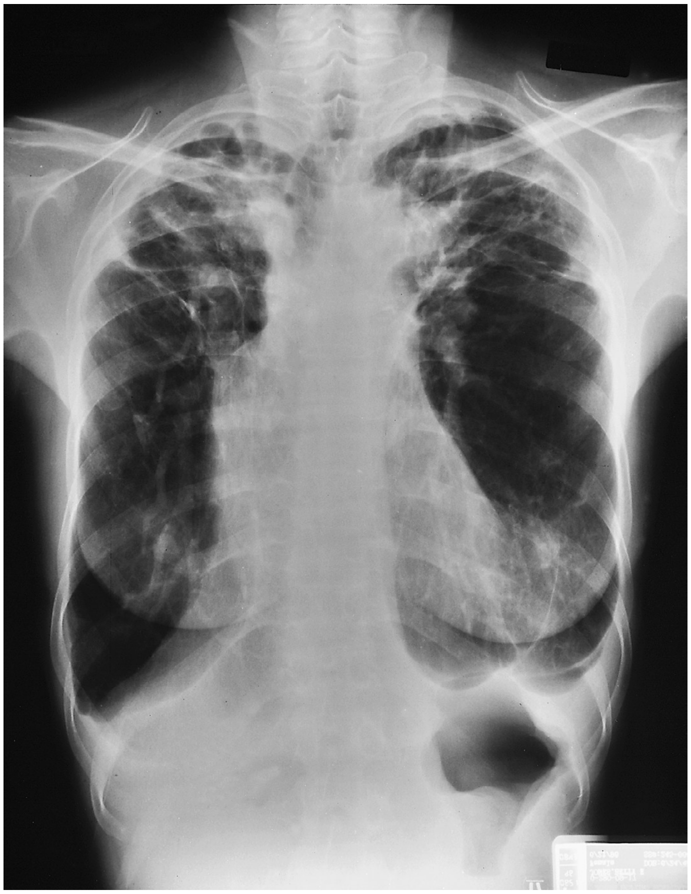


FIG 73.4 Stage IV chest radiograph of fibrocystic pulmonary sarcoidosis with typical upward hilar retraction and multiple cystic and bullous changes.

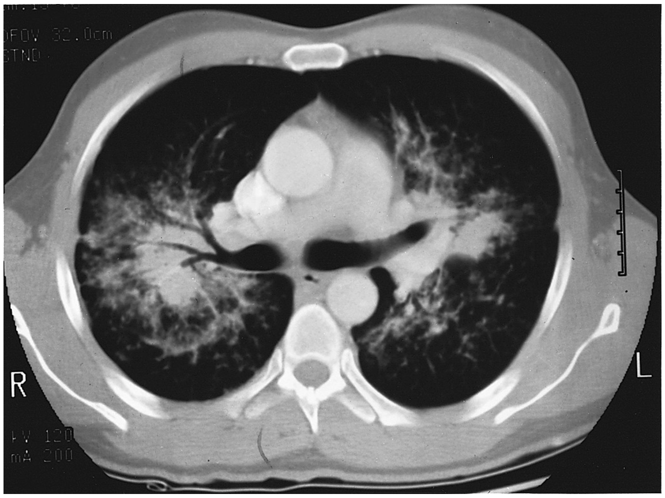


FIG 73.5 Computed tomography scan demonstrating bilateral hilar, paratracheal, and paraaortic lymphadenopathy with interstitial infiltrates in a dominantly central bronchovascular distribution.

Pulmonary hypertension is an underrecognized complication of pulmonary sarcoidosis affecting at least 5% patients and up to 50% of patients with dyspnea that is not explained by the level of pulmonary function impairment.[44] Screening for pulmonary hypertension is indicated in patients with advanced lung disease because of the associated increased mortality.

Extrapulmonary Sarcoidosis

Most patients have clinically important involvement of more than one organ, either with or without pulmonary sarcoidosis (see Table 73.1).

Sarcoidosis of the Upper Respiratory Tract

Sarcoidosis can cause nasal congestion, crusting, epistaxis, anosmia, rhinorrhea, sinusitis, or nasal septal perforation; a "saddle nose" deformity, palatal perforation, or bone erosion may occur. Laryngeal sarcoidosis occurs in 1–5% of patients, with hoarseness, dysphonia, dysphagia, dyspnea, or, rarely, stridor and acute respiratory failure.

Ocular Sarcoidosis

Unilateral or bilateral anterior uveitis is the most common manifestation of ocular sarcoidosis, often in association with bilateral hilar adenopathy. Chronic uveitis occurs in up to 20% patients with chronic sarcoidosis, more frequently in Black populations. Other manifestations include posterior uveitis, granulomatous conjunctivitis, severe chorioretinitis, or optic neuritis that can cause blindness (Chapter 74).

Cutaneous Sarcoidosis

Erythema nodosum, which is a feature of Löfgren syndrome, may recur. Chronic skin sarcoidosis usually manifests as nontender, nonpruritic plaques and subcutaneous nodules around the hairline, eyelids, ears, nose, mouth, and the extensor surfaces of arms and legs. Lupus pernio is a disfiguring form of facial cutaneous sarcoidosis, with violaceous plaques and nodules covering the nose, nasal alae, and malar areas and the area around the eyes.

Cardiac Sarcoidosis

Cardiac involvement is clinically apparent in <5–10% of patients with sarcoidosis in North America and Europe, although autopsy studies have suggested that the prevalence may exceed 25%. In Japan, cardiac involvement is more common, occurring in more than 50% of patients. Complete heart block, bundle branch blocks, ventricular arrhythmias, sudden death, cardiomyopathy, supraventricular arrhythmias, and valvular dysfunction may occur. Current advice is to screen for cardiac sarcoidosis at presentation, via assessment of symptoms (palpitations, syncope), electrocardiography (ECG), and echocardiography.[45]

Hepatic Sarcoidosis

Symptoms of hepatic sarcoidosis include fever, tender hepatomegaly, and/or pruritus. Characteristically, serum alkaline phosphatase and γ-glutamyl transferase are elevated proportionately higher than aspartate and alanine aminotransferases or bilirubin, although all patterns can occur. "Abdominal sarcoidosis" refers to a constellation of liver, spleen, and abdominal lymph node involvement, which is often associated with hypercalcemia; pulmonary involvement may not be evident.

Joints and Bones

Arthralgias are common in multisystem sarcoidosis, although joint radiographs are usually normal. Acute, often incapacitating, polyarthritis involving the ankles, feet, knees, and wrists may occur in Löfgren syndrome. Persistent joint disease with pain, swelling, and tenderness of the phalanges of the hands and feet is found in <5% of patients with chronic sarcoidosis. Joint radiographs may demonstrate "punched-out" lesions with cystic changes without erosive chondritis. Cystic lesions of the long bones, pelvis, sternum, skull, and vertebrae are rare.

Neurosarcoidosis

Neurosarcoidosis occurs in 5–15% of patients with sarcoidosis. The most common manifestation is cranial neuropathy, most frequently bilateral or unilateral seventh-nerve (facial) palsy. The palsy may resolve spontaneously or with corticosteroids but sometimes recurs years later. Optic neuritis may result in blindness. Manifestations of central nervous system (CNS) involvement include mass lesions, aseptic meningitis, obstructive hydrocephalus, and hypothalamic/pituitary dysfunction. Seizures, headache, change in mental status, confusion, and diabetes insipidus may be presenting symptoms. Myelopathy with paraparesis, hemiparesis, and back and leg pain have been described. Peripheral neuropathies account for ≈15% of cases of neurosarcoidosis, often presenting as mononeuritis multiplex or a primary sensory neuropathy. Small-fiber neuropathies with or without autonomic neuropathy are increasingly recognized in patients with sarcoidosis, particularly those with pain and chronic disease.

Salivary, Parotid, and Lacrimal Gland Sarcoidosis

Parotid or lacrimal gland enlargement and sicca syndrome may be the dominant clinical manifestations of sarcoidosis. Heerfordt syndrome, also known as *uveoparotid fever,* manifests as fever, parotid and lacrimal gland enlargement, uveitis, and bilateral hilar adenopathy, sometimes with cranial neuropathies (usually facial palsy).

Hematological Sarcoidosis

Peripheral lymph node enlargement occurs in 20–30% patients as an early manifestation of sarcoidosis but then typically remits spontaneously. Persistent, bulky lymphadenopathy occurs in <10% of patients. Splenomegaly occurs in <5% of cases. It is occasionally massive and often associated with hepatomegaly or hypercalcemia. Polyclonal hypergammaglobulinemia is present in ≥25% of patients.

Sarcoidosis Myopathy

Although random muscle biopsy specimens at autopsy have demonstrated muscle granulomas in most patients with sarcoidosis, symptomatic myopathy with weakness and tenderness is uncommon. Rarely, sarcoidosis can present as a polymyositis with profound weakness and elevated serum creatine kinase and aldolase.

Hypercalcemia, Hypercalciuria, and Renal Disease

Hypercalcemia is present in 2–5% of patients; hypercalciuria is more common. Abnormal calcium regulation may cause kidney stones or nephrocalcinosis and is thought to result from enhanced conversion of 25-OH vitamin D_3 to the active 1,25-$(OH)_2$ vitamin D_3 by macrophages and epithelioid cells in granulomas expressing 1α-hydroxylase (also known as *cytochrome p450 27B1*). Importantly, if patients with sarcoidosis are screened for vitamin D deficiency, they may have low 25-OH vitamin D_3 levels and elevated active 1,25-$(OH)_2$ vitamin D_3 levels, placing them at risk of nephrocalcinosis with renal failure if they are given vitamin D supplements.

Psychosocial Manifestations

Depression and/or fatigue have been found in 30–60% patients with symptomatic sarcoidosis. Depression is associated with female sex, lower socioeconomic status, poor access to care, and increased disease severity, but not with race.

KEY CONCEPTS

Clinical Relevance

- Multisystem disease with heterogeneous manifestations and clinical course
- Diagnosis based on consistent manifestations and biopsy showing typical pathology
    - No useful biomarkers for diagnosis or prognosis
    - Differential includes infections (mycobacterial, fungal), malignancy (*e.g.,* lymphoma), immunoglobulin G4 (IgG4)–related disease; interstitial lung diseases
- Association with T-helper 1 (Th1) –promoting therapies
    - After treatment with interferon (IFN)-α, IFN-γ, and interleukin (IL)-2 and other biologicals
    - After immune reconstitution in patients with human immunodeficiency virus (HIV) infection undergoing highly active retroviral therapy (HAART)
- Goal of treatment with immunosuppressive medications is to preserve function of affected organs while improving quality of life
    - Clinical trials needed to establish roles of individual therapies
    - Benefit of anti–tumor necrosis factor (TNF) therapies point to a role for TNF in sarcoidosis

Associated Conditions

Sarcoidosis and Pregnancy

Pregnancy usually has little effect on the course of sarcoidosis, although some patients experience spontaneous improvement. In those whose condition improves during pregnancy, exacerbations frequently follow several months after delivery. The temporary clinical improvement may relate to suppressed Th1 immunity during pregnancy.

Th1-Promoting Therapeutics

Administration of Th1-promoting therapeutics, such as IFN-α, IFN-γ, IL-2, and IFN-β, has been linked to initiation or recrudescence of sarcoidosis.

Common Variable Immunodeficiency

Sarcoidosis is associated with common variable immunodeficiency (CVID) in both adults and children. A high index of suspicion must be maintained for this in patients with sarcoidosis who have recurrent infections.

IgG4-Related Disease

This lymphoproliferative disorder involves multiple organ systems and can be mistaken for sarcoidosis, particularly when central or peripheral adenopathy are present. Distinguishing features of IgG4-related disease include the presence of dense lymphoplasmacytic infiltrates on biopsy, autoimmune pancreatitis and other autoimmune phenomena, and the presence of fibrosis in organs not commonly involved in sarcoidosis, such as the peritoneum, aorta, and thyroid gland.

Human Immunodeficiency Virus

Sarcoidosis may develop in patients infected by human immunodeficiency virus (HIV) with immune reconstitution after starting antiretroviral therapy, perhaps from reconstituted Th1 immunity. Granulomatous inflammation of skin or lungs is most often reported.

Autoimmune Disorders

Sarcoidosis has been described in association with autoimmune diseases, such as Crohn disease, ulcerative colitis, primary biliary cirrhosis, scleroderma, Sjögren syndrome, autoimmune hemolytic anemia, and autoimmune endocrinopathies. These associations could result from a common, predisposing altered Th1/Th17 immunity.

Cancer

Multisystem sarcoidosis may develop in patients with a recent history of cancer or following chemotherapy treatment, perhaps related to a rebound in immunological responsiveness following successful treatment.

Diagnosis

The diagnosis of sarcoidosis is based on a compatible clinical picture, histological evidence of noncaseating granulomas, and the absence of other known causes of this pathological response. Chronic beryllium disease and hypersensitivity pneumonitis must be excluded when there is a compatible history, and clinical findings are confined to the lung. In the absence of defined multisystem disease, sarcoidosis is a presumptive diagnosis, as local "sarcoid" reactions can occur in response to infection, tumor, or foreign material. Biopsy confirmation of sarcoidosis is not usually necessary in Löfgren syndrome, except in regions where histoplasmosis is endemic.

In general, biopsy of the easiest, most accessible abnormal tissue site is used to confirm the diagnosis. Biopsy of a skin nodule, superficial lymph node, lacrimal gland, nasal mucosa, conjunctivae, or salivary gland (lip biopsy) helps establish a diagnosis.

Biopsy by fiberoptic bronchoscopy is frequently used to diagnose pulmonary sarcoidosis because of its relative safety and high yield. The yield of transbronchial biopsy is operator dependent but approaches 50–85% of patients when pulmonary infiltrates are present. Endoscopic bronchial ultrasonography has improved the diagnostic yield of sampling intrathoracic lymph nodes in sarcoidosis. Mediastinoscopy or surgical lung biopsy is considered when lymphoma or other intrathoracic malignancy cannot be excluded and less invasive techniques have not given definitive answers. Imaging techniques, such as gadolinium-enhanced MRI, ^{18}F-fluorodeoxyglucose (FDG)-positron emission tomography (PET) scanning, or ^{67}gallium scanning, are nonspecific but may assist in assessing inflammation in the heart, brain, and bone or in directing biopsy.

Initial diagnostic evaluation of a patient with possible sarcoidosis should include tests for the presence and extent of pulmonary involvement and screening for extrathoracic disease. If cardiac symptoms are present, echocardiography and Holter monitoring should be performed. Cardiac magnetic resonance imaging (MRI) or cardiac PET have greater sensitivity in demonstrating patchy inflammation or scarring compared with gated technetium(^{99m}Tc) sestamibi scanning or echocardiography.[46] Endomyocardial biopsy is positive in <10–25% cardiac sarcoidosis cases because of the patchiness of the granulomatous inflammation, so a negative biopsy result never excludes the diagnosis. Diagnosis of cardiac sarcoidosis is usually made by histological confirmation of sarcoid at a noncardiac site in association with compatible cardiac imaging, conduction abnormalities, or arrhythmias.

CLINICAL PEARLS

Tests Recommended for an Initial Evaluation of a Patient With Sarcoidosis

- Chest radiography or chest computed tomography (CT)
- Pulmonary function tests
  - Spirometry
  - Diffusing capacity
  - Lung volumes
  - Flow–volume loop (if suspected upper airway obstruction)
- Slit-lamp examination (to exclude subclinical uveitis)
- Blood tests
  - Comprehensive metabolic panel (calcium; liver and renal functions)
  - Complete blood count, including differential and platelet count
  - Measurement of active $1,25\text{-(OH)}_2$ vitamin D_3
- Electrocardiography
- Purified protein derivative skin test or blood test for latent tuberculosis

MRI with gadolinium enhancement is the best way to detect the characteristic inflammatory lesions of CNS or spinal cord sarcoidosis, particularly in the periventricular and leptomeningeal areas. A normal scan does not exclude neurosarcoidosis, particularly for cranial neuropathies or during corticosteroid therapy. Cerebral spinal fluid characteristically demonstrates lymphocytic pleocytosis and/or elevated protein levels. Diagnosis of neurosarcoidosis is usually confirmed by biopsy of a non-CNS site (*e.g.*, lung or lymph node). Rarely, brain biopsy is needed to exclude infectious or malignant disease. In suspected cases of peripheral neuropathy or myopathy, electromyography (EMG) or nerve conduction studies should be considered. Small-fiber neuropathy can be confirmed by performing quantitative immunohistochemistry of appropriate skin biopsy specimens.

Other Diagnostic Studies

No noninvasive tests are useful in making a diagnosis of sarcoidosis. Serum ACE levels are elevated in 40–90% of patients with clinically active disease, but these are nonspecific and may be found in other inflammatory and granulomatous diseases as well.

CLINICAL COURSE AND PATIENT MANAGEMENT

The clinical course of sarcoidosis is highly variable. Overall, 50–65% of patients undergo spontaneous remission, usually within the first 2–3 years. Löfgren syndrome is associated with spontaneous remission in >70–80% of patients. Peripheral adenopathy, salivary and parotid gland enlargement, and Bell palsy generally subside spontaneously or with treatment and do not recur. Elevated serum liver function tests also may revert to normal without treatment. Remission is observed in approximately 50–80% of patients with a stage I CXR, 30–60% with a stage II CXR, and 20–30% with a stage III CXR. Patients with a stage IV CXR with fibrocystic changes rarely (<5%) experience remission.

Extrapulmonary disease that is severe at presentation tends to persist and require treatment. Overall, mortality ranges from 1–6%, with the major causes being advanced pulmonary, cardiac, and neurological involvement.[10]

The intensity of surveillance of sarcoidosis depends on the severity of clinical presentation. When sarcoidosis undergoes remission, the disease rarely recurs; exceptions often involve neurological or ocular manifestations. No biomarkers have been found to be useful in management decisions.

TREATMENT

THERAPEUTIC PRINCIPLES

Indications for Corticosteroid Therapy in Patients With Sarcoidosis

- Pulmonary involvement
  - Moderate or severe, symptomatic pulmonary disease
  - Progressive, symptomatic pulmonary disease
  - Persistent pulmonary infiltrates or abnormal lung function for 1–2 years with mild symptoms to assess reversibility
  - Advanced fibrocystic disease
- Extrapulmonary involvement
  - Threatened organ failure: severe ocular, cardiac, or central nervous system (CNS) disease
  - Posterior uveitis or anterior uveitis not responding to local steroids
  - Persistent hypercalcemia
  - Persistent renal or hepatic dysfunction
  - Pituitary disease
  - Myopathy
  - Palpable splenomegaly or evidence of hypersplenism, such as thrombocytopenia
  - Severe fatigue and weight loss
  - Painful lymphadenopathy
  - Disfiguring skin disease

Current therapies are nonspecific and largely unproven.[47] If there is no organ-threatening involvement at presentation, a period of observation is usually indicated to evaluate whether the disease will remit spontaneously.

Corticosteroids remain the mainstay of therapy for progressive organ-threatening or life-threatening manifestations of sarcoidosis. The optimal doses and duration of corticosteroid treatment have not been established by rigorous clinical studies. For pulmonary and cutaneous disease, initial treatment is usually with no more than 20–40 mg/day of prednisone for 2–4 weeks, followed by a slow tapering regimen over several months to a maintenance dose of 5–15 mg/day. Alternate-day therapy has been suggested, although it may be ineffective in a subgroup of patients who subsequently respond to daily dosing. Treatment is normally continued for at least 6–12 months, as premature tapering is likely to result in relapse. Recurrent, progressive pulmonary disease occurs in 16–74% of patients as corticosteroid therapy is tapered; patients with repetitive relapses usually need chronic suppressive therapy to minimize loss of lung function. Weight gain, hypertension, hyperglycemia, glaucoma, and osteoporosis are serious potential complications. Bisphosphonate therapy is recommended for patients on chronic corticosteroid therapy. Inhaled corticosteroids may help reduce symptoms of cough or airway irritability but are not useful as sole agents in pulmonary sarcoidosis.

Pentoxifylline, a phosphodiesterase inhibitor, was shown to be beneficial in one study of mild pulmonary sarcoidosis, although wider experience suggests only very few patients respond to it. Gastrointestinal side effects often limit the dose that is tolerated.

ALTERNATIVE CYTOTOXIC THERAPIES

Corticosteroid-sparing therapies are frequently used for chronic, progressive sarcoidosis when corticosteroid adverse effects significantly degrade quality of life.[2,47,48] Low-dose weekly methotrexate (10–20 mg/week) is recommended for pulmonary, cardiac, ocular (panuveitis), cutaneous, and neurological sarcoidosis requiring higher doses of corticosteroids. Success rates range from 50% to 70%, but it may take ≥6 months to be effective. Potential serious complications include hepatotoxicity, opportunistic infections, bone marrow suppression, and pulmonary toxicity. Azathioprine can also be useful in sarcoidosis with similar response rates to methotrexate but a slightly increased risk of infection. Mycophenolate mofetil and leflunomide have been used as steroid-sparing agents in small case series. Potential drug toxicities for all four drugs include bone marrow suppression, gastrointestinal symptoms, skin rashes, and an increased risk of malignancy. Cyclophosphamide has been used in steroid-recalcitrant sarcoidosis, particularly refractory neurosarcoidosis, but its use is severely limited because of its oncogenic potential. Calcineurin inhibitors suppress T-cell activation but do not work in sarcoidosis.

Clinical trials support the effectiveness of biological agents targeting specific immunological pathways in sarcoidosis. Infliximab (anti-TNF) had a small impact on pulmonary function over 6 months but was effective for many extrapulmonary manifestations. Case series have suggested that adalimumab may also be effective in some patients with sarcoidosis. Etanercept, golimumab (anti-TNF), and ustekinumab (anti–IL12/IL23) were ineffective in pulmonary sarcoidosis. Interestingly, there have been some case reports of successful use of rituximab (anti–B cell) in treatment-refractory sarcoidosis, particularly neurosarcoidosis, although how it may work in sarcoidosis remains unclear. Further studies are needed to define the role of biological agents (Chapter 89) in sarcoidosis, not least because these therapies carry risks of serious, life-threatening infections and malignancy.

SPECIFIC SITUATIONS

Löfgren Syndrome

Nonsteroidal antiinflammatory drugs (NSAIDs) are recommended for relief of constitutional symptoms and joint pains. For disabling arthritis or severe constitutional symptoms, corticosteroids are almost immediately effective and can generally be tapered over a few weeks to months.

Mucocutaneous and Joint Sarcoidosis and Hypercalcemia

Hydroxychloroquine is often tried as a first-line treatment for cutaneous, joint, nasal, or sinus manifestations when corticosteroids are not immediately needed. Hypercalcemia may also respond to hydroxychloroquine as maintenance therapy. Chloroquine may be effective in some patients unresponsive to hydroxychloroquine but has a greater potential for ocular toxicity and is rarely used. Minocycline and doxycycline have been effective in a small number of patients with cutaneous sarcoidosis, but wider experience shows that these drugs are rarely effective.

Ocular Sarcoidosis

Anterior uveitis is usually treated with topical corticosteroids. Oral or intravenous corticosteroids are used initially for posterior uveitis, chorioretinitis, or optic neuritis.

Pulmonary Hypertension

Small retrospective case series have suggested that pulmonary vasodilator treatment can reduce pulmonary artery pressure and improve exercise dyspnea and 6-minute walk distance in patients with sarcoidosis who have pulmonary hypertension.[49] The benefit of vasodilator therapy on survival rates remains uncertain.[50]

Cardiac Sarcoidosis

Initial treatment consists of cardiac medications (antiarrhythmic therapy, diuretics, and afterload-reducing agents) along with corticosteroid therapy that may reverse heart block, reduce arrhythmias, and improve cardiac function. Corticosteroids are recommended in moderate doses as some studies have found no difference in 5-year survival rates for patients treated with prednisone >30 mg/day versus <30 mg/day. Higher doses are often used initially for intractable arrhythmias or heart block. Cardiomyopathy may require long-term treatment, and cytotoxic drugs are often used for steroid sparing. Small case series have suggested that TNF inhibitors may be useful, but caution is advised with regard to their use, since these agents can worsen congestive heart failure in nonsarcoidosis cardiomyopathy. Automatic implantable defibrillators and/or pacemakers are indicated in patients at risk for sudden death from serious arrhythmias or heart block, although their prophylactic use remains controversial. The effectiveness of radiofrequency ablation for prevention of arrhythmias in cardiac sarcoidosis remains uncertain.

Neurosarcoidosis

High doses of oral corticosteroids (60–80 mg/day) or high-dose pulse intravenous therapy are often used for serious CNS involvement. Tapering should be performed over several months after confirming suppression of inflammation by objective criteria (*e.g.,* serial MRI scans). With the exception of cranial neuropathies, neurosarcoidosis tends to be chronic and requires long-term therapy.

Depression/Fatigue/Pain

Small case series have highlighted the existence of small-fiber neuropathy and autonomic neuropathy in sarcoidosis, particularly in patients with pain. A few case reports have suggested that anti-TNF therapies may be beneficial in small-fiber neuropathy, but their role in sarcoidosis remains undefined. Intravenous immunoglobulin (IVIG) has also been effective in a subset of patients with small-fiber neuropathy (Chapter 84). One study suggested that the stimulant dexmethylphenidate hydrochloride may be helpful in treating chronic fatigue in sarcoidosis. Therapies directed at treating depression may improve quality of life in these patients, although clinical trials are lacking.

Role of Transplantation in Sarcoidosis

Lung, heart, kidney, and liver transplantations have been performed successfully in patients with sarcoidosis. Recurrent granulomas are found in the transplanted organs of some patients but are generally responsive to increased immunosuppression. Overall, survival rates with lung and cardiac transplantations in sarcoidosis are similar to those in other organ-specific diseases.

OPPORTUNITIES FOR PROGRESS IN SARCOIDOSIS

ON THE HORIZON

- Diagnostic and prognostic tools based on:
  - Genotyping
  - Genomic/microbiome/proteomic and immunological signatures
- Improvements in imaging techniques to assist in assessment of organ involvement
- Multicenter clinical trials of immune-altering therapies for treatment of sarcoidosis

We have the necessary foundations for making rapid progress in understanding sarcoidosis. Although there is considerable clinical heterogeneity, distinct clinical phenotypes are well established, from acute sarcoidosis (Löfgren syndrome) to chronic progressive organ disease. It is known that the genetic basis of sarcoidosis predominantly resides within the MHC. Immunological hallmarks of the disease include noncaseating granulomas and highly polarized Th1 and Th17 immunity with proinflammatory cytokines, such as TNF, at sites of disease, associated with reduced Treg activity. Evidence from multiple centers suggests that mycobacterial and possibly propionibacterial organisms may trigger sarcoidosis. The recent identification of innate immune pathways, one of which involves SAA, suggests mechanisms that may explain the pathogenesis of chronic sarcoidosis in the absence of active chronic infection. The challenge for the next 5–10 years is to translate these laboratory discoveries into clinical tools to assist the clinician in providing the diagnosis and prognosis for individual patients. Newer therapies are sorely needed for the treatment of patients with sarcoidosis, particularly those who have chronic disease and generally require long-term maintenance therapy. The potential for curing this disease is suggested by a remission rate of >50%, although progress toward this goal awaits a better understanding of the mechanisms that allow remission to occur. Research that integrates assessment of clinical phenotype, genetic background, and individual immunophenotypical and environmental triggers in different subgroups of patients may offer the best chance for rapid progress.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Chen ES, Moller DR. Sarcoidosis-scientific progress and clinical challenges. Nat Rev Rheumatol 2011;7:457–67.
2. Valeyre D, Prasse A, Nunes H, et al. Sarcoidosis. Lancet 2014;383:1155–67.
3. Newman LS, Rose CS, Bresnitz EA, et al. A case control etiologic study of sarcoidosis: environmental and occupational risk factors. Am J Respir Crit Care Med 2004;170:1324–30.
4. Rybicki BA, Iannuzzi MC. Epidemiology of sarcoidosis: recent advances and future prospects. Semin Respir Crit Care Med 2007;28:22–35.
5. Morimoto T, Azuma A, Abe S, et al. Epidemiology of sarcoidosis in Japan. Eur Respir J 2008;31:372–9.
6. Grunewald J, Eklund A. Sex-specific manifestations of Lofgren's syndrome. Am J Respir Crit Care Med 2007;175:40–4.
7. Varron L, Cottin V, Schott AM, et al. Late-onset sarcoidosis: a comparative study. Medicine (Baltimore) 2012;91:137–43.
8. Mirsaeidi M, Machado RF, Schraufnagel D, et al. Racial difference in sarcoidosis mortality in the United States. Chest 2015;147:438–49.

9. Dumas O, Abramovitz L, Wiley AS, et al. Epidemiology of sarcoidosis in a prospective cohort study of U.S. women. Ann Am Thorac Soc 2016;13:67–71.
10. Swigris JJ, Olson AL, Huie TJ, et al. Sarcoidosis-related mortality in the United States from 1988 to 2007. Am J Respir Crit Care Med 2011;183:1524–30.
11. Rybicki BA, Iannuzzi MC, Frederick MM, et al. Familial aggregation of sarcoidosis. A case-control etiologic study of sarcoidosis (ACCESS). Am J Respir Crit Care Med 2001;164:2085–91.
12. Fischer A, Grunewald J, Spagnolo P, et al. Genetics of sarcoidosis. Semin Respir Crit Care Med 2014;35:296–306.
13. Medica I, Kastrin A, Maver A, et al. Role of genetic polymorphisms in ACE and TNF-alpha gene in sarcoidosis: a meta-analysis. J Hum Genet 2007;52:836–47.
14. Bello GA, Adrianto I, Dumancas GG, et al. Role of NOD2 pathway genes in sarcoidosis cases with clinical characteristics of Blau syndrome. Am J Respir Crit Care Med 2015;192:1133–5.
15. Rybicki BA, Walewski JL, Maliarik MJ, et al. The BTNL2 gene and sarcoidosis susceptibility in African Americans and Whites. Am J Hum Genet 2005;77:491–9.
16. Morais A, Lima B, Peixoto MJ, et al. BTNL2 gene polymorphism associations with susceptibility and phenotype expression in sarcoidosis. Respir Med 2012;106:1771–7.
17. Wennerstrom A, Pietinalho A, Lasota J, et al. Major histocompatibility complex class II and BTNL2 associations in sarcoidosis. Eur Respir J 2013;42:550–3.
18. Mirsaeidi M, Gidfar S, Vu A, et al. Annexins family: insights into their functions and potential role in pathogenesis of sarcoidosis. J Transl Med 2016;14:89.
19. Fischer A, Ellinghaus D, Nutsua M, et al. Identification of immune-relevant factors conferring sarcoidosis genetic risk. Am J Respir Crit Care Med 2015;192:727–36.
20. Crowley LE, Herbert R, Moline JM, et al. Sarcoid like" granulomatous pulmonary disease in World Trade Center disaster responders. Am J Ind Med 2011;54:175–84.
21. Loupasakis K, Berman J, Jaber N, et al. Refractory sarcoid arthritis in World Trade Center-exposed New York City firefighters: a case series. J Clin Rheumatol 2015;21:19–23.
22. Gupta D, Agarwal R, Aggarwal AN, et al. Molecular evidence for the role of mycobacteria in sarcoidosis: a meta-analysis. Eur Respir J 2007;30:508–16.
23. Teirstein AS. Kveim antigen: what does it tell us about causation of sarcoidosis? Semin Respir Infect 1998;13:206–11.
24. Chen ES, Wahlstrom J, Song Z, et al. T cell responses to mycobacterial catalase-peroxidase profile a pathogenic antigen in systemic sarcoidosis. J Immunol 2008;181:8784–96.
25. Drake WP, Dhason MS, Nadaf M, et al. Cellular recognition of Mycobacterium tuberculosis ESAT-6 and KatG peptides in systemic sarcoidosis. Infect Immun 2007;75:527–30.
26. Eishi Y, Suga M, Ishige I, et al. Quantitative analysis of mycobacterial and propionibacterial DNA in lymph nodes of Japanese and European patients with sarcoidosis. J Clin Microbiol 2002;40: 198–204.
27. Kobak S, Yilmaz H, Sever F, et al. The prevalence of antinuclear antibodies in patients with sarcoidosis. Autoimmune Dis 2014;2014:351852.
28. Rossi G, Cavazza A, Colby TV. Pathology of sarcoidosis. Clin Rev Allergy Immunol 2015;49:36–44.
29. Zissel G, Prasse A, Muller-Quernheim J. Immunologic response of sarcoidosis. Semin Respir Crit Care Med 2010;31:390–403.
30. Koth LL, Solberg OD, Peng JC, et al. Sarcoidosis blood transcriptome reflects lung inflammation and overlaps with tuberculosis. Am J Respir Crit Care Med 2011;184:1153–63.
31. Christophi GP, Caza T, Curtiss C, et al. Gene expression profiles in granuloma tissue reveal novel diagnostic markers in sarcoidosis. Exp Mol Pathol 2014;96:393–9.
32. Ramstein J, Broos CE, Simpson LJ, et al. Interferon-gamma-producing Th17.1 cells are increased in sarcoidosis and more prevalent than Th1 cells. Am J Respir Crit Care Med 2016;193:1281–91.
33. Kaiser Y, Lepzien R, Kullberg S, et al. Expanded lung T-bet+RORgammaT+ CD4+ T-cells in sarcoidosis patients with a favourable disease phenotype. Eur Respir J 2016;48:484–94.
34. Taflin C, Miyara M, Nochy D, et al. FoxP3+ regulatory T cells suppress early stages of granuloma formation but have little impact on sarcoidosis lesions. Am J Pathol 2009;174:497–508.
35. Broos CE, van Nimwegen M, Kleinjan A, et al. Impaired survival of regulatory T cells in pulmonary sarcoidosis. Respir Res 2015;16:108.
36. Prasse A, Zissel G, Lutzen N, et al. Inhaled vasoactive intestinal peptide exerts immunoregulatory effects in sarcoidosis. Am J Respir Crit Care Med 2010;182:540–8.
37. Gabrilovich MI, Walrath J, van Lunteren J, et al. Disordered Toll-like receptor 2 responses in the pathogenesis of pulmonary sarcoidosis. Clin Exp Immunol 2013;173:512–22.
38. Veltkamp M, van Moorsel CH, Rijkers GT, et al. Genetic variation in the Toll-like receptor gene cluster (TLR10-TLR1-TLR6) influences disease course in sarcoidosis. Tissue Antigens 2012;79:25–32.
39. Preusse C, Goebel HH, Pehl D, et al. Th2-M2 immunity in lesions of muscular sarcoidosis and macrophagic myofasciitis. Neuropathol Appl Neurobiol 2015;41:952–63.
40. Boot RG, Hollak CE, Verhoek M, et al. Plasma chitotriosidase and CCL18 as surrogate markers for granulomatous macrophages in sarcoidosis. Clin Chim Acta 2010;411:31–6.
41. Prasse A, Pechkovsky DV, Toews GB, et al. A vicious circle of alveolar macrophages and fibroblasts perpetuates pulmonary fibrosis via CCL18. Am J Respir Crit Care Med 2006;173:781–92.
42. Zhang Y, Chen X, Hu Y, et al. Preliminary characterizations of a serum biomarker for sarcoidosis by comparative proteomic approach with tandem-mass spectrometry in ethnic Han Chinese patients. Respir Res 2013;14:18.
43. Sileo C, Epaud R, Mahloul M, et al. Sarcoidosis in children: HRCT findings and correlation with pulmonary function tests. Pediatr Pulmonol 2014;49:1223–33.
44. Maimon N, Salz L, Shershevsky Y, et al. Sarcoidosis-associated pulmonary hypertension in patients with near-normal lung function. Int J Tuberc Lung Dis 2013;17:406–11.
45. Birnie DH, Sauer WH, Bogun F, et al. HRS expert consensus statement on the diagnosis and management of arrhythmias associated with cardiac sarcoidosis. Heart Rhythm 2014;11:1305–23.
46. Mehta D, Lubitz SA, Frankel Z, et al. Cardiac involvement in patients with sarcoidosis: diagnostic and prognostic value of outpatient testing. Chest 2008;133:1426–35.
47. Bradley B, Branley HM, Egan JJ, et al. Interstitial lung disease guideline: the British Thoracic Society in collaboration with the Thoracic Society of Australia and New Zealand and the Irish Thoracic Society. Thorax 2008;63(Suppl. 5):v1–58.
48. Baughman RP, Lower EE. Treatment of sarcoidosis. Clin Rev Allergy Immunol 2015;49:79–92.
49. Barnett CF, Bonura EJ, Nathan SD, et al. Treatment of sarcoidosis-associated pulmonary hypertension. A two-center experience. Chest 2009;135:1455–61.
50. Bonham CA, Oldham JM, Gomberg-Maitland M, et al. Prostacyclin and oral vasodilator therapy in sarcoidosis-associated pulmonary hypertension: a retrospective case series. Chest 2015;148:1055–62.
51. Chen ES, Song Z, Willett MH, et al. Serum amyloid A regulates granulomatous inflammation in sarcoidosis through Toll-like receptor-2. Am J Respir Crit Care Med 2010;181:360–73.

MULTIPLE-CHOICE QUESTIONS

1. Is cardiac sarcoidosis a rare manifestation?
2. True or false, sarcoidosis rarely results in airway obstruction.
3. True or false, stage IV sarcoidosis indicates advanced, widespread, multisystem disease.
4. True or false, Sarcoidosis represents an immune deficiency state.
5. True or false, calcineurin inhibitors are useful as treatment for sarcoidosis.

74

Immunological Ocular Disease

James T. Rosenbaum, Phoebe Lin

The immune system can induce disease in virtually any portion of the eye. Examples include conjunctivitis, keratitis, keratoconjunctivitis, uveitis, scleritis, optic neuritis, and orbital inflammation. The anatomy of the eye is depicted schematically in Fig. 74.1. Uveitis is defined as inflammation of the uveal tract, which is the middle layer of the eye, divided into the anterior uvea (iris, ciliary body) and posterior uvea (choroid). The uvea is sandwiched between an outer layer (sclera) and an inner layer (retina). The anterior segment is separated from the posterior segment by the lens. Two relatively common immunologically mediated ocular diseases, discussed elsewhere in this volume, are sicca secondary to Sjögren syndrome (Chapter 54) and anterior ischemic optic neuropathy secondary to giant cell arteritis (Chapter 59). Before considering the various ocular inflammatory disorders, it is critical to review some unique considerations related to ocular immunology.

OCULAR IMMUNE PRIVILEGE

Many of the mechanisms that drive inflammation in the eye are identical to those operating at other tissue sites. The major difference between the immunopathology of intraocular inflammation and that of systemic inflammatory disease relates to the fact that the eye, like the brain and the testis, is an immunologically privileged site. During uveitis, keratitis, and scleritis, as well as following corneal transplantation, a variety of local immunosuppressive mechanisms act to limit the damage caused by infiltrating leukocytes and, consequently, to influence the patient's clinical course. One of the most important factors is the constitutive expression of Fas ligand (FasL) within the eye. In addition, normal ocular tissues produce relatively high levels of immunomodulatory cytokines and immunosuppressive neuropeptides, as well as complement. Other factors include the blood–aqueous and blood–retinal anatomical barriers, limited major histocompatibility complex (MHC) expression, a paucity of lymphatic drainage channels within the eye, and, in the case of the cornea, the complete absence of blood vessels. The term *anterior chamber–associated immune deviation* (ACAID) describes the suppression of a cell-mediated immune response when soluble antigen is directly injected into the aqueous humor. All of the above factors contribute to ACAID.

FAS LIGAND

Most immune cells, including neutrophils, monocytes, macrophages, and lymphocytes, express Fas (CD95) on their surface. The interaction between Fas and its receptor FasL (CD95L) triggers apoptosis of the Fas-bearing cell. FasL is constitutively expressed within the eye, being detected in normal cornea, anterior uvea, and retina. The importance of FasL to ocular immune privilege has been demonstrated primarily in experimental models of ocular inflammation. Corneal allografts from FasL-negative mice into recipients of normal phenotype are rejected in all cases, whereas approximately half of FasL-positive grafts survive.[1] When injected into the eyes of FasL-deficient mice, herpes simplex virus (HSV) causes a severe invasive infection. A similar procedure in normal phenotype control animals results in relatively minor inflammation. Levels of FasL in the aqueous humor during human acute anterior uveitis are capable of inducing apoptosis in Fas-positive lymphoid cells. In this self-limiting condition, as well as in relevant rodent models, apoptosis of infiltrating T cells is observed early in the course of the inflammation. In addition to FasL, the local expression of tumor necrosis factor (TNF)–related apoptosis inducing ligand (TRAIL) by corneal endothelium also contributes to enhanced apoptosis and immune privilege within the eye.[2]

CYTOKINES, NEUROPEPTIDES, AND COMPLEMENT IN THE PROMOTION AND REGULATION OF OCULAR INFLAMMATION

The aberrant immune activation that occurs in uveitis appears to be caused by a combination of environmental triggers and genetic risk factors that tip the balance away from immune regulation and toward inflammation. Inflammatory pathways of the adaptive immune system that play a role in the pathogenesis of uveitis include T-helper 1 (Th1) lymphocytes (that produce interferon-γ [IFN-γ], interleukin-2 [IL-2], and IL-12), and IL-23-responsive Th17 lymphocytes (that produce IL-17, IL-22, and IL-6).

An example of a subgroup of immune cells that keep the immune system in check are regulatory T cells (Tregs). Tregs produce antiinflammatory cytokines, such as IL-10, transforming growth factor (TGF)-β, and IL-35. Tissues in both the anterior and the posterior segments of the human eye constitutively express TGF-β. The concentration of TGF-β in the aqueous humor is sufficient to inhibit T-cell activation and proliferation in a variety of assays.[3] As might be anticipated, significantly lower levels of this activated cytokine are measured in the aqueous humor of patients with a variety of uveitis syndromes compared with levels present in normal eyes. Another antiinflammatory cytokine, IL-27, is produced constitutively by resident cells in the retina, including retinal ganglion and photoreceptor cells.[4] IL-27 is upregulated by IFN-γ and suppresses Th17 cells.[4]

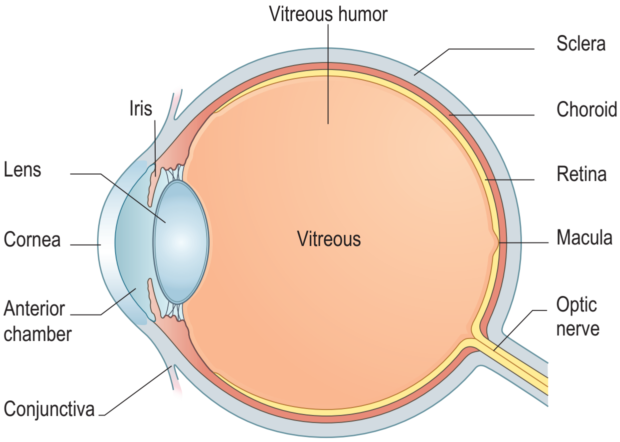


FIG 74.1 Schematic representation of the eye.

Furthermore, retinal pigment epithelial (RPE) cells are thought to play an important role in inducing Tregs through the production of retinoic acid in the presence of TGF-β.

Several cytokines, such as IFN-γ or IL-17, can be either protective or inflammatory, depending on the context in which they are produced. For example, although IL-17 is recognized as an inflammatory cytokine contributing to the pathogenesis of uveitis, one study showed that recombinant human IL-17 given systemically protected against inflammation in two different animal models of experimental uveitis.[5]

The inflammatory cytokine IL-1 has been implicated in the pathogenesis of various ocular inflammatory diseases. IL-1 receptor antagonist (IL-1Ra) is expressed in normal cornea[6] and retinal pigment epithelium, and this implies that this tissue contains a control mechanism for responses mediated by IL-1. Interestingly, topical application of IL-1Ra promotes experimental corneal allograft survival. Furthermore, the aqueous humor contains neuropeptides, including α-melanocyte stimulating hormone, vasoactive intestinal peptide, substance P, and calcitonin gene–related peptide, which inhibit T cells and macrophages.[7] Complement is active at low levels in the healthy eye, regulated by complement regulatory proteins that are expressed both on intraocular cell membranes and within the intraocular fluid; this system may participate in the destruction of pathogens invading the eye. Interestingly, iC3b, generated because of this activation, appears to contribute to immune tolerance.[8]

ANTERIOR CHAMBER–ASSOCIATED IMMUNE DEVIATION

ACAID is "a stereotypic, systemic immune response to antigens placed in the anterior chamber (of the eye) in which delayed hypersensitivity is avoided and suppressed." The culmination of cellular events in ACAID leads to the production of two different Treg populations, $CD4^+$, $CD25^+$ Tregs that inhibit the afferent immune response, and $CD8^+$ Tregs that block the efferent phase characterized by previously sensitized CD4 T cells. The molecular events that are responsible for ACAID include the following: entry of an antigen into the eye stimulates production of TNF-α, and hence the upregulation of cell adhesion molecules; antigens processed by antigen-presenting cells (APCs) in the eye migrate to the spleen and present to B cells in the marginal zone in the context of a nonconventional MHC molecule; induction of $CD8^+$ ACAID Tregs in the presence of IL-10 from $CD4^+$, $CD25^+$ Tregs, and from γδ T cells; $CD8^+$ Tregs upregulate FasL, IL-10, and TGF-β to promote FasL-mediated killing of T cells that recognize the ocular antigen; TGF-β also induces ocular APCs to secrete IL-10 during antigen processing. Apoptotic T cells from the eye pass in the bloodstream to the spleen, where they are phagocytosed and induce activation of Th2-type CD4 T cells. The Th2-type cells control Th1 function by secreting various immunomodulatory cytokines. T-cell receptor α-chain fragments from apoptotic cells are presented in the class I pathway. This event generates CD8 natural killer (NK) cells, which are capable of deleting the CD4 T cells that would otherwise mediate a delayed-type hypersensitivity response. One study demonstrated that IFN-γ appears to be important in CD8 Treg suppressive function.

Invariant natural killer T (iNKT) cells, which interact with lipid antigens presented in the context of an MHC-like molecule, Cd1d, may play a role in the induction of ACAID in the thymus and the spleen. In fact, one study showed that activation of iNKT cells could suppress experimental autoimmune uveitis.

IMMOBILITY OF DENDRITIC CELLS WITHIN THE ANTERIOR CHAMBER

The vast majority of cells that phagocytose foreign antigen within the iris fail to migrate after antigen uptake. This is clearly demonstrable by intravital microscopy and correlates with the failure of these antigen-bearing cells to migrate to the local lymph nodes.[9] The inability to migrate is consistent with the known lack of lymphatics within the eye and must mean that soluble antigen injected into the anterior chamber is not presented in the regional lymph node in a manner that is comparable with what follows antigen exposure in other peripheral tissues.

Role of the Commensal Microbiota in Ocular Immunity

Although the eye is thought to be relatively devoid of microorganisms, it is now known that there is a small number of normally residing ocular surface bacteria and viruses.[10,11] Perturbations in the ocular surface microbiota can result in infection and/or manifest as sterile inflammation. For instance, the presence of torque teno virus was very strongly associated with sterile endophthalmitis (a condition in which there is severe intraocular inflammation usually associated with infection) and was also found in some cases of culture-positive bacterial endophthalmitis.[12] Whether torque teno virus was the causal factor, merely an association, or indirectly causal via viral immune-modifying effects is unknown. Deep DNA sequencing may uncover previously unknown causes of uveitis.

UVEITIS

Uveitis as a Diagnostic Entity

Uveitis is a spectrum of different diseases that includes infections and immune-mediated diseases (Table 74.1). It is the third leading preventable cause of blindness worldwide. Inflammatory disorders of the retina (retinitis) and sclera (scleritis) frequently involve the adjacent uvea. Mechanisms contributing to uveitis include an immune response to a sequestered self-antigen, molecular mimicry, immune complex deposition, or a toxin.

The differential diagnosis of uveitis is facilitated by identifying characteristic clinical features. Uveitis can be classified by location[13]: anterior (iritis, iridocyclitis), intermediate (pars planitis, vitritis), or posterior (retinitis, choroiditis, retinochoroiditis, chorioretinitis). Some forms of uveitis involve all portions of

TABLE 74.1 Diagnostic Categories of Uveitis

Diagnostic Group		Diagnosis
Infectious causes	Viral	Herpes simplex, Herpes zoster
		Cytomegalovirus (CMV) infection
		Chikungunya
		Human T-lymphotropic virus 1 (HTLV-1) infection
		Mumps
		West Nile virus infection
		Ebola
		Zika virus infection
	Bacterial or spirochetal	Atypical *Mycobacterium* infection
		Bacterial endocarditis
		Bartonella infection
		Brucellosis
		Leprosy
		Leptospirosis
		Lyme disease
		Propionibacterium infection
		Tuberculosis
		Leprosy
		Syphilis
		Whipple disease
	Parasitical (protozoan or helminthic)	*Acanthamoeba* infection
		Cysticercosis
		Onchocerciasis
		Pneumocystis jiroveci infection
		Toxoplasmosis
		Toxocariasis
	Fungal	Histoplasmosis
		Coccidioidomycosis
		Candidiasis
		Aspergillosis
		Sporotrichosis
		Blastomycosis
		Cryptococcosis
Immune-mediated		Ankylosing spondylitis
		Behçet disease
		Crohn disease
		Drug or hypersensitivity reaction (*e.g.,* to rifabutin or cidofovir)
		Familial granulomatous synovitis with uveitis
		Interstitial nephritis
		Juvenile idiopathic arthritis
		Kawasaki disease
		Multiple sclerosis
		Psoriatic arthritis
		Reactive arthritis
		Relapsing polychondritis
		Rheumatic fever
		Sarcoidosis
		Scleritis
		Sjögren syndrome
		Sweet syndrome
		Systemic lupus erythematosus
		Ulcerative colitis
		Vasculitis
		Vogt-Koyanagi-Harada syndrome

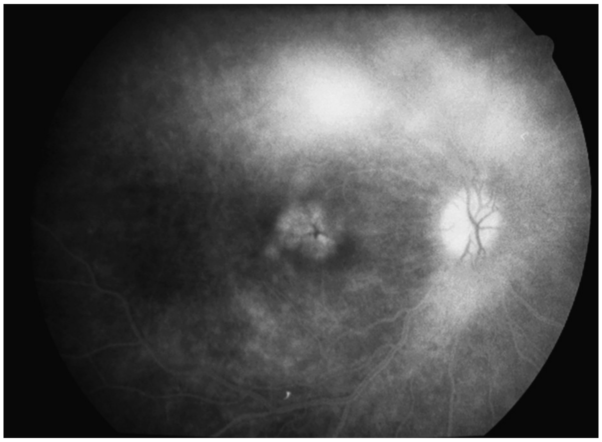


FIG 74.2 Fundus Fluorescein Angiogram Showing Cystoid Macular Edema. Fluorescein, which appears as a white stain, should be absent from the center of this photograph because the macular area is avascular.

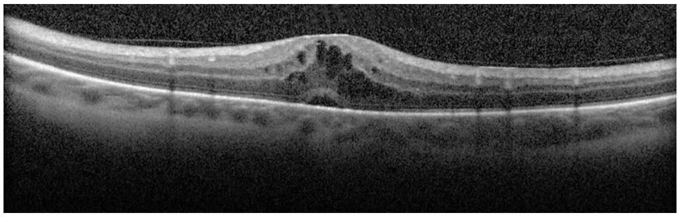


FIG 74.3 Spectral Domain Optical Coherence Tomography (SD-OCT) of Cystoid Macular Edema caused by Panuveitis. SD-OCT allows for precise quantification of macular thickening and photoreceptor damage. Note the intraretinal fluid–filled cysts resulting in distortion of the foveal contour. This patient had 20/100 visual acuity.

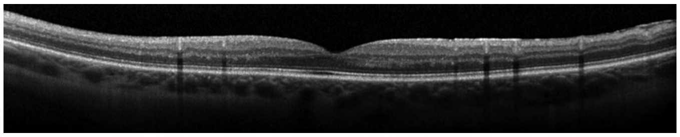


FIG 74.4 Spectral Domain Optical Coherence Tomography (SD-OCT) Image of a Normal Eye. This is the same macula as seen in Fig. 74.3, after an intravitreal steroid implant was placed. In this SD-OCT image, the macula shows a normal dimple or indentation at the normal fovea.

the uveal tract (panuveitis). Uveitis can be classified by course (self-limiting, chronic, or recurrent); by onset (sudden, insidious); by symmetry (unilateral, bilateral); by associated complications, such as glaucoma, cystoid macular edema (Figs. 74.2–74.4), synechiae (*e.g.,* adhesion of the iris to the lens) (Fig. 74.5), retinal detachment, or band keratopathy (the deposition of calcium in the epithelium of the cornea) (Fig. 74.6); and by the appearance of inflammatory keratic precipitates on the endothelium of the cornea within the eye (granulomatous, nongranulomatous) (Fig. 74.7). Granulomatous diseases with large cellular concretions on the cornea or nodules within the iris include tuberculosis, syphilis, sarcoidosis, Vogt-Koyanagi-Harada disease, and sympathetic ophthalmia. The group of nongranulomatous diseases includes ankylosing spondylitis, reactive arthritis, and juvenile idiopathic arthritis. Table 74.2 shows how these parameters contribute to the differential diagnosis.

Additionally, ethnic and geographical considerations factor into the differential diagnosis. For example, sarcoidosis, Behçet syndrome, and Vogt-Koyanagi-Harada disease have strong ethnic predispositions, whereas certain infections, such as cytomegalovirus (CMV) in association with acquired immunodeficiency syndrome

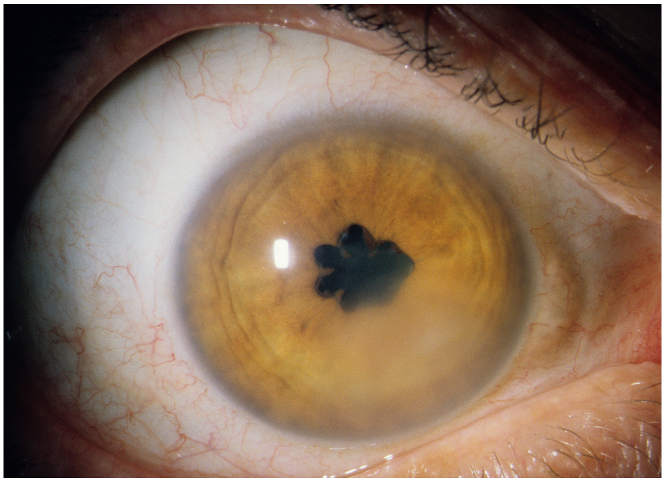


FIG 74.5 **Posterior Synechiae.** The cloverleaf appearance of the pupil is caused by the adherence of the iris to the lens capsule.

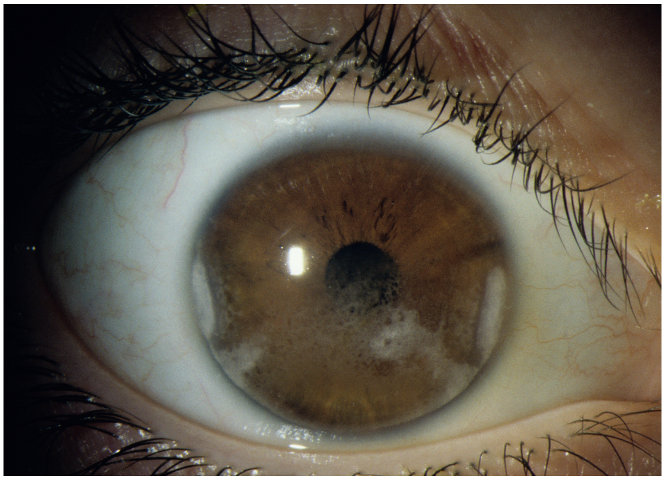


FIG 74.6 **Band Keratopathy.** Calcium deposition in the corneal endothelium complicates the iridocyclitis in a patient with juvenile idiopathic arthritis.

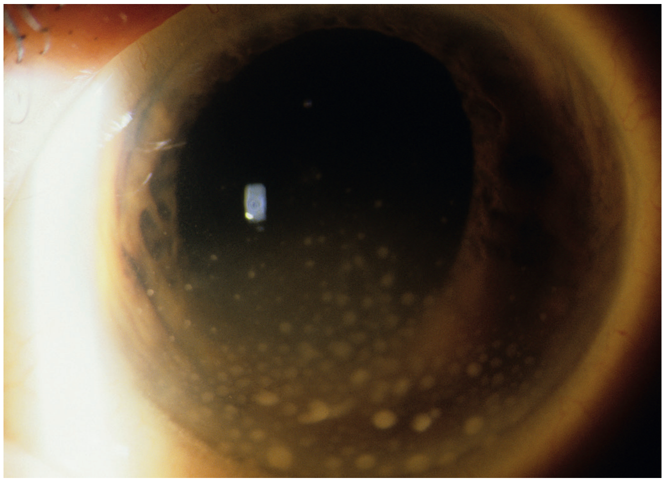


FIG 74.7 **Granulomatous Keratic Precipitates.** Round white dots are scattered in a triangular shape. The dots represent concretions of leukocytes adherent to the corneal endothelium.

TABLE 74.2 Characteristic Features of Common Forms of Uveitis

Parameter	
Location	Anterior: ankylosing spondylitis, reactive arthritis, juvenile idiopathic arthritis Intermediate: Pars planitis Posterior: Vogt-Koyanagi-Harada syndrome (panuveitis most common)
Onset	Sudden: ankylosing spondylitis, reactive arthritis Insidious: pars planitis, juvenile idiopathic arthritis
Symmetry	Unilateral: ankylosing spondylitis, toxoplasmosis, herpetic Bilateral: pars planitis, lymphoma, juvenile idiopathic arthritis
Course	Self-limiting: toxoplasmosis Recurrent: Behçet disease, ankylosing spondylitis Chronic: pars planitis

TABLE 74.3 Likelihood of Developing Uveitis in Association With a Specific Disease

Diagnosis	Likelihood (%)
Ankylosing spondylitis	30
Sarcoidosis	25–50
Behçet disease	80
Inflammatory bowel disease	1–9
Psoriatic arthritis	7
Juvenile idiopathic arthritis	<2–53
Multiple sclerosis	<1–2

(AIDS), leprosy, or onchocerciasis, vary in prevalence depending on geographical area. Together with a medical history, sex, and age, these findings help to narrow the differential diagnosis of uveitis. For example, a 22-year-old man with low back pain and a red painful eye caused by episodic, unilateral, sudden-onset anterior uveitis is highly likely to have a spondyloarthropathy (Chapter 57). A 6-year-old girl with no ocular complaints but bilateral band keratopathy and leukocytes in the anterior chamber is likely to suffer from pauciarticular juvenile idiopathic arthritis (Chapter 53).

The most obvious sign of uveal inflammation is the presence of leukocytes in the anterior chamber or the vitreous humor of the eye. Most patients with anterior uveitis will experience pain, redness, photophobia, miosis, and a variable degree of visual loss. In contrast, many forms of uveitis that affect the posterior segment will cause no redness, pain, or change in pupil size. Instead, ocular disturbances may vary from normal vision to seeing floaters, flashing lights, or photopsia, blurred vision, or blindness.

Epidemiology

The incidence and prevalence of the different types of uveitis varies throughout the world, depending on numerous factors: human leukocyte antigen (HLA) distribution, prevalence of different infectious diseases, and the methodology employed to evaluate and classify uveitis (Table 74.3). Gritz and Wong[14] reported an incidence of 52.4 per 100 000 per year and a prevalence of 115.3/100 000. Anterior uveitis is much more common than intermediate or posterior uveitis.

Prevention and Patient Management

Uveitis can be the first sign of systemic disease (*e.g.*, sarcoidosis, syphilis, or central nervous system [CNS] lymphoma). It is essential to identify the cause of uveitis as accurately as possible. A targeted approach is preferable, with a limited workup guided by the type and severity of uveitis and the presence of systemic findings. The history is of great importance. The minimal workup for uveitis of unknown etiology requires an extensive and careful review of systems, a syphilis serology, and a chest X-ray.[15]

Most practitioners use a stepladder approach to treat uveitis. The first step is corticosteroids for all patients with noninfectious uveitis (topical, especially for anterior uveitis, regional injection, or systemic administration). A surgically implantable intravitreal device to release fluocinolone for approximately 2.5 years is an option that is limited by cost and complications, which can include cataract, glaucoma, and scleral thinning. New delivery approaches include injectable corticosteroids complexed with polymers to prolong the beneficial effects; iontophoresis to deliver medication across the sclera; and small molecules, which seem to have a much broader biodistribution within the eye compared with topical prednisolone. Systemic immunosuppressive therapy should be used in cases of inadequate benefit from corticosteroids or unacceptable steroid side effects. Immunosuppression is usually reserved for bilateral forms of uveitis with a severity sufficient to alter activities of daily living and is frequently chosen for eye involvement with Behçet disease. Immunosuppression may be contraindicated by active infection in the eye or elsewhere.

The role for therapy with biologics to treat uveitis is being assessed in randomized controlled trials. Antibodies to IL-1β or to anti–IL-17A have proven disappointing in the treatment of uveitis.[16-20] However, adalimumab (an anti-TNF-α biologic) was shown to be effective in a randomized clinical trial for active uveitis[17] and as a corticosteroid sparing agent in inactive uveitis.[18] In an open label trial, infliximab has proven effective but had frequent severe toxicities for this indication.[19] The authors frequently prescribe infliximab to treat Behçet disease (see below). Biologic therapy for uveitis might also be guided by the underlying disease, such as Crohn disease, juvenile idiopathic arthritis, or psoriatic arthritis, all of which are particularly responsive to biologics.

Overview of the Most Common Uveitis Diagnoses

Ankylosing Spondylitis–Associated Uveitis

Ankylosing spondylitis (AS) is the most common systemic diagnosis associated with uveitis in Western nations. A genetic predisposition with the HLA-B27 allele is found in 88–96% of patients. About 30–50% of patients with AS develop anterior uveitis transiently during their lifetime. Approximately 50% of patients with sudden-onset anterior uveitis are HLA-B27 positive.[20] In some studies, 80–90% of individuals with HLA-B27–associated iritis have spondyloarthritis. AS should be considered in any patient with sudden-onset, primarily anterior, unilateral uveitis associated with redness and pain. Recurrent disease, fibrin in the anterior chamber, posterior synechiae, relatively brief inflammation that resolves in <2 months, and lowered intraocular pressure are additional hallmarks that help to distinguish HLA-B27–associated iritis from other categories.[21] Inflammation is sometimes so severe that a hypopyon (Fig. 74.8) develops. The posterior segment of the eye is not commonly involved in this disease except for the development of a vision-reducing cystoid macular edema (see Figs. 74.2 and 74.4). HLA-B27–associated anterior uveitis usually responds to topical corticosteroids.

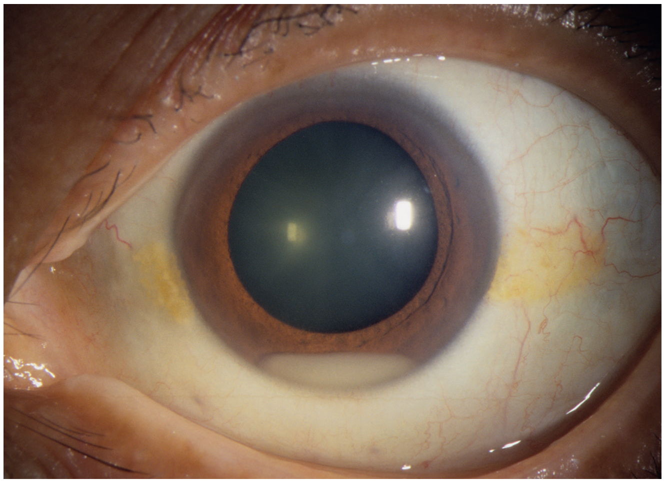


FIG 74.8 Hypopyon. Intense inflammation has resulted in a creamy exudate (pus), seen as whitening over the inferior portion of the anterior chamber. In this example, hypopyon was secondary to rifabutin, not ankylosing spondylitis.

In frequently recurrent cases, therapy with sulfasalazine[22] or TNF inhibitors is effective in reducing the frequency of the attacks.[23]

Several studies have emphasized that spondyloarthritis may be overlooked in patients with acute anterior uveitis (AAU). A study from Ireland evaluated patients with AAU in the emergency room. About 40% had spondyloarthritis that had not been recognized.[24] A study from Spain included 798 patients who had either HLA-B27–associated AAU or recurrent, HLA-B27–associated anterior uveitis. As in the Irish study, all patients with known spondyloarthritis were excluded. The study utilized the Assessment of Ankylosing Spondylitis (ASAS) criteria.[25] Fifty percent of this large group had axial spondyloarthritis and close to an additional 20% had peripheral spondyloarthritis.[26]

Inflammatory Bowel Disease–Associated Uveitis

Ocular manifestations of inflammatory bowel disease (IBD) (Crohn disease and ulcerative colitis; Chapter 75) include anterior uveitis, episcleritis, and scleritis (1–9% of all cases). Less common ocular inflammatory manifestations include peripheral keratitis, intermediate uveitis, optic neuritis, and retinal vasculitis.[27] Twenty percent of patients with IBD may have sacroiliitis; 60% of them are HLA-B27 positive. The anterior uveitis in patients with IBD is often similar to that seen with the HLA-B27 spectrum of joint diseases: sudden in onset, unilateral, and self-limiting. Like sarcoidosis, it may result in intraocular inflammation that is insidious in onset, bilateral, posterior to the lens, chronic, or associated with a retinal vasculitis.[28]

Crohn disease is a complex genetic disease—that is, multiple genes contribute to susceptibility. One of the best characterized of these genes is *NOD2.* This gene functions as an intracellular Toll-like receptor (TLR) that recognizes muramyl dipeptide from bacteria. The *NOD2* gene is also implicated in autoinflammatory diseases. A mutation in the portion of the *NOD2* gene known as the *nucleotide-binding domain* results in a rare autosomal dominant form of uveitis known as *Blau syndrome.*

Psoriatic Arthritis–Associated Uveitis

According to one study in the UK, 1–2% of the population have psoriasis (Chapter 64); roughly 10% of patients with psoriasis have arthritis[29]; and 7% of patients with psoriasis and arthritis have uveitis. The diagnosis of psoriatic arthritis is made clinically with the typical findings of cutaneous changes (erythematous, hyperkeratotic rash) and joint disease. Patients with sacroiliac disease and psoriatic arthritis are especially likely to develop uveitis.

Sarcoidosis-Associated Uveitis

Ocular disease is the initial manifestation of sarcoidosis (Chapter 73) in 20% of patients. The systemic illness begins as ocular inflammation almost as often as it initially presents as pulmonary disease. Of the patients with sarcoidosis, 25–50% exhibit ocular inflammation during the course of the illness. In most series, sarcoidosis accounts for 3–10% of all patients with uveitis, second to spondyloarthropathy as the most common systemic illness associated with uveitis in North America and Europe. The diagnosis of sarcoidosis should always be considered in any patient presenting with a uveitis of unknown etiology. Sarcoidosis occurs in the United States 10 times more frequently among African Americans than among whites. Because it can affect almost every part of the eye, symptoms of ocular sarcoidosis vary widely. Most commonly, patients have a bilateral chronic granulomatous anterior uveitis.[30] The disease often regresses clinically, with two-thirds of patients becoming symptom free after 3 years. Typical findings include "mutton-fat" keratic precipitates of the cornea (see Fig. 74.7), Koeppe and Busacca iris nodules, posterior synechiae (see Fig. 74.5), and white clumps of cells ("snowballs") in the inferior vitreous humor. Fundus findings include multifocal perivenular sheathing, typical punched out chorioretinal lesions, and/or optic nerve head granulomata. Diffuse retinal vasculitis, predominantly of the retinal venules, can also be seen. A chest X-ray (or computed tomography [CT] scan of the chest[31]) should be included in every basic workup of patients with uveitis of unknown cause. The diagnosis is supported by a biopsy that demonstrates the noncaseating granuloma in the absence of an infection or beryllium exposure. Possible biopsy sites include lymph nodes, lung, skin, oral mucosa, and conjunctiva. Serological abnormalities are not sufficiently unique to establish a diagnosis. In general, the combination of uveitis and symmetrical hilar adenopathy is considered presumed sarcoidosis and can be treated like sarcoidosis. Topical, periocular, and systemic corticosteroids are the mainstay of therapy. Steroid-sparing therapy can be required for more severe or chronic cases affecting the posterior segment, with the antimetabolites as first-line steroid-sparing agents, and TNF-α inhibitors as alternative treatment.[32,33]

Behçet Uveitis

Behçet syndrome is a generalized occlusive vasculitis of unknown cause (Chapter 58). It accounts for about 2.5% of patients at a North American referral uveitis clinic and has a strong ethnic relationship. It is most common in Japan (20% of patients with uveitis) and prevalent in Middle Eastern countries along the Silk Road.

The classic complex includes aphthous stomatitis, genital ulceration, and iritis, sometimes with hypopyon (30% of patients) and skin lesions. Ocular manifestations can range from a hypopyon anterior uveitis to intermediate uveitis, retinitis, and occlusive retinal vasculitis (Fig. 74.9). Extraocular manifestations include arthritis, erythema nodosum, diarrhea (mimicking IBD), and sterile meningitis. There is a strong tendency for symptoms to remit and exacerbate spontaneously. HLA-B51 is more commonly found among patients with Behçet syndrome. Behçet syndrome may lead to blindness if ischemic optic neuropathy and retinopathy are not adequately treated. Patients are often inadequately managed by corticosteroids alone: A regimen with immunosuppressive drugs (cyclosporine or azathioprine) has been proved useful. Some investigators advocate the use of IFN-α.[34] Infliximab has demonstrated dramatic benefit in studies reported from Europe and Japan.[35] It is now approved for this indication in Japan.

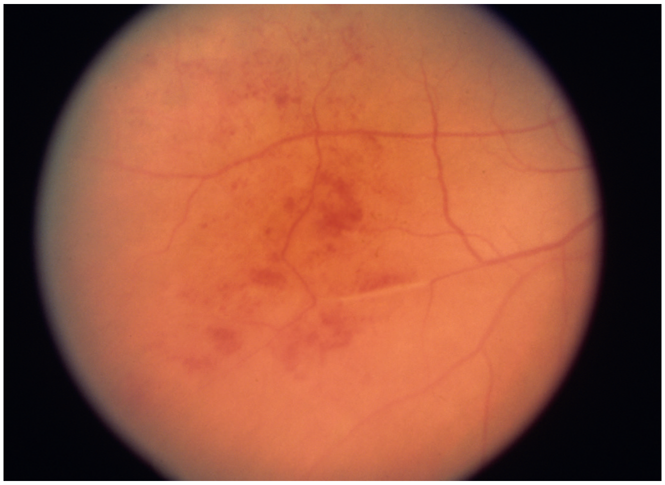


FIG 74.9 Fundus Photograph Showing Retinal Vasculitis. The scattered hemorrhages are secondary to the vasculitis. The patient has Behçet syndrome.

Multiple Sclerosis–Associated Uveitis

Although retrobulbar optic neuritis is the ocular inflammation most associated with multiple sclerosis (MS), intermediate uveitis and bilateral granulomatous anterior uveitis are well described in association with this disease[36] (Chapter 66). Bilateral granulomatous uveitis may indicate a worse prognosis for the neurological disease. The HLA-DR2 antigen, strongly associated with MS, is also associated with uveitis.

Juvenile Idiopathic Arthritis–Associated Uveitis

Patients with juvenile idiopathic arthritis (JIA; Chapter 53) need special consideration as the disease often does not present with a "red eye": The uveitis often occurs as a smoldering, silent inflammation in a quiet, white eye. In 20% of all JIA cases, the joint involvement is pauciarticular (≤5 joints affected). Uveitis occurs far more commonly in this pauciarticular subset, and these patients require routine ophthalmic screening every 3–4 months. Girls are affected four times more often than boys, although boys are more likely to develop complications.[37] Eighty percent of uveitis cases are positive for antinuclear antibodies (ANAs) and negative for rheumatoid factor. The onset of ocular disease is usually within 5 years of arthritis but may precede arthritis as well. Biomicroscopically, bilateral fine keratic precipitates, band keratopathy (see Fig. 74.6), flare and cells in the anterior chamber, posterior synechiae, glaucoma, and cataract formation can be found. Topical and periocular corticosteroid treatment with a slow taper is the mainstay of therapy. Systemic therapy with antimetabolites, such as methotrexate or mycophenolate, can be used to spare corticosteroids and may be indicated to reduce the morbidity of chronic corticosteroid use. A TNF inhibitor, such as adalimumab, is useful in patients with refractory disease.[38] Cataract surgery is a special challenge and may require intense perioperative immunosuppression.

KEY CONCEPT

Unilateral, recurrent, acute anterior uveitis is one of the most common forms of uveitis; it is frequently associated with spondyloarthritis.

CLINICAL PEARL

It is important to rule out infectious causes of uveitis before initiating immunosuppression for uveitis.

Advancements and New Treatments for Uveitis

As molecular diagnostics, such as next-generation deep sequencing techniques and bioinformatics, are becoming refined and more broadly available, additional etiological agents for uveitis may be discovered. Additionally, with an increased understanding of the cellular pathways involved in various uveitic entities, new targets for therapy of uveitis may emerge. For instance, targeting IL-6 with a biologic agent versus targeting lymphocyte trafficking to the eye via an antibody against adhesion molecules may represent emerging therapies. Additional local drug delivery methods may improve ocular efficacy while lowering systemic side effects.

ON THE HORIZON

Newer molecular diagnostic techniques will improve identification of etiological agents in uveitis, and potentially direct novel treatments both locally and systemically for uveitis.

SCLERITIS

Description and Natural History

Scleritis is a relatively uncommon inflammation affecting the *sclera*, the white tunic that encases the eye, and the overlying episcleral tissues. Scleritis is usually presumed to represent a vasculitis of the deep episcleral vessels that overly the sclera. At least one-third of patients with scleritis will have clinical evidence for a systemic vasculitis elsewhere in the body.[39] The most common systemic diseases are granulomatosis with polyangiitis (GPA; Chapter 58), previously known as *Wegener granulomatosis* (often in a limited form), or rheumatoid arthritis (RA; Chapter 52). The subset of patients with RA who develop scleritis are especially likely to develop rheumatoid nodules, high-titer rheumatoid factor, and pleuropericarditis, as well as small vessel vasculitis of the lower extremities. Other systemic diseases commonly associated with scleritis are listed in Table 74.4. The epidemiology of scleritis has not been adequately reported. However, RA accounts for about one-third of all cases of scleritis. Because only 1% of patients with RA develop scleritis, and about 1% of the population has RA, one can estimate that the prevalence of scleritis is 0.01 × 0.01 × 3, or 3/10 000.

Scleritis is frequently divided into five subsets based on the clinical presentation: diffuse anterior (Fig. 74.10); nodular anterior; necrotizing; scleromalacia perforans (Fig. 74.11); and posterior scleritis.[40] Scleromalacia perforans is also known as *necrotizing scleritis without inflammation.* The clinical presentation will vary according to the type, but scleritis is usually a very painful and persistent condition. Even with treatment, median duration reaches 5 years.

TABLE 74.4 Systemic Immune-Mediated Diseases Associated With Scleritis

Rheumatoid arthritis
Granulomatosis with polyangiitis
Inflammatory bowel disease
Polyarteritis nodosa
Temporal arteritis/giant cell arteritis
Systemic lupus erythematosus
Ankylosing spondylitis
Relapsing polychondritis

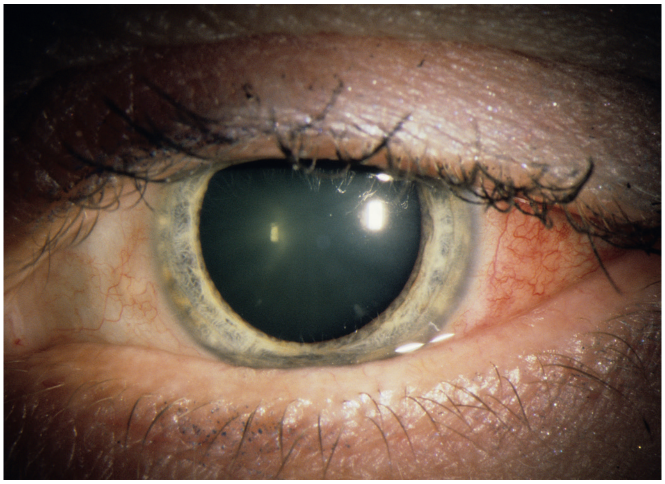


FIG 74.10 Diffuse Anterior Scleritis. The red patch of vessels temporally is secondary to scleritis.

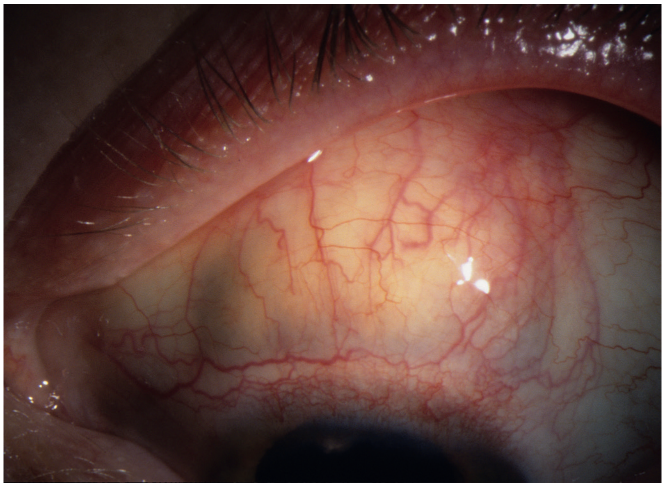


FIG 74.11 Scleromalacia Perforans. Blue sclera is visible secondary to inflammation, resembling a rheumatoid nodule in the eye.

Genetic and Environmental Factors

Genetic factors affecting scleritis have not been adequately studied. Genes that affect the associated systemic diseases, such as RA, GPA, and polyarteritis nodosa (PAN), presumably also influence susceptibility to scleritis. Environmental triggers for scleritis have not been defined, except that some cases of scleritis can be a manifestation of infection, such as herpes zoster ophthalmicus, syphilis, Lyme disease, tuberculosis, or other bacterial and fungal infections.

Immunology and Pathology

Biopsy of the sclera entails some risk. Accordingly, many of the published histological observations are based on end-stage or extremely severe disease. In one pathology study of 55 examples of necrotizing scleritis, the histology was divided into four types: zonal necrotizing granulomatous scleritis; nonzonal diffuse scleritis; necrotizing scleritis with microabscesses; and sarcoidal granulomatous scleritis.[41] Eighty-five percent of patients with zonal necrotizing granulomatous pathology had a systemic disease, usually rheumatoid arthritis. The pathology associated with RA was not distinct from other systemic diseases, such as GPA. None of 19 patients with nonzonal diffuse scleral inflammation had a systemic disease. Just over half the patients with microabscesses had an identifiable infection as the cause of scleritis. Only one patient had "sarcoidal granulomatous inflammation," and that patient had sarcoidosis.

Prevention and Management

Topical medications are only slightly beneficial for most patients with scleritis. Topical nonsteroidal antiinflammatory drugs (NSAIDs), such as ketorolac, have not been proven efficacious. A topical corticosteroid can sometimes help in symptomatic control but also has risk, especially the promotion of cataract formation and elevation of intraocular pressure. The role for topical cyclosporine has not been adequately evaluated, but the authors' clinical experience has not been favorable.

Although most forms of scleritis represent a vasculitis, oral NSAIDs can be immensely beneficial. Although not all patients derive adequate control from oral NSAIDs, a subset will benefit sufficiently such that no other medication is required. We speculate that the subset responsive to an NSAID might have a pathogenesis other than vasculitis, but this hypothesis remains unproven.

Many patients with scleritis will not benefit adequately from an NSAID and will require immunosuppression. In general, for those with an associated systemic disease, pharmacologic control of that disease will control the scleritis. For those without a systemic illness, oral corticosteroids are an accepted initial approach to treatment. For example, therapy might begin with a dose of prednisone of 1 mg/kg bodyweight per day. The use of calcium and vitamin D, as well as other measures to preserve bone mineral density, should be considered for any patient who will be receiving corticosteroid on a chronic basis. If prednisone is not adequate for disease control or if the medication is poorly tolerated and cannot be safely tapered to a modest dose, it is reasonable to add a steroid-sparing medication. There is considerable interest in the use of biologic agents as corticosteroid-sparing agents in scleritis. Multiple recently published case reports and small case series have described the successful use of drugs, including TNF blockers or rituximab, in some patients with resistant disease.[42]

Locally injected corticosteroids may be effective for uveitis, but they are contraindicated for patients with necrotizing scleritis. The corticosteroid might enhance the likelihood that the necrotic sclera will perforate. However, for nonnecrotizing anterior scleral inflammation, there is accumulating evidence that subconjunctivally injected corticosteroid is safe and may be highly effective.[43] Scleral inflammation limited to the posterior sclera can probably be safely treated with a periocular corticosteroid injection.

In choosing to treat with oral corticosteroid or an immunosuppressive, the clinician will, of course, weigh the risk/benefit ratio. The rationale for aggressive therapy is much stronger if the disease is bilateral and is affecting activities of daily living, because of pain or reduction in acuity.

CLINICAL PEARL

Although noninfectious scleritis may be considered a vasculitis of the deep episcleral vessels, oral nonsteroidal antiinflammatory drugs (NSAIDs) are effective in its treatment, and when the disease is refractory, oral corticosteroids, antimetabolites, or biologics can be quite effective.

Evaluation and Differential Diagnosis

Patients with scleritis generally have characteristic findings that allow an accurate diagnosis; scleritis can usually be readily distinguished from other causes of a red, painful eye. The most difficult distinction is between scleritis and episcleritis. The differences between scleritis and episcleritis are highlighted in Table 74.5.

The scleritis associated with a systemic vasculitis is often more severe than scleritis not associated with a systemic illness.[44] The scleritis associated with GPA may be particularly destructive and refractory to therapy. Scleritis in association with RA is often a poor prognostic sign. Before the routine use of disease-modifying antirheumatic drugs, the small subset of patients with scleritis in association with RA had a shortened life expectancy compared with those without evidence for ocular involvement.

Posterior scleritis can be an extremely difficult disease to diagnose. If the anterior sclera is uninvolved, no redness is present. Pain with posterior scleritis is much more variable. The examination may show elevation of the adjacent retina and choroid. The diagnosis can be confirmed with an ultrasound examination or, less commonly, with CT of the orbit demonstrating thickening of the sclera.

Pitfalls and Controversy

Some experts believe that scleritis can sometimes be a *forme fruste* of GPA. In patients with scleritis, a positive antineutrophil cytoplasmic antibody (ANCA) test and no other evidence for GPA has been described.[45]

With the exception of the ANCA test, laboratory studies are generally selected according to the history and general physical examination. For example, although RA is associated with scleritis, the patients almost always have long-standing severe RA. A rheumatoid factor and/or anticyclic citrullinated peptide antibody titer would not be appropriate tests if no joint disease were

TABLE 74.5 **Contrasting Features of Scleritis and Episcleritis**

	Scleritis	Episcleritis
Pain	Prominent	Minimal
Duration	Years	Days to months
Association with systemic disease	Frequent	Uncommon
Vessels blanch with topical vasoconstrictor	No	Yes
Associated ocular complications, including visual loss	Sometimes present	Rarely present
Vessel color	Violaceous	Light pink

present clinically. The findings from one study, however, do support screening patients with scleritis with serological tests to detect RA.[46] Similarly, systemic lupus erythematosus (SLE) can be associated with scleritis, but one would not diagnose SLE on the basis of a positive ANA result, scleritis, and the absence of other findings to suggest SLE. An erythrocyte sedimentation rate (ESR) can sometimes be helpful in ensuring that a systemic process is not present. Other laboratory tests are largely dictated by the medication that is chosen to treat the scleritis.

KERATITIS

The cornea is the anteriorly situated window to the eye. Normally, it should be clear. It can become opacified as a result of trauma, exposure to toxins, infection, dryness, calcium deposition, or genetic diseases, such as corneal dystrophies. The three main immune-mediated diseases that affect the cornea are peripheral ulcerative keratitis (PUK) (Fig. 74.12), Mooren ulcer, and Cogan syndrome.

Most patients with PUK have a systemic disease, most commonly RA, but can have SLE, GPA, or PAN. PUK is usually concomitant with scleritis. Treatment is similar to that for scleritis, with emphasis on controlling the underlying systemic illness and aggressive use of immunosuppression. A complication of uncontrolled PUK is corneal melt, which can lead to corneal perforation. This is usually treated with antimetalloproteinases/collagenases and surgically with tissue adhesives, conjunctival flaps, and possibly corneal transplantation, in addition to aggressive systemic immunosuppression, such as with rituximab. PUK is fortunately a very rare disease. PUK must also be distinguished from other causes of corneal thinning, such as Terrien marginal degeneration or senile marginal furrow, both of which are generally more benign and rarely lead to perforation.

Mooren ulcer is clinically very similar to PUK, except that there is no accompanying scleritis and no evidence for a systemic illness. The superior cornea is especially likely to be involved. Mooren ulcer is too rare for therapeutic options to be studied in randomized clinical trials. It may respond to oral immunosuppression, which is warranted if the severity threatens visual acuity because of impending corneal perforation. Some patients with Mooren ulcer have an autoantibody to calgranulin C. This antigen is expressed in the cornea, in neutrophils, and by filarial nematodes.

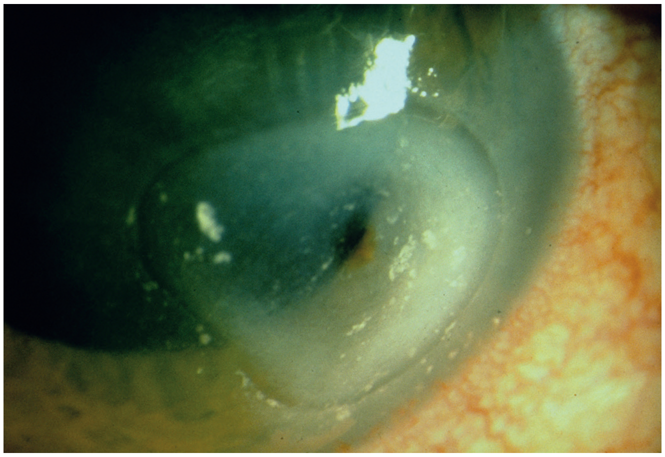


FIG 74.12 Corneal Melt. A pie-shaped wedge of the cornea is thinned. The eye is red secondary to an associated scleritis.

The cornea is divided into five layers: the epithelium facing the atmosphere, Bowman membrane, the stroma, Descemet membrane, and the endothelium abutting the aqueous humor. Opacification of the stroma is known as *interstitial keratitis.* Congenital syphilis and other infections, such as herpes simplex, are important causes of the latter. Cogan syndrome is defined as an autoimmune disease of the eighth nerve combined with interstitial keratitis, presumably on an autoimmune basis.[47] Some ophthalmologists also recognize autoimmune eighth-nerve disease associated with other forms of ocular inflammation, such as uveitis, as examples of Cogan syndrome. The pathogenesis of Cogan syndrome is presumed to be a vasculitis, and many patients will have evidence for vasculitis elsewhere in the body. The treatment usually requires aggressive use of systemic immunosuppressive medications.

CORNEAL TRANSPLANTATION AND TRANSPLANT REJECTION

Corneal transplantation involves the repair of a diseased cornea with healthy cadaver tissue. Since the 1970s, advances in microsurgical techniques and eye banking procedures have led to widespread acceptance of this procedure. Currently, more than 40 000 corneal transplantations are performed annually in the United States alone. The usual indication for a full-thickness corneal graft is poor vision from deep stromal scarring. The most common medical conditions leading to corneal transplantation are *keratoconus,* a condition in which myopic astigmatism develops as the cornea becomes progressively more conical in shape; corneal edema following intraocular surgery; and a failed previous corneal graft.

Although the eye is an immune-privileged site, and corneal transplants enjoy a 91% 1-year survival as shown by Kaplan-Meier survival analyses, only 62% of grafts are functional at 10 years. The most common cause of transplantation failure is immunologic rejection. Corneal allograft rejection rarely occurs within 2 weeks, and may occur as late as 20 years after surgery. Animal studies using monoclonal antibodies (mAbs) directed against different T-cell subsets indicate that CD4 T cells play a critical role in the rejection response. However, the mechanisms responsible for this process are still being studied.

Early recognition of a rejection episode is the most important factor in achieving survival of the corneal transplant. In its most florid form, the anterior eye is obviously inflamed, with intense conjunctival injection, a cellular anterior chamber reaction, and a Khodadoust line. This line, which is visible with the slit-lamp biomicroscope, is a classic sign of corneal graft rejection. It appears as a linear formation of inflammatory precipitates stretching across the corneal endothelium and represents a wave of lymphocytes marching across the cornea and destroying the endothelium in their path. As endothelial pump function is lost, the cornea becomes waterlogged and opaque. At an early stage of rejection, the patient may be asymptomatic, but later, ocular redness, photophobia, halos, and blurred vision are frequent complaints.

An intensive and extended course of topical corticosteroids is the mainstay of treatment for a rejecting corneal graft, and in severe cases intravenous and/or oral corticosteroids may also be administered. Patients considered at high risk of transplant rejection, such as those with corneal neovascularization or a history of other anterior segment inflammation, are often given perioperative systemic immunosuppression. The ideal prophylactic regimen has not been defined, although various combinations of prednisone, cyclosporine, and azathioprine may be used. Despite the critical influence of donor–recipient histocompatibility matching for solid organ transplant survival, the chance of corneal graft survival is not significantly improved by HLA matching.

Corneal transplantation has traditionally been accomplished by penetrating keratoplasty, in which all layers of the cornea are simultaneously replaced. Many centers are increasingly performing lamellar keratoplasty, in which layers of the cornea, such as the endothelium and Descemet membrane, are selectively transplanted.[48] Descemet membrane endothelial keratoplasty (DMEK) is an example of endothelial keratoplasty. Although technically more challenging, lamellar keratoplasty results in faster visual recovery and a reduced likelihood of transplant rejection.

CANCER-ASSOCIATED RETINOPATHY

Cancer-associated retinopathy (CAR) is a rare paraneoplastic syndrome that is most commonly associated with small cell carcinoma of the lung.[49] In addition, the disease has been documented in association with various tumors of the female reproductive tract, carcinoma of the breast, and neuroendocrine bronchial carcinoma. For 50% of patients, CAR is the presenting feature of their malignancy. Interestingly, there are now reports of retinopathy mimicking CAR but occurring in an apparently healthy individual. Melanoma-associated retinopathy is a related syndrome, occurring in patients with metastatic cutaneous melanoma.

Histopathological examinations of postmortem specimens taken from patients with CAR consistently demonstrate loss of inner and outer segments of the retinal photoreceptors. This destruction was initially attributed to the release of a hormone-like substance by malignant cells, but evidence has now accumulated in support of an autoimmune etiology. Affected individuals produce antibodies against one or more retinal photoreceptor antigens.[50] These antibodies induce experimental CAR when injected into laboratory animals. Although more than 15 antigens have been described in relation to CAR, the most common is the so-called CAR antigen.[50] This 23-kilodalton (kDa) protein has been identified as recoverin, a photoreceptor protein that participates in visual adaptation. Experimental work supports the hypothesis that a single mutational event simultaneously activates the recoverin gene and eliminates functional p53, a tumor suppressor protein. Consequently, there is development of a tumor that encodes CAR antigen and stimulates formation of antirecoverin antibody. Antirecoverin antibodies are capable of inducing photoreceptor apoptosis, leading to the characteristic loss of both rods and cones.

CAR generally occurs after the age of 60 years. Patients usually complain of decreased vision, although other symptoms may include transient visual obscurations, night blindness, scotomata, glare, and photosensitivity. Although visual acuity can be dramatically reduced, other clinical signs are often subtle. An afferent pupillary defect can be present if the retinopathy is asymmetrical. There may be mild iridocyclitis and/or anterior vitritis, narrowing of retinal arterioles, mottling of retinal pigment epithelium, and optic disc pallor. Visual field abnormalities occur. Electroretinography shows either reduced or completely flattened amplitudes. The disease must be differentiated from optic neuritis which, unlike CAR, typically occurs in younger persons, some of whom suffer from MS, and is painful. In contrast to CAR, optic neuritis typically has an abrupt onset. In patients with malignancy, ocular metastatic tumor spread and the effects of drugs, including chemotherapy agents, must be excluded before CAR is diagnosed.

The natural history of CAR is one of progressive visual loss, although this occurs over a variable period. Corticosteroid and other immunosuppressive drugs have been successfully used to treat the ocular disease.[51] It is suggested that serum antibody levels can be monitored to guide therapy. In some cases, advanced photoreceptor damage can be irreversible despite therapy.

IMMUNOLOGIC ETIOLOGIC FACTORS IN “NONIMMUNOLOGIC” OCULAR DISEASE

Age-related macular degeneration (ARMD) is the leading cause of irreversible visual loss in older persons in Western nations. A spectrum of pathologies can occur at the macula, ranging from relatively benign lipid deposits, called *drusen,* to the sometimes visually devastating retinal pigment epithelium atrophy or subretinal/choroidalneovascular membranes. The pathogenic stimulus for these age-related changes is unknown. Genetic polymorphisms of the complement regulatory protein, complement factor H, influences the susceptibility to ARMD, suggesting that this disease is related to chronic inflammation.[52] Subretinal scar tissue removed from patients with advanced, neovascular disease may contain immunoglobulins and complement components. The histopathology of ARMD includes inflammatory cells, including lymphocyte subsets, and local cell populations can express class II MHC antigens. Inhibition of complement activation is under study to treat or prevent ARMD.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Stuart PM, Griffith TS, Usui N, et al. CD95 ligand (FasL)-induced apoptosis is necessary for corneal allograft survival. J Clin Invest 1997;99:396–402.
2. Wang S, Boonman ZF, Li HC, et al. Role of TRAIL and IFN-gamma in CD4+ cell-dependent tumor rejection in the anterior chamber of the eye. J Immunol 2003;171:2789–96.
3. Cousins SW, McCabe MM, Danielpour D, et al. Identification of transforming growth factor- b as an immunosuppressive factor in aqueous humor. Invest Ophthalmol Vis Sci 1991;32:2201–11.
4. Amadi-Obi A, Yu CR, Liu X, et al. TH17 cells contribute to uveitis and scleritis and are expanded by IL-2 and inhibited by IL-27/STAT1. Nat Med 2007;13:711–18.
5. Ke Y, Liu K, Huang GQ, et al. Anti-inflammatory role of IL-17 in experimental autoimmune uveitis. J Immunol 2009;182:3183–90.
6. Kennedy MC, Rosenbaum JT, Brown J, et al. Novel production of interleukin-1 receptor antagonist peptides in normal human cornea. J Clin Invest 1995;95:82–8.
7. Taylor AW, Kaplan HJ. Ocular immune privilege in the year 2010: ocular immune privilege and uveitis. Ocul Immunol Inflamm 2010;18:488–92.
8. Sohn JH, Bora PS, Suk HJ, et al. Tolerance is dependent on complement C3 fragment iC3b binding to antigen-presenting cells. Nat Med 2003;9:206–12.
9. Dullforce PA, Garman KL, Seitz GW, et al. APCs in the anterior uveal tract do not migrate to draining lymph nodes. J Immunol 2004;172:6701–8.
10. Willcox MD. Characterization of the normal microbiota of the ocular surface. Exp Eye Res 2013;117:99–105.

11. Huang Y, Yang B, Li W. Defining the normal “core microbiome” of conjunctival microbial communities. Clin Microbiol Infect 2016;22:643.e7–12.
12. Lee AY, Akileswaran L, Tibbetts MD, et al. Identification of torque teno virus in culture-negative endophthalmitis by representational deep DNA sequencing. Ophthalmology 2015;122:524–30.
13. Jabs DA, Nussenblatt RB, Rosenbaum JT, Standardization of Uveitis Nomenclature Working G. Standardization of uveitis nomenclature for reporting clinical data. Results of the First International Workshop. Am J Ophthalmol 2005;140:509–16.
14. Gritz DC, Wong IG. Incidence and prevalence of uveitis in Northern California; the Northern California Epidemiology of Uveitis Study. Ophthalmology 2004;111:491–500, discussion 500.
15. Rosenbaum JT. An algorithm for the systemic evaluation of patients with uveitis: guidelines for the consultant. Semin Arthritis Rheum 1990;19:248–57.
16. Dick AD, Tugal-Tutkun I, Foster S, et al. Secukinumab in the treatment of noninfectious uveitis: results of three randomized, controlled clinical trials. Ophthalmology 2013;120:777–87.
17. Nguyen QD, Merrill PT, Jaffe GJ, et al. Adalimumab for prevention of uveitic flare in patients with inactive non-infectious uveitis controlled by corticosteroids (VISUAL II): a multicentre, double-masked, randomised, placebo-controlled phase 3 trial. Lancet 2016;388:1183–92.
18. Jaffe GJ, Dick AD, Brezin AP, et al. Adalimumab in Patients with Active Noninfectious Uveitis. N Engl J Med 2016;375:932–43.
19. Suhler EB, Smith JR, Giles TR, et al. Infliximab therapy for refractory uveitis: 2-year results of a prospective trial. Arch Ophthalmol 2009;127:819–22.
20. Brewerton DA, Caffrey M, Nicholls A, et al. Acute anterior uveitis and HL-A 27. Lancet 1973;302:994–6.
21. Rosenbaum JT. Acute anterior uveitis and spondyloarthropathies. Rheum Dis Clin North Am 1992;18:143–51.
22. Munoz-Fernandez S, Hidalgo V, Fernandez-Melon J, et al. Sulfasalazine reduces the number of flares of acute anterior uveitis over a one-year period. J Rheumatol 2003;30:1277–9.
23. Braun J, Baraliakos X, Listing J, et al. Decreased incidence of anterior uveitis in patients with ankylosing spondylitis treated with the anti-tumor necrosis factor agents infliximab and etanercept. Arthritis Rheum 2005;52:2447–51.
24. Haroon M, O’Rourke M, Ramasamy P, et al. A novel evidence-based detection of undiagnosed spondyloarthritis in patients presenting with acute anterior uveitis: the DUET (Dublin Uveitis Evaluation Tool). Ann Rheum Dis 2015;74:1990–5.
25. Rudwaleit M, van der Heijde D, Landewe R, et al. The Assessment of SpondyloArthritis International Society classification criteria for peripheral spondyloarthritis and for spondyloarthritis in general. Ann Rheum Dis 2011;70:25–31.
26. Juanola X, Loza Santamaria E, Cordero-Coma M, et al. Description and Prevalence of Spondyloarthritis in Patients with Anterior Uveitis: The SENTINEL Interdisciplinary Collaborative Project. Ophthalmology 2016;123:1632–6.
27. Salmon JF, Wright JP, Murray AD. Ocular inflammation in Crohn’s disease. Ophthalmology 1991;98:480–4.
28. Lyons JL, Rosenbaum JT. Uveitis associated with inflammatory bowel disease compared with uveitis associated with spondyloarthropathy. Arch Ophthalmol 1997;115:61–4.
29. Ibrahim G, Waxman R, Helliwell PS. The prevalence of psoriatic arthritis in people with psoriasis. Arthritis Rheum 2009;61:1373–8.
30. James DG. Ocular sarcoidosis. Ann N Y Acad Sci 1986;465:551–63.
31. Kosmorsky GS, Meisler DM, Rice TW, et al. Chest computed tomography and mediastinoscopy in the diagnosis of sarcoidosis-associated uveitis. Am J Ophthalmol 1998;126:132–4.
32. Liu D, Birnbaum AD. Update on sarcoidosis. Curr Opin Ophthalmol 2015;26:512–16.
33. Pasadhika S, Rosenbaum JT. Ocular Sarcoidosis. Clin Chest Med 2015;36:669–83.
34. Kotter I, Zierhut M, Eckstein AK, et al. Human recombinant interferon alfa-2a for the treatment of Behcet’s disease with sight threatening posterior or panuveitis. Br J Ophthalmol 2003;87:423–31.
35. Van Staa TP, Leufkens HGM, Abenhaim L, et al. Use of oral corticosteroids and risk of fractures. J Bone Miner Res 2000;15: 993–1000.
36. Messenger W, Hildebrandt L, Mackensen F, et al. Characterisation of uveitis in association with multiple sclerosis. Br J Ophthalmol 2015;99:205–9.
37. Edelsten C, Lee V, Bentley CR, et al. An evaluation of baseline risk factors predicting severity in juvenile idiopathic arthritis associated uveitis and other chronic anterior uveitis in early childhood. Br J Ophthalmol 2002;86:51–6.
38. Simonini G, Taddio A, Cattalini M, et al. Prevention of flare recurrences in childhood-refractory chronic uveitis: an open-label comparative study of adalimumab versus infliximab. Arthritis Care Res 2011;63:612–18.
39. Raiji VR, Palestine AG, Parver DL. Scleritis and systemic disease association in a community-based referral practice. Am J Ophthalmol 2009;148:946–50.
40. Watson PG, Hayreh SS. Scleritis and episcleritis. Br J Ophthalmol 1976;60:163–91.
41. Riono WP, Hidayat AA, Rao NA. Scleritis: a clinicopathologic study of 55 cases. Ophthalmology 1999;106:1328–33.
42. Suhler EB, Lim LL, Beardsley RM, et al. Rituximab therapy for refractory scleritis: results of a phase I/II dose-ranging, randomized, clinical trial. Ophthalmology 2014;121:1885–91.
43. Albini TA, Zamir E, Read RW, et al. Evaluation of subconjunctival triamcinolone for nonnecrotizing anterior scleritis. Ophthalmology 2005;112:1814–20.
44. Sainz de la Maza MF, Foster CS, Jabbur NS. Scleritis associated with rheumatoid arthritis and with other systemic immune mediated diseases. Ophthalmology 1994;101:1281–6.
45. Hoang LT, Lim LL, Vaillant B, et al. Antineutrophil cytoplasmic antibody-associated active scleritis. Arch Ophthalmol 2008;126:651–5.
46. Lin P, Bhullar SS, Tessler HH, et al. Immunologic markers as potential predictors of systemic autoimmune disease in patients with idiopathic scleritis. Am J Ophthalmol 2008;145:463–71.
47. Cogan DG. Syndrome of nonsyphilitic interstitial keratitis and vestibuloauditory symptoms. Arch Ophthalmol 1945;33:144–9.
48. Liu S, Veldman P. Evidence-based endothelial rehabilitation. Semin Ophthalmol 2017;32:96–103.
49. Weleber RG, Watzke RC, Shults WT, et al. Clinical and electrophysiologic characterization of paraneoplastic and autoimmune retinopathies associates with antienolase antibodies. Am J Ophthalmol 2005;139:780–94.
50. Thirkill CE, Keltner JL, Tyler NK, et al. Antibody reactions with retina and cancer-associated antigens in 10 patients with cancer-associated retinopathy. Arch Ophthalmol 1993;111:931–7.
51. Ferreyra HA, Jayasundera T, Khan NW, et al. Management of autoimmune retinopathies with immunosuppression. Arch Ophthalmol 2009;127:390–7.
52. Hughes AE, Bridgett S, Meng W, et al. Sequence and expression of complement factor H gene cluster variants and their roles in age-related macular degeneration risk. Invest Ophthalmol Vis Sci 2016;57:2763–9.

MULTIPLE-CHOICE QUESTIONS

1. What is anterior chamber–associated immune deviation (ACAID)?
   A. Uveal or lens proteins, when exposed to the immune system, result in a robust reaction against these self antigens
   B. When an autoimmune process only affects the eye, but no other tissue
   C. Describes ocular immune privilege; antigens injected into the anterior chamber induce selective suppression of antigen-specific T-lymphocyte responses
   D. When viral infections affect only the anterior chamber but not the vitreous or retina

2. Acute anterior uveitis is:
   A. Associated with human leukocyte antigen B27 (HLA-B27) positivity 50% of the time
   B. Seldom associated with systemic disease
   C. Associated with spondyloarthritis without prior recognition 40% of the time
   D. None of the above
   E. A and C

3. Juvenile idiopathic arthritis (JIA)–associated uveitis:
   A. Often presents with pain, redness, photophobia
   B. Causes a smoldering asymptomatic bilateral uveitis most commonly
   C. Is typically associated with polyarticular arthritis in antinuclear antibody (ANA)–negative patients
   D. Cannot be treated with anti–tumor necrosis factor (TNF) treatment

75

Immunological Diseases of the Gastrointestinal Tract

Peter J. Mannon

The gastrointestinal (GI) tract contains the largest mass of immune cells of any organ in the body. To manage the continuous exposure to environmental and dietary antigens as well as pathogens, commensal microbes, and their metabolic products, the gut uses a number of strategies that support mucosal structural integrity and immune regulation to maintain homeostasis. In spite of these defenses, the GI tract is susceptible to chronic immune-mediated inflammation from pathogens, autoimmunity, and immunodeficiency. This chapter discusses some of the most frequent immunological diseases of the GI tract likely to be encountered in the clinic, including autoimmune gastritis, chronic *Helicobacter pylori* gastritis, celiac disease, inflammatory bowel diseases (IBDs; Crohn disease, ulcerative colitis [UC], microscopic colitis, and eosinophilic esophagitis) and GI complications of primary immunodeficiency. The coverage focuses on the clinical presentation, mechanisms of disease, and approaches to diagnosis and treatment of these conditions.

GASTRITIS

Gastritis is a histological term that describes stomach inflammation resulting from toxic exposures, infection, idiopathic inflammation, and autoimmunity. Although symptoms tend to be very nonspecific, the etiology and treatment of gastritis can be extremely specific, particularly for *H. pylori* infection, which has important implications for outcomes and natural history.

Atrophic Gastritis/Pernicious Anemia

The most classic autoimmune disease of the GI tract is atrophic gastritis characterized by antiintrinsic factor and antiparietal cell autoantibodies and an association with autoimmune thyroiditis, vitiligo, and type I diabetes.[1] Loss of acid-producing parietal cells leads to atrophic gastritis or "pernicious anemia (PA)." The age-old stereotype of the patient with atrophic gastritis/PA is an older Caucasian woman with progressive effects of iron and vitamin B_{12} deficiency. However, a female preponderance (approximately 2 : 1) notwithstanding, 25% of cases are diagnosed below age 50 years, and disease does occur in African and Asian ethnicities. The detection of both antiintrinsic factor and antiparietal cell serum antibodies has 73% sensitivity and 100% specificity for PA.[2]

Murine models suggested that induced immunity to H^+-K^+-ATPase (the parietal cell membrane protein that secretes H ions into the gastric lumen) may result in PA. Although potentially pathogenic H^+-K^+-ATPase–responsive CD4 T cells can be isolated from the gastric mucosa of patients with PA, it remains unclear how these cells arise, but observations implicate a role for *H. pylori* infection. Antral inflammation (up to 92%) and mucosal atrophy (up to 30%) together with histological evidence of *H. pylori* infection (up to 30%) occur in PA, with antigastric antibodies present in up to 65% of patients with *H. pylori* infection (and none in noninfected patients with gastritis).[3] In addition, gastric mucosal CD4 T cells can display cross-reactivity to H^+-K^+-ATPase and *H. pylori* antigens.[4] The actual identity of *H. pylori* infection as the major causal link to PA has not been established.

Helicobacter pylori Gastritis

H. pylori infection of the gastric mucosa has been identified as the leading cause of peptic ulcer disease and is a WHO-designated class I carcinogen for gastric carcinoma. Current data consistently show that primary infection is largely acquired in childhood, at which stage in life poor sanitation enhances acquisition via fecal/gastrooral pathways.[5] Although acute infection can cause abdominal pain and dyspepsia, there is typically no clinical recognition of acute infection. Rather, the burden of *H. pylori* results from chronic infection of the stomach; *H. pylori* is uniquely adapted to the acidic environment of the stomach through its ability to metabolize urea to ammonia, which provides a buffered microenvironment that allows prolonged asymptomatic colonization.

The ability of *H. pylori* to confer risk of peptic ulcer disease in 15% of chronically infected persons has been linked to *H. pylori* infection lowering the mucus pH adjacent to the gastric epithelium and increasing the permeability of the mucus to hydrogen ions. In addition to inhibition of bicarbonate secretion by the epithelium, expression of cytotoxins (vacuolating cytotoxin, VacA) and upmodulation of the mucosal inflammatory response by virulence factors, such as CagA, may contribute to pathogenesis. The recognition of *H. pylori* as a causative agent in the vast majority of recurrent peptic ulcers and that its elimination could cure this disease led to the Nobel Prize in Medicine or Physiology to Marshall and Warren in 2005.

The development of peptic ulcer disease and adenocarcinoma caused by chronic *H. pylori* infection correlates with the anatomical distribution of inflammation. When *H. pylori* chronic gastritis affects the antrum predominantly, there is an association with duodenal ulcers, increased serum gastrin levels and excess acid production, and no gastric mucosal atrophy.[6] However, when *H. pylori* affects the body and the antrum in a confluent or patchy manner, intestinal metaplasia develops, oxyntic mucosa atrophies, and acid production decreases. This latter type of *H. pylori* chronic gastritis is associated with gastric ulcerations and increased risk for adenocarcinoma and mucosa-associated lymphoreticular tissue (MALT) B-cell lymphoma. Although eradication of *H. pylori* can reverse the mucosal atrophy and restore acid production in this setting, mucosal restoration occurs only in a minority of patients and does not necessarily reverse the intestinal metaplasia. Histologically, this form of atrophic gastritis is differentiated from primary autoimmune atrophic gastritis by the presence of *H. pylori* organisms in the specimen (and concurrent involvement of the antrum). Specific immunohistochemical methods for detecting *H. pylori* are required when organisms are not seen on hematoxylin and eosin staining, when intestinal metaplasia occurs widely (*H. pylori* does not colonize intestinal metaplasia heavily), or when confirming *H. pylori* eradication after antibiotic treatment.[7] Although there are many pieces of evidence to support immune mechanisms for the persistence of *HP* infection in the stomach,[8] data suggest that pro-regulatory effects of *H. pylori* infection, including local IL-10 production, increases in regulatory T cells (Tregs) in the gastric mucosa and increased antigen-presenting cell (APC) phagocytosis of apoptotic cells all contribute to persistence of chronic *H. pylori* gastritis.[9-11] The indications for diagnosis and treatment of active *H. pylori* infection include active gastroduodenal ulcer disease, gastric MALT lymphoma, and, in highly endemic areas, dyspepsia symptoms (upper abdominal pain, bloating, early satiety, and nausea). Active disease can be diagnosed with endoscopic biopsy, which has high sensitivity and specificity, while simultaneously assessing peptic and malignant complications.[12] Noninvasive testing for *H. pylori* infection includes serum antibody detection (best used in highly endemic areas to predict active infection), urea breath testing (limited by expense and possible false-positive results), and fecal antigen testing (which has potential advantages in the setting of intestinal metaplasia and after antibiotic treatment).

Once *H. pylori* infection is diagnosed, there are many effective eradication therapies that need to be tailored to patients' drug tolerance and allergy history as well as local antibiotic resistance patterns. In general, a 14-day course with a proton pump inhibitor (histamine 2 [H2] blockers may be substituted) and two antibiotics (clarithromycin with amoxicillin or metronidazole) is recommended as first-line treatment.[13] Alternative regimens, including bismuth or sequential therapy, may be needed in cases of antibiotic resistance. Eradication of infection can be confirmed by either invasive or noninvasive (but not serum antibody) methods. In addition to the cure of recurrent gastroduodenal ulcer disease, eradication of *H. pylori* can cause regression of gastric MALT lymphoma in a majority of coinfected patients.[14] Investigation continues into enhancing eradication or preventing chronic infection through targeting molecules critical for *H. pylori* survival in the gastric environment as well as vaccine strategies.

KEY CONCEPTS

Gastritis

- Atrophic gastritis/Pernicious anemia results from loss of the acid-producing cells of the gastric body and is associated with antiparietal cell antibodies, an intact antral gastric mucosa, and no evidence of *Helicobacter pylori* infection.
- However, *H. pylori* infection has been linked to atrophic gastritis/pernicious anemia via increased rates of antigastric antibodies, autoantibodies to H^+-K^+-adenosine triphosphatase (ATPase), and gastric mucosal CD4 T cells with cross-reactivity to H^+-K^+-ATPase and *H. pylori* antigens.
- *H. pylori* infection of the stomach is chronic because of metabolic and immune adaptations by the pathogen that allow persistence in the acidic gastric environment.
- *H. pylori* infection is a leading cause of peptic ulcer disease and a World Health Organization (WHO) class I carcinogen because of its link to gastric adenocarcinoma.

ON THE HORIZON

Gastritis

- Development of an ultimate *Helicobacter pylori* eradication treatment with easy global implementation yet without significant adverse effects or antibiotic resistance issues.

CELIAC DISEASE

Celiac disease (also known as *gluten enteropathy*) is the most prevalent immune-mediated disease of the human GI tract. Gluten peptide–specific T cells drive specific autoantibody formation, small intestinal villus atrophy, and malabsorption caused by subsequent intestinal surface injury. The identification of the specific antigen-presenting human leukocyte antigen (HLA) molecules and the gluten peptide cognate ligands that activate the T cells have advanced the understanding of the pathophysiology and genetics of celiac disease and are driving innovative therapies.

Presentation

The classic clinical presentation of celiac disease—chronic diarrhea, weight loss, and abdominal bloating—results from defective nutrient absorption by the small intestine as a result of inflammatory injury. However, patients with celiac disease are more likely to present with complications of nutrient deficiency (anemia, osteoporosis, failure to thrive in children) without overt GI complaints. Celiac disease can also present with a gluten-reactive skin eruption (dermatitis herpetiformis), cerebellar ataxia, infertility and miscarriage, chronic fatigue, and associated autoimmune disorders, such as diabetes and thyroiditis, Addison disease, Sjögren syndrome, autoimmune hepatitis, and primary biliary cirrhosis.[15] A high frequency of nonspecific GI symptoms, such as abdominal pain and constipation, reported in over one-fifth of newly diagnosed subjects, complicates targeted disease screening. The variations of celiac disease presentation may reflect several features: *(i)* celiac disease severity follows a dose effect of risk alleles; *(ii)* the actual occurrence of celiac disease requires genetic and environmental factors in addition to the well-characterized HLA molecules; *(iii)* celiac disease expression depends on the quantity and quality of gluten exposure; and *(iv)* celiac disease complications reflect the severity of the inflammation and the length of involved bowel.[16] The concept of a vast subclinical celiac disease prevalence has led to an "iceberg model," where the visible tip is the group with symptomatic GI disease and the characteristic gut mucosal lesion, and below the surface are those with silent disease (no overt symptoms but gut mucosal lesions) and latent disease (no symptoms or mucosal lesions but genetic susceptibility possibly with positive serology).

The key to diagnosing celiac disease is simply to consider it in the differential diagnosis of classic gut malabsorption as well as subtle manifestations of malabsorption (*e.g.*, unexplained iron deficiency anemia). Conversely, the differential diagnosis of villus blunting or intraepithelial lymphocytosis includes small-intestinal bacterial overgrowth, tropical sprue, autoimmune enteropathy, common variable immunodeficiency (CVID) enteropathy, and *H. pylori* gastritis, emphasizing that celiac disease is also not solely a histological diagnosis.

Immune Pathophysiology

Celiac disease results from the activation of T cells by gluten peptide–MHC complexes on APCs in the lamina propria of the gut (primarily the small intestine). In the absence of dietary gluten or the specific MHC molecules HLA-DQ2 or -DQ8, celiac disease, theoretically, cannot occur. Dietary gluten, largely from wheat, barley, and rye, exists in polymeric (glutenin) and monomeric (gliadin) form and is incompletely digested to small peptides by gut luminal enzymes because of their high glutamine and proline content.[16] These large gluten peptides cross the epithelial barrier and bind to the specific HLA-DQ2 or -DQ8 molecules. The enzyme tissue transglutaminase can deamidate gluten peptides, and the resulting negatively charged gluten peptides actually have increased affinity for the HLA binding site. Gut microbes also can affect the immunogenicity of gluten peptides via their own proteolytic enzymes.[17] The gluten peptide-activated T cells produce proinflammatory cytokines interferon-γ (IFN-γ), interleukin-18 (IL-18), tumor necrosis factor (TNFGr-), and IL-21. The activated T cells also induce B-cell maturation to plasma cells producing antibodies to gluten peptides as well as the tissue transglutaminase. The reason for tissue transglutaminase being targeted for autoantibody production and its role in disease is unknown.

According to animal models, the activated T cells are not sufficient to induce the epithelial damage and villus blunting. The characteristic villus atrophy is induced by gut APC- and epithelial cell–produced IL-15, which enhances infiltration of CD8 T cells into the epithelium and upregulates NKG2D receptors on intraepithelial T lymphocytes (IELs) conferring cytotoxicity against the epithelium. More recently, it was found that the IEL signaling induced by NKG2D ligands MICA/B upregulated on epithelial cells stimulates a leukotriene pathway (together with IL-15 exposure) that itself produces cysteinyl leukotrienes (CystLT) that appear to drive the IEL-based killing. The impact of this observation is the possible use of montelukast, a CystLT inhibitor and drug widely used in allergic asthma, to treat celiac disease.[18]

Beyond the HLA associations with disease, genome-wide association studies (GWAS) have linked over 30 disease susceptibility loci with non-HLA regions. One such locus is a polymorphism in *lnc13*, a long noncoding RNA, which affects its binding with a heterogeneous nuclear ribonucleoprotein to undo its repression of other genes, including inflammatory mediators.[19]

Diagnosis

The diagnostic workup of celiac disease is initiated by both suggestive symptoms and signs (weight loss, chronic diarrhea) as well as by atypical presentations, such as specific micronutrient deficiencies or unexplained hyperamylasemia or hypertransaminasemia.[20] Initial tests include measuring serum immunoglobulin A (IgA) antibodies against tissue transglutaminase and endomysial proteins, which have an estimated specificity/sensitivity of 95%/95% and 100%/>90%, respectively.[21] It is necessary to measure a total serum IgA level at the same time to be certain that these IgA-based screening tests do not yield false–negative results. However, in the setting of IgA deficiency (where celiac disease has an increased incidence), elevated IgG anti–tissue transglutaminase (TTG) or deamidated gliadin levels and identification of celiac disease susceptibility HLA genes can help determine the risk and presence of disease.

In any case, an important part of celiac diagnosis is biopsy of the upper small intestine mucosa (Fig. 75.1B,C); typically three to four endoscopic biopsy specimens are obtained from both the duodenal bulb and the distal duodenum. Although the absolute requirement for histological diagnosis of celiac disease may be debated (and cannot be used alone because of lack of specificity), it remains important for several reasons: *(i)* Currently, the serologic markers should only be used as a screening test, identifying which patients are at highest risk for the disease and appropriate for biopsy confirmation; *(ii)* even in the presence of celiac-susceptible HLA genes, only a minority of persons will develop celiac disease (latent disease), so evaluating nonspecific symptoms will require histological examination; and *(iii)* because the treatment can be life altering, the most comprehensive information is needed to make a definitive diagnosis.

Beyond celiac disease, a true gluten-induced enteropathy, other gluten-centric entities have been identified, namely, non–celiac disease gluten sensitivity and IgE-mediated gluten allergy. Non–celiac disease gluten sensitivity includes a set of poorly characterized gut and extraintestinal symptoms that lack the defining features of celiac disease yet respond to a gluten-free diet (GFD).[22] Increased levels of serum markers for bacterial translocation (lipopolysaccharide [LPS]-binding protein, soluble CD14) and epithelial damage (circulating fatty acid binding protein 2) that is reversed following a wheat-free diet in a subset of subjects suggests altered gut permeability related to wheat and cereal exposure may be a factor contributing to symptoms.[23]

Treatment

The treatment of celiac disease is avoidance of gluten, specifically foods containing wheat, barley, and rye. The expected response to gluten restriction is the resolution of symptoms and malabsorption. Follow-up endoscopy to assess response to therapy should be done only after 6–12 months of a strict GFD, although full recovery of the villus mucosa may take several years. There are no accurate biomarkers to monitor adherence to a GFD, although one indication may be a fall in the pretreatment level IgA anti-TTG serum antibodies; therefore, follow-up endoscopy with biopsy is needed to document restoration of the villus architecture. Symptom improvement, correction of malabsorption, and regrowth of villi (seen by endoscopy or video capsule) may not reflect complete full histological healing, yet the long-term risk of partial histological recovery is unknown.

About 5% of patients with celiac disease do not respond to a GFD. Ensuring a strict adherence to a GFD is important to identify the reasons for nonresponse, whether through inadvertent gluten exposure or whether the inflammation is truly refractory to a strict GFD. One group of patients with the so-called refractory celiac disease with polyclonal IEL populations may respond to corticosteroids and immunosuppressant treatment; another group with monoclonal IELs do not respond to such treatment and are at increased risk of lymphoma.[24]

The majority of patients with celiac disease respond positively to a GFD with return of normal gut absorption. However, ongoing inflammation is associated with risk of small bowel lymphoma, so ensuring adherence to a GFD and documenting mucosal healing can affect the natural history of this disease. Finally, since first-degree relatives are at increased risk of celiac disease, patients should be advised of serum antibody screening of these family members.

CROHN DISEASE

Crohn disease is a chronic idiopathic inflammation of the gut characterized by transmural involvement of the bowel wall (mucosa, muscle layer, and serosa) (Fig. 75.1D). Although often referred to as "terminal ileitis," the majority of patients with Crohn disease have colonic inflammation in addition or solely.

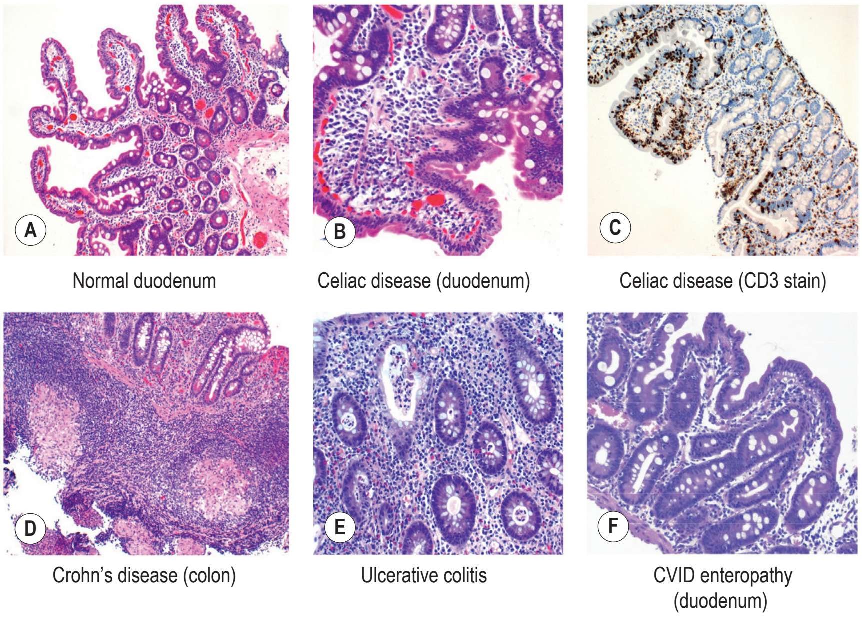

A
Normal duodenum
B
Celiac disease (duodenum)
C
Celiac disease (CD3 stain)
D
Crohn's disease (colon)
E
Ulcerative colitis
F
CVID enteropathy (duodenum)

FIG 75.1 Gastrointestinal Histology in Health and Immune-Mediated Disease. (A) Normal duodenal histology. (B**)** Celiac disease with blunted villus, increased plasma cell infiltrate, increased intraepithelial lymphocytes. (C) CD3 staining in celiac disease showing increased intraepithelial lymphocytes. (D) Crohn colitis showing mucosa expanded with lymphoplasmacytic infiltrate and two granulomata. (E) Ulcerative colitis showing crypt dropout, cryptitis, crypt abscess, and lymphoplasmacytic infiltrate. (F) Common variable immunodeficiency (CVID) enteropathy showing villus blunting, increased intraepithelial lymphocytes, and epithelial apoptosis. (Photomicrographs courtesy of Dr. Leona Council, Department of Pathology, University of Alabama at Birmingham, Birmingham, Alabama, USA.)

KEY CONCEPTS

Celiac Disease

- Patients with celiac disease present more often with complications of malabsorption than chronic diarrhea and weight loss.
- Human leukocyte antigen (HLA)-DQ2 or -DQ8 alleles are necessary but not sufficient for celiac disease to develop, as they are present in more persons unaffected by celiac disease.
- Immunoglobulin A (IgA) antibodies to tissue transglutaminase and endomysial proteins should be used for screening only (not diagnosis) and should be measured along with total IgA for validity.
- The goals of gluten-free diet treatment are relief of symptoms, reversal of malabsorption, and restoration of villi.

ON THE HORIZON

Celiac Disease

- Discovery of additional genetic and environmental factors (beyond gluten) that strongly confer risk for celiac disease in the setting of human leukocyte antigen (HLA)-DQ2 and -DQ8 alleles leading to strategies to modify or eliminate that risk
- Novel therapies that modify gluten to nonimmunogenic forms or that induce tolerance to gluten in persons at risk for or suffering from celiac disease or are alternatives to a gluten-free diet
- Identification of specific celiac disease–relevant components of the microbiome that influence the expression of disease

Crohn disease typically runs a chronic, relapsing course often complicated by bowel obstruction as a result of fibrous strictures as well as abscesses and fistulae caused by extension of inflammation beyond the bowel wall. Most of these patients will require surgical treatment at rates up to 80% after 20 years' disease duration. Crohn disease is treated with corticosteroids and immunosuppressants and more recently with antibodies targeting TNF-α and the integrin molecules α_4 and $\alpha_4\beta_7$.

The cause of Crohn disease is unknown, but it is thought to result from a dysregulated immune response to gut microbes. There is clear evidence for heritable disease susceptibility in twin and multiplex family studies, and several genes involved in innate immune function, notably mutations in *NOD2,* have been linked to Crohn disease risk. However, it seems clear that the complex interactions of environmental exposures (including the gut microbiome and its metabolome), innate and adaptive immune dysfunction, and complex genetic and epigenetic features are all complicit in disease causation and expression.[25]

Presentation

Patients with Crohn disease most often come to medical attention because of abdominal pain, altered bowel habits, and rectal bleeding. Abdominal pain may indicate bowel obstruction (especially if the pain is postprandial), an inflamed viscus, or a penetrating complication, such as an abscess or fistula. Diarrhea is related to malabsorption and dysmotility secondary to the effects of inflammatory cytokines on gut function. Diarrhea can also result from noninflammatory mechanisms (bile salt wasting, small intestinal bacterial overgrowth, etc.). Conversely, constipation in Crohn disease can be a sign of stricturing of the bowel. Rectal bleeding results from mucosal friability and ulceration. In addition, fever, unexplained weight loss, fatigue, anemia, and growth retardation (in children) can accompany the GI complaints or even be the primary presentation. Extraintestinal manifestations of Crohn disease include arthritis, uveitis, inflammatory skin lesions (pyoderma gangrenosum and erythema nodosum), and stomatitis. The arthritis can affect the axial (spine and pelvis) and articular skeleton, with the latter more often mirroring the activity of the gut disease. The joint complaints range from usual arthralgias to frank synovitis with swelling and tenderness (without erosive joint destruction). The uveitis most commonly occurs as episcleritis and iritis. Many of these lesions will subside with effective therapy aimed at the gut, but they can also have independent courses that require site-targeted treatment.

The incidence of Crohn disease in North American populations has been estimated to be ≈3.1–14.6 cases/100 000 person-years. Ashkenazi Jewish heritage confers increased risk in Caucasians, whereas African Americans seem to have rates similar to those of non-Jewish Caucasians, and Hispanics and Asians have much lower rates.[26,27] There is a genetic risk with up to 10-fold increased rates of IBD in relatives of patients with Crohn disease and a 30% concordance rate in monozygotic twin pairs.[28] The typical patient is diagnosed in his or her second or third decade, and there is no significant gender preference. The only environmental exposure that has been repeatedly linked to risk of Crohn disease has been tobacco use.

The majority (up to 70%) of Crohn patients experience a remitting and relapsing course, but some have chronically active symptoms refractory to medication. The recognized phenotypes of disease include inflammatory (manifesting primarily as intestinal edema and ulceration), fibrostenotic (luminal narrowing by fibrous strictures dominate with symptoms of painful obstruction), and fistulizing (inflammatory tracts between the bowel and other intestines, the bladder, vagina, skin, and other structures). Although the majority of patients have inflammatory disease at the time of diagnosis, over time this phenotype changes so that after 20 years of disease duration, up to 70% and 18% of patients with Crohn disease report penetrating and fibrostenotic complications, respectively.[29] Accompanying this change in phenotype is the need for surgical treatment in most patients after 20 years of disease duration.

Immune Pathophysiology

The current paradigm of Crohn disease inflammatory origins is a dysregulated immune response to gut commensal microbial components (antigens and pathogen-associated molecular patterns). Initial rodent colitis models showed a predominant T-helper cell-1 (Th1) inflammation, where the colitis could be blocked or reversed by treatment with anti–IL-12 antibodies. Roles for IL-23 and IL-17 in Crohn disease were later supported by the IL-10–deficient mouse model of spontaneous colitis and the cell transfer model of induced colitis, where colitogenic naïve $CD4^+CD45RB^{high}$ T cells from wild-type mice are infused into T cell–deficient mice.[30,31] However, despite the important role of the IL-23/IL-17 axis in the transfer model of colitis, blockade of IFN-γ activity also prevented development of gut inflammation.[32] Although studies of Crohn disease confirm that production of IL-12, IL-23, IFN-γ, and IL-17 are significantly elevated,[33] the hierarchy of cytokine control of inflammation in Crohn disease has been tested by clinical trials using agents that have targeted IFN-γ, IL-12, IL-23, and IL-17. The anti–IFN-γ monoclonal antibody (mAb) fontolizumab showed no clinical efficacy in Crohn disease (but did decrease C-reactive protein [CRP] levels[34]). However, anti–IL-12/-23 p40 subunit mAbs targeting the Th1 and Th17 pathways resulted in clinical improvement that led to the current use of ustekinumab for Crohn disease.[35,36] Unexpectedly, antibodies targeting IL-17A alone (secukinumab) or the IL-17AR receptor subunit (transducing IL-17A, -17E, and -17F intracellular signals) either did not show clinical improvement (and was associated with increased fungal infections[37]) or even induced clinical deterioration (the anti–IL-17AR antibody).

To date, it is estimated from multiple genome polymorphism studies that over 160 IBD susceptibility loci exist, most associated with risk of both Crohn disease and UC, 30 loci associated only with Crohn disease, and 23 loci associated only with UC. However, it was further estimated that all the loci together could account for no more than 13.6% of Crohn disease and 7.5% of UC disease risk.[38] Although most of these loci are in the noncoding regions of genes thought to modulate gene expression, the actual genes implicated have roles in the immune response and epithelial integrity, providing biological plausibility for their involvement in IBD. Several examples stand out: *(i)* the *NOD2* gene, encodes an intracellular protein that binds muramyl dipeptide (MDP), a component of the bacterial cell wall TLR2 ligand peptidoglycan. Disease-associated mutations in the MDP-sensing leucine-rich repeat domain of *NOD2* are associated with defective activation of nuclear factor (NF)-κB. This disrupts the microbe recognition pathway, but because *NOD2* is also expressed by small intestinal epithelial Paneth cells, mucosal production of antimicrobial defensins is decreased as well. *(ii)* The *ATG16L1* autophagy gene is important for the metabolism of autologous cell proteins as well as intracellular microbes. Expression of the Crohn disease–associated Thr300Ala polymorphic *ATG16L1* sequence in a colon cancer cell line showed *in vitro* inhibition of packaging of intracellular *Salmonella* into autophagosomes, supporting the hypothesis that this mutation could lead to impaired clearance of microbes and chronic inflammation.[39] This *ATG16L1* polymorphism provides a caspase cleavage site that enhances degradation of this allele and leads to defective autophagy.[40] *(iii)* A polymorphism in the coding region of the IL-23 receptor gene (Arg381Gln) found in 14% of healthy controls is associated with protection from Crohn disease (less so UC) and is associated with decreased Th17 cascade effector cells.[41]

The emergence of the role of the commensal microbiota in immune homeostasis has produced renewed interest in its role in IBD, where a number of studies show a dysbiosis compared with the healthy gut microbiota. This change in the microbiota in IBD has been characterized as a flipping of the Bacteroidetes-to-Firmicutes phyla ratio with reduction in Firmicutes.[42] In addition, Crohn disease has been associated with a reduction in the butyrate-producing *Faecalibacterium prausnitzii* compared with healthy controls, but it is not known whether this is a cause or effect of the gut inflammation. The roles of antibiotics, probiotics, prebiotics, or actively changing the microbiota through fecal transplant have not been clarified sufficiently to employ these strategies as conventional treatments in Crohn disease.

Diagnosis

The diagnosis of Crohn disease is based on findings from radiographic, endoscopic, and histological examinations. In general, a combination of colonoscopy (with ileal intubation, if possible) and small bowel examination (barium small bowel follow-through, computed tomography [CT] or magnetic resonance enterography, and capsule video endoscopy) is usually sufficient to demonstrate active inflammatory disease of the colon and small bowel. Endoscopically, mucosal ulceration and friability in a patchy distribution separated by unaffected mucosa ("skip areas") are hallmarks of the disease. Radiographically, evidence of patchy bowel wall thickening, mucosal hyperemia, stricturing, and penetrating complications, such as fistulae, abscesses, and extraintestinal inflammatory masses involving the bowel, all suggest Crohn disease. Histologically, although the appearance of noncaseating granulomata is highly supportive of a diagnosis of Crohn disease, in practice, they are not often detected by endoscopic biopsy, particularly in adults. More often evidence of chronic inflammation, such as architectural crypt distortion and basal lymphoplasmacytosis, are sought to differentiate the inflammation from an acute, self-limiting colitis or enteritis. Other findings, such as fecal leukocytes or elevated fecal calprotectin, may indicate an inflammatory colitis but is not specific for diagnosis of a chronic idiopathic IBD, such as Crohn disease or UC. In the setting of supportive findings through imaging or endoscopy, the measurement of certain serum antibodies can further strengthen the diagnosis of Crohn disease and even help differentiate it from UC, but they should not be used by themselves as diagnostic tests.[43] It has been shown that up to 68% of patients with Crohn disease are seropositive for antibodies targeting microbial antigens, such as anti–*Saccharomyces cerevisiae* antibody (up to 16% of patients with UC are seropositive). Additional antimicrobial antibodies, such as anti-OmpC, anti-I2, anti-flagellin 3, X, and CBir, also develop in Crohn disease.

It goes without saying that prior to treatment, Crohn disease may need to be differentiated from other similarly presenting conditions, including UC, chronic ischemic colitis, infectious enteritis/colitis (amebiasis, *Yersinia enterocolitica* infection, *Mycobacterium tuberculosis* infection), intestinal lymphoma, celiac disease, diverticula-associated colitis, and radiation- and nonsteroidal antiinflammatory drug (NSAID)–induced enteropathy.

Treatment

The treatment of Crohn disease includes medical and surgical approaches. Since there is no cure for Crohn disease, the principles of therapy are to first make sure that symptoms are secondary to the underlying idiopathic inflammation and not caused by infectious or noninflammatory factors, such as coexisting irritable bowel syndrome (IBS). The goal is to induce quick remission of symptoms and establish therapy to maintain the remission with aggressive initial therapy commensurate with the extent and activity of the disease. A traditional approach to therapy has been to use short-term corticosteroids and mesalamine preparations to start with in mild disease (these agents are not useful for long-term maintenance of a clinical response), add immunosuppressants (azathioprine, 6-mercaptopurine, or methotrexate) if symptoms persist or are moderate, and finally start anti–TNF-α agents for recalcitrant or initially severe disease.[44] Only corticosteroids and the anti–TNF-α agents are relied on for inducing clinical responses in short periods, within days to weeks. Therefore corticosteroid use may need to be prolonged while immunosuppressants are begun concomitantly, since the latter's clinical effects may not be maximized until months later. For maintenance of remission, only immunosuppressants and anti–TNF-α agents reliably exert long-term control of symptoms. In this way, they may be used for years as long as the disease is responding. Currently, medically refractory inflammatory disease may be treated with natalizumab, an anti-α_4 integrin antibody, although this is being supplanted by vedolizumab, an anti-$\alpha_4\beta_7$ antibody inhibiting the trafficking of lymphocytes from the blood specifically into the gut lamina propria. Because of substantially reduced concerns about a rare infectious side effect with vedolizumab (progressive multifocal leukoencephalopathy has been associated with natalizumab), it is a more attractive antiintegrin therapy for IBD. An established approach to immunosuppressant- and biological-naïve patients is to begin with a combination of azathioprine and infliximab, as this has been shown to be superior to either a single agent alone for inducing remission, especially 1 year later.[45] This "top-down" approach continues to be studied for long-term efficacy as well as for identifying those patients with Crohn disease with a pretreatment risk profile predicting more aggressive disease.

Surgery is required in cases of complications, such as bleeding, pain/obstruction, and fistulae that are refractory to medical therapy. Surgery typically involves resection limited to inflamed segments of small intestine and colon; small strictures can be treated *in situ* by stricturoplasty. In addition, surgery is required for treatment of intestinal adenocarcinoma, which also complicates the chronic inflammation of the bowel.[46] There is a high incidence of endoscopic and symptomatic recurrence of inflammation by 2 years after surgery, but to date, there is no consensus about who should routinely receive preventive therapy.

KEY CONCEPTS

Crohn Disease

- Crohn disease affects the full thickness of the bowel wall resulting in fistulae and abscesses in 70% of patients by 20 years' disease duration.
- Crohn disease is a chronic, relapsing inflammation of the bowel, and up to 80% of patients will require surgical treatment at some point.
- Several gene mutations and many genetic loci are associated with disease risk, including *NOD2* and *ATG16L1*, which are important in innate immune function.
- T-helper 1 (Th1) and Th17 cytokine pathways mediate disease in animal models, but the hierarchy of these cytokine pathways in human disease is still being tested by targeting strategies in clinical trials.

ULCERATIVE COLITIS

UC also involves chronic idiopathic inflammation of the gut, but this is limited to the mucosal layer of the colon (Fig. 75.1E). It can affect the rectum alone (ulcerative proctitis), the distal transverse colon to rectum (left-sided colitis), or the entire colon (pancolitis). Because it lacks the transmural inflammation of Crohn disease, penetrating complications, such as fistulae and abscesses, are generally not features of UC. Oral and topical (per rectum) mesalamine preparations are most often used to treat UC, although patients also may require corticosteroids, immunosuppressants, and anti–TNF-α antibody agents to induce and maintain remission. Unlike Crohn disease, total colectomy eliminates the disease. However, there can be ongoing complications of the surgery, such as pouchitis, when an ileal pouch–anal anastomosis is performed. Over 40% of patients with UC require surgery to treat medically refractory symptoms or development of dysplasia over their lifetime.[47]

Presentation

The presentation of UC reflects the primary involvement of the rectum and extent of proximal colitis. Bloody stool and diarrhea (including nocturnal) are common symptoms, with proctitis-specific complaints of rectal urgency and incomplete evacuation also being prominent. As in Crohn disease, fever, unexplained weight loss, fatigue, and anemia can accompany the GI complaints or be the primary presentation. Extraintestinal manifestations may include arthritis, uveitis, inflammatory skin lesions, and stomatitis. An interesting genetic connection exists between UC and HLA-B27–positive spondyloarthropathy (Chapter 57), with 60% of patients with ankylosing spondylitis showing inflammation on colonoscopy. UC is also closely associated with primary sclerosing cholangitis (PSC; Chapter 76); up to 3% of UC patients develop PSC, and up to 75% of all patients with PSC have UC.[48]

With an incidence up to 20.3 cases per 100 000 person-years, typically the diagnosis of UC is in the second or the third decade, and there is no significant gender preference. Approximately 6–15% of first-degree relatives of patients with UC have a history of UC, but in general, the genetic contribution to risk is lower than in Crohn disease. There is a higher incidence of UC in European and North American populations compared with that in Asian and Latin American countries. The only environmental exposures linked to risk of UC are the protective effect of tobacco use and appendectomy in the first decade of life.

The natural history of UC shows that most patients experience a remitting and relapsing course (60%), although some have prolonged remission after one episode of disease (20%), and others have chronically active symptoms refractory to medical remission (20%). Total colectomy is performed to treat refractory symptoms or development of epithelial dysplasia. Chronic inflammatory UC (and Crohn colitis) is accompanied by an increased incidence of colorectal cancer (18% after 30 years' disease duration), so much so that recurring colonoscopic surveillance for dysplasia with biopsy is recommended starting 8–10 years after diagnosis.

Immune Pathophysiology

UC has been characterized as a Th2-like disease because of the increased IL-5 and IL-13 production in inflamed gut tissue seen in an animal model of UC, oxazolone-induced colitis, as well as from patient specimens.[49] In this model, not only were mucosal natural killer T (NKT) cells the source of excess IL-13, but the colitis was reversed by immunoneutralizing anti–IL-13.[50,51] When translated to humans, patients with UC were found to have high capacity for IL-13 production, also by type II NKT cells. It turns out that IL-13 is a biologically plausible effector cytokine in UC injury because it disrupts the epithelial tight junction by upregulating claudin-2 and has a direct toxic effect on human gut epithelial cells *in vitro.*[52] However, results from a clinical trial using an anti–IL-13 antibody in UC did not show significant treatment efficacy.[53]

Data from genetic susceptibility studies in UC have provided less compelling examples of disease-specific mechanisms compared with those in Crohn disease, but there are associations with HLA class II genes distinct from Crohn disease. Furthermore, other associated loci contain genes involved in epithelial barrier function, such as *GNA12* for tight junction formation, *CDH* for epithelial cadherin-1, and *LAMB1* for the laminin constituent of basement membranes.[28] When specifically assaying epithelial cells (vs whole tissue) for epigenetic markers associated with genomic risk loci, there is an enhanced association supporting the epithelium as a primary target for dysfunction in UC.[54]

Diagnosis

The diagnosis of UC is based on a colitis that typically involves the rectum and adjacent proximal colonic segments in a confluent inflammatory injury. This is best accomplished by colonoscopy (ileal intubation can confirm that the inflammation is limited to the colon). Biopsy specimens should contain histological features of chronic inflammation, including crypt distortion, crypt dropout, and lymphoplasmacytosis. The presence of acute inflammatory features alone (polymorphonuclear cells, crypt abscess, and cryptitis) may also be seen but, when in isolation, these features suggest other etiologies, such as acute infectious, drug-induced, ischemia, and toxic exposures. Although no blood test can be used to diagnose UC, the presence of high-titer perinuclear antineutrophil cytoplasmic antibodies (pANCAs) can be seen in up to 70% of patients with UC; this information can help differentiate chronic indeterminate colitis when coupled with anti-*Saccharomyces cerevisiae mannan* antibodies (ASCAs; see above). At all times, acute infections from enteric pathogens, including *Clostridium difficile,* should also be excluded, as these may also occur during active IBD treatment and mimic exacerbation of disease (evaluation for cytomegalovirus [CMV] infection should be performed in the setting of UC seemingly refractory to immunosuppressants). Once a diagnosis is established, elevations in transaminases or alkaline phosphatase should prompt an evaluation for PSC starting with magnetic resonance cholangiopancreatography.

Treatment

Treatment is tailored to the extent and activity of disease.[55] For instance, mild to moderate proctitis can respond to topical corticosteroids or mesalamine alone (enemas and/or suppositories). Most often with more extensive colonic involvement, oral mesalamine is required, which can be useful for induction and maintenance of remission. In moderate to severe disease, corticosteroids may be required to induce rapid responses, whereas immunosuppressants, such as azathioprine or 6-mercaptopurine, are used for remission. Anti–TNF-α agents are used to induce fairly rapid clinical responses and remission and may also be used as maintenance therapy. Vedolizumab has been a recent addition to conventional therapy as well.

Mesalamine-based drugs are a cornerstone of therapy in UC and are generally included in most ongoing UC medical regimens. Whether use of mesalamine even in quiescent disease confers chemoprotection from developing dysplasia is still being debated, but because of their generally low adverse event rate and high tolerance, long-term use is a reasonable choice.

As discussed, surgery has a definite role in treating medically refractory disease or addressing dysplasia discovered by surveillance colonoscopy (dysplasia surveillance is done every 1–2 years after 8–10 years of disease by taking four-quadrant biopsy specimens every 10 cm). Total colectomy is performed, and options include an ileal pouch–anal anastomosis or a permanent end ileostomy. However, even the pouch can become chronically inflamed; generally antibiotic-responsive, this inflammation can also become refractory and require steroid, immunosuppressant, or biological therapy and even pouch removal.

KEY CONCEPTS

Ulcerative Colitis

- Ulcerative colitis is a chronic, relapsing inflammation of the colon that is limited to the mucosa and is not transmural.
- Animal models and human disease show the dominant cytokines to be interleukin (IL)-5 and IL-13, suggesting a T-helper 2 (Th2)–like (without IL-4) inflammatory response, although IL-17 and interferon (IFN)-γ are also active.
- Although less strongly associated with genetic inheritance compared with Crohn disease, several disease susceptibility loci have been identified and are associated with epithelial barrier function.
- Over the course of the disease up to 40% of patients will undergo total colectomy for refractory symptoms or detection of epithelial dysplasia.

ON THE HORIZON

Inflammatory Bowel Disease (IBD)

- Establishment of new therapies targeting key cytokines and signaling pathways, especially anti-SMAD7 for transforming growth factor (TGF)-β enhancement, JAK/STAT inhibition, and immune cell–based treatments
- Improvement in detection of dysplasia using molecular techniques for better understanding of the incidence, natural history, treatment, and prevention of colonic epithelial dysplasia and neoplasia
- Refining genome-wide association studies (GWAS) using large-scale fine-mapping to identify the actual polymorphisms in coding and noncoding regions and to better test the mechanisms of IBD susceptibility
- Defining whether the dysbiosis in IBD is a cause or effect and how the microbiome is influenced by genetic background and inflammatory status

OTHER IDIOPATHIC INFLAMMATORY BOWEL DISEASES

Microscopic Colitis

Microscopic colitis is an increasingly recognized condition that links chronic watery diarrhea with isolated intraepithelial lymphocytosis (lymphocytic colitis) or with increased subepithelial collagen deposition (collagenous colitis). It differs from Crohn disease and UC because it does not display endoscopic mucosal damage or show evidence of histological chronicity (no architectural crypt distortion, lymphoplasmacytosis, or loss of Goblet cells). However, it causes significant morbidity and may require chronic immunosuppression for treatment. Although its etiology remains elusive, there are associations with autoimmune conditions as well as with certain drug exposures.

The hallmark symptom of microscopic colitis is chronic, watery, nonbloody diarrhea that is frequently accompanied by marked complaints of poor quality of life. Fatigue, arthralgias, and weight loss may also be reported. Microscopic colitis has its typical onset in the sixth and seventh decades, has a female predominance, and is associated with a history of autoimmunity especially thyroid disease, rheumatoid arthritis, and CREST (calcinosis, Raynaud phenomenon, esophageal dysmotility, sclerodactyly, and telangiectasia) syndrome.[56] A connection between microscopic colitis and celiac disease has also be observed, with lymphocytosis of the colonic epithelium found in one-third of patients with classic malabsorption symptoms and small-bowel lesions of celiac disease. This may be a sequela of gluten antigen reactivity extended to colonic epithelial responses, but it becomes important to address as an independent process related to the microscopic colitis itself if diarrhea and colonic lymphocytosis persist despite a GFD. Conversely, although celiac-like villus blunting may be seen in <10% of patients with microscopic colitis, celiac serologies will not be positive, indicating that classic gluten enteropathy is not playing a role.

There is no current information on the immune mechanism of microscopic colitis. Activated NF-κB is seen in the mucosa of microscopic colitis, consistent with the general inflammatory picture. Increased nitric oxide and prostaglandin production have also been measured, possibly contributing more to the diarrhea rather than the inflammation. Excessive transforming growth factor (TGF)-β has also been measured in collagenous colitis, consistent with its role in collagen production and fibrosis. Interestingly, medications have been associated with microscopic colitis, including H2 blockers, proton pump inhibitors (PPIs), selective serotonin reuptake inhibitors (SSRIs), and ticlopidine, among others.[57] Reports also suggest that long-term use of NSAIDs may induce or sustain microscopic colitis.

The diagnosis of microscopic colitis relies on histology. The *sine qua non* for microscopic colitis is increased numbers of intraepithelial lymphocytes (>20 lymphocytes/100 epithelial cells) on colonic mucosal biopsy. This can be accompanied by a chronic inflammatory infiltrate in the lamina propria and, less often, a limited appearance of neutrophils (especially cryptitis), the latter finding suggesting that the etiology of the microscopic colitis may actually be related to an injurious drug effect, such as NSAID exposure. In collagenous colitis, a prominent subepithelial collagen band ≥10 μm is seen in addition to the intraepithelial lymphocytosis.

As a general treatment approach to the patient with microscopic colitis, the clinician should review and eliminate suspect concomitant medications, such as NSAIDs. Coexisting celiac disease should be considered, where appropriate, and addressed with gluten withdrawal. Overall the choice of treatment begins with low-risk medications in mild or moderate disease in an attempt to settle on a regimen that delivers the most relief of symptoms with the lowest side effects.[58] Therefore treatment can begin with antidiarrheals (loperamide, diphenoxylate/atropine), adding a trial of cholestyramine (bile salt malabsorption has been hypothesized to play a role in microscopic colitis) and even a trial of mesalamine. Bismuth subsalicylate has been reported to benefit a minority of patients with microscopic colitis; three 262 mg tablets PO three times daily may result in long-term remission in some (some have reservations about the long-term toxic effects of bismuth). In patients with initially severe symptoms or who are refractory to these first treatments, corticosteroids have been very effective. In particular, oral budesonide 9 mg taken once per day has reliably improved diarrhea and induced remission, a result supported by placebo-controlled studies.[59] The challenge is balancing relief of symptoms with the long-term side effects of corticosteroid use (including budesonide) in relapsing or steroid-dependent disease. In this case, emerging strategies include use of lower doses of budesonide while monitoring for side effects or using steroid-sparing immunosuppressive medications, such as azathioprine and methotrexate, for long-term maintenance. Reports of anti–TNF-α drugs for refractory microscopic colitis have been published.[60] Occasionally, patients may need colectomy and ileostomy to manage refractory symptoms or drug intolerance.

Eosinophilic Esophagitis

Eosinophilic esophagitis (EoE) is a more recently recognized disease defined by symptomatic idiopathic eosinophilic inflammation in the absence of other known causes, especially chronic gastroesophageal reflux disease (GERD).[61] Although its etiopathology is unclear, there are gene associations (*FLG* for epithelial barrier effects and *eotaxin-3* and *TSLP* for immune response effects), and data from animal models and human disease have implicated central roles for loss of tolerance to food antigens involving Th2 cytokines IL-5 and IL-13. EoE is being recognized with increasing frequency against a background of increased incidence of inflammatory allergic diseases.

EoE is estimated to occur at 4–5 cases/10 000 children, has a male predominance (up to 70%), and peak incidence in neonates to 3-year-olds. Symptoms include failure to thrive and feeding difficulty in infants (*e.g.,* food refusal, limited variety diet, prolonged feeding times) and abdominal pain and vomiting in older children and adolescents. In adults, the primary symptom is typically intermittent dysphagia, with the first presentation possibly being an acute food impaction in the esophagus. Adult patients may report GERD symptoms that do not respond to adequate acid-suppression therapy. Patients report a high rate (>50%) of atopy (rhinoconjunctivitis, wheezing, or family history of atopy) as well as food allergies (including positive skin prick test, allergen-specific IgE test, or anaphylactic response to a dietary allergen).[62] There is also an association of esophageal disease (strictures or EoE) in the parents of up to 10% of patients.

The diagnosis of EoE requires endoscopic biopsy of the esophagus, since increased eosinophils in the esophageal epithelium must be measured. The endoscopic appearance of the esophagus can show multiple thin rings (“feline esophagus”), with linear longitudinal furrows and whitish papules that represent eosinophilic microabscesses at the surface of the squamous epithelium. Biopsy should show an infiltrate of eosinophils in the epithelium of at least 15—20 eosinophils/high-power field (hpf). These often concentrate just under the epithelial surface and also form microabscesses in groups of ≥4 eosinophils. It is important to take at least three biopsy specimens, since involvement may be variable and patchy; in addition, it is advisable to take biopsy specimens from the distal and mid-to-upper esophagus (to help differentiate changes seen in GERD that can be limited to the distal esophagus) and specimens from the gastric and duodenal mucosae (to show that the eosinophilic infiltration is limited to the esophagus and does not represent a diffuse process, such as that found in eosinophilic gastroenteritis or hypereosinophilic syndrome). Systemic eosinophilia can be seen in over 70% of patients with EoE.

In terms of GERD, it is important to make sure that any excessive acid reflux is treated and controlled; persistent symptoms (and persistent biopsy abnormalities) may prompt a 24-hour ambulatory pH study of the distal esophagus to ensure that the acid-suppression treatment results in a normal acid-contact time. In fact, there seems to be a form of EoE that, while seeming to be without excess exposure to gastric acid, does, in fact, respond well to PPI treatment; this may be attributed to the non–acid-suppressive effects of a drug, such as omeprazole, that can suppress eotaxin-3 secretion from squamous mucosae.[63] Multiple types of esophageal dysmotility, often reversible with treatment, have been described in EoE, contributing more to dysphagia than to strictures.[64]

The pathogenesis of EoE seems to be linked to allergen hypersensitivity. Given a familial association of EoE, atopy, and food allergy, a genetic component may be contributing to disease susceptibility. Several candidate disease susceptibility gene/gene loci in EoE include the 3’ untranslated region of eotaxin (*CCL26*), the TGF-β_1 promoter, a filaggrin (*FLG*) exon, and a thymic stromal lymphopoietin (*TSLP*) intron and TSLP receptor (*CRLF2*) exon.[65] These associations are biologically plausible, since eotaxin is excessively expressed in EoE mucosae, filaggrin is a structural skin protein that helps maintain barrier function (and is downregulated by IL-13), and TSLP has been shown to stimulate IL-13 production by innate helper cells in the lamina propria. Moreover, the inflammation in EoE is characterized by increased IL-13 and IL-5 production; animal models of aeroallergen induction of an EoE-like lesion is blocked in both IL-13–deficient and signal transducer and activator of transcription 6 (STAT6)–deficient (an intracellular molecule important for IL-13 receptor α_1 signaling) animals.[66] These data suggest that IL-13 secretion induces production of eotaxin from epithelial cells, which, together with IL-5, drives the local eosinophilic infiltration. Finally, the association with food allergy has led to successful therapy of EoE by using strict elimination diets (sometimes informed by skin testing) or even the use of elemental diet tube feedings.

Given the strong association with food allergies, strict elimination of suspect foods or those identified by *in vitro* or skin prick testing may be restricted first. Lack of improvement in symptoms would lead to a trial of amino acid–based elemental liquid diet necessitating nasogastric (or later percutaneous gastrostomy) placement, but this can be uncomfortable, impractical, and expensive. If this dietary approach is successful, then after several weeks, individual foods may be added back every 5–7 days. For patients not responding to dietary therapy or with no identifiable dietary allergens, corticosteroid treatment has been used successfully. Both systemic oral and swallowed topical corticosteroids (*e.g.,* fluticasone propionate metered dose inhalers) for 4–6 weeks have been shown to relieve symptoms and resolve histological inflammation; one or the other may be more or less effective based on body weight, dose, steroid resistance, or severity of the EoE. However, relapse rates are high over the year following a course of steroids. Lastly, endoscopic therapy to treat strictures using dilation may incur higher risk of mucosal tears so that conservative treatment (smaller dilators, assessment for tears during the procedure before proceeding further) is encouraged.

Eosinophilic gastroenteritis encompasses an additional group of infrequent conditions that link increased intestinal eosinophilia (from mucosal to full-thickness intestinal wall infiltration) and symptoms, such as abdominal pain, diarrhea, and malabsorption.[67] These are assumed to be primary inflammatory responses, although exclusion of parasitic infections, especially with helminths, is a necessary part of the workup.

GASTROINTESTINAL COMPLICATIONS OF PRIMARY IMMUNODEFICIENCIES

Certain primary immunodeficiency diseases increase the risk for GI complications. These complications fall into three main categories: infectious, idiopathic inflammatory/autoimmune, and neoplastic. Common variable immunodeficiency (CVID) and chronic granulomatous disease (CGD) have some of the most frequent and significant gut manifestations, and these are discussed in detail below. However, a variety of inherited conditions of broad lymphocyte or innate immune cell dysfunction can also lead to gut disease. The study of monogenic diseases that are associated with inflammatory diseases of the gut have come to the attention of investigators, as they may be models for the mechanisms that could contribute to the pathophysiology of more genetically complex conditions, such as Crohn disease and UC.[68]

Common Variable Immunodeficiency

CVID is a syndrome of hypogammaglobulinemia (low levels of IgG and IgA and/or IgM) accompanied by recurrent sinopulmonary infections (Chapter 34). Patients have no isohemagglutinins and cannot mount an adequate antibody response to many vaccine antigens. In general, Ig replacement therapy improves the sinopulmonary infection rate but does not affect other complications, such as autoimmune disease and GI symptoms, including intestinal infections. Case series of patients with CVID have shown that up to 60% of patients experience GI symptoms[69] and that endoscopic and histological abnormalities can occur in the majority of patients with CVID. When considering the differential diagnostic possibilities for the GI complications of CVID, it is helpful to separate them into infectious, immune-mediated, and neoplastic processes. Among the infectious agents, *Giardia lamblia,* nontyphoidal *Salmonella,* and *Campylobacter jejuni* are frequently seen, but *Cryptosporidium* may also be found, and *C. difficile* and viral agents (CMV) may be encountered. In addition to the higher rate of gastric *H. pylori* infection, small-intestinal bacterial overgrowth (SIBO) is present in up to 30% of patients with CVID. The immune-mediated GI complications of CVID include idiopathic enteropathy (villous atrophy, increased intraepithelial lymphocytes/microscopic colitis, nodular lymphoid hyperplasia) (Fig. 75.1F) manifesting as severe malabsorption (see below) and, much less frequently, macroscopic ulcerating disease resembling UC or Crohn disease.[69] CVID-associated autoimmune disease involving the GI tract also includes type II gastritis that can lead to achlorhydria and vitamin B_{12} deficiency ("pernicious anemia") and even autoimmune hepatitis and primary biliary cirrhosis. Last, neoplastic complications of intestinal lymphoma and gastric adenocarcinoma (related to achlorhydric autoimmune gastritis) have been reported.

Presentation

The most common symptom of GI complications of CVID is episodic diarrhea that can progress to chronic diarrhea, regardless of the etiology. Weight loss and evidence of vitamin or mineral deficiencies may occur, with underlying maldigestion or malabsorption. In addition, patients may present with abdominal pain related to splenomegaly (with or without portal hypertension) or ascites (portal hypertension secondary to hepatic nodular regenerative hyperplasia [NRH] as a known complication of CVID). Fever, together with intestinal obstruction and GI bleeding, may indicate the development of small-bowel lymphoma. Patients with CVID may report increased frequency of GERD symptoms as well as dyspepsia, but these complaints are nonspecific and have been difficult to link with immune defects or autoimmune complications.

Immune Pathophysiology

It is not clear if lack of Igs alone causes susceptibility to gut infections and inflammation because patients with X-linked agammaglobulinemia infrequently develop overt GI disease (they lack B cells but have functionally intact T-cell populations). Even persons with selective IgA deficiency have little in the way of GI disease, including infections. In fact, the T-cell dysfunction that often accompanies CVID may contribute substantially to susceptibility to GI disease.

In terms of an increased GI infection risk, autoimmune gastritis-induced achlorhydria could increase small-bowel exposure to swallowed commensals and pathogens escaping this innate gastric barrier to infection. In addition, although the loss of secreted IgA alone is not enough to meaningfully increase susceptibility to intestinal infections, perhaps the additional loss of IgG and IgM and presence of T-cell dysfunction is more permissive, but no mechanism for this has been established. Mucosal nodular lymphoid hyperplasia, characterized by disorganized secondary lymphoid nodules with poorly formed germinal centers, is likely related to the inability of B cells to undergo class switching when presented with antigen *in situ.* Last, CVID enteropathy is characterized clinically by severe malabsorption and histologically by blunted villi and increased intraepithelial lymphocytes, and epithelial apoptosis is associated with excess Th1 cytokine secretion (IL-12 and IFN-γ).[70] It remains unknown why this inflammatory lesion occurs in the first place.

Diagnosis

The workup of GI complications of CVID often takes place as symptoms are usually first episodic but may eventually turn into chronic, progressive complaints.[71] The key to successful diagnosis is the initial search for treatable infectious causes. This requires stool assay and culture for bacterial (especially *Campylobacter*) and protozoal pathogens. In addition, hydrogen breath testing can help detect SIBO to provide another antibiotic-responsive etiology of the symptoms.

Negative results for infectious causes require endoscopy for biopsy of the proximal small intestine and colon. Routine histology will detect the features of enteropathy, including blunted villi and increased IELs, which are often interpreted as celiac disease. Unlike celiac disease, there usually is a lack of plasma cell infiltrate and crypt hyperplasia, as well as a preserved brush border and Goblet cells, and there is often an increase in epithelial cell apoptosis, particularly in the colon.[72] In fact, some of the villus changes in CVID enteropathy can be mimicked by SIBO, which should be investigated as described above. If the diagnosis of celiac disease must be considered, then genetic testing for celiac susceptibility HLA alleles should be performed; the absence of combinations of A and B alleles for HLA-DQ2 or -DQ8 that confer risk will rule this out. However, even if celiac disease gene alleles are detected, it still does not indicate that this is a gluten-driven celiac disease lesion but will require a trial of GFD nonetheless. The functional significance of any histological lesion in the small intestine can be evaluated by a measure of steatorrhea (fecal fat excretion) and small-bowel absorption (d-xylose absorption test), values that can be used to track improvement following treatments.

Treatment

None of the GI complications of CVID is treated by intravenous or subcutaneous immunoglobulin (IV/SCIG) (initial treatment might alleviate GI symptoms, it does not treat overt gut inflammation and ulcers),[73] and in fact, these complications typically occur in patients who have adequate trough levels of IgG. Patients with any bacterial or protozoal pathogens should be treated with a recommended course of conventional antibiotics, including those positive for the *C. difficile* toxin. SIBO should be treated, and recurrent SIBO may need cycling antibiotic regimens.

The treatment of the idiopathic enteropathy is very challenging. Although this seems to be a late complication in a subset of patients, it can be fatal. In the early stages, it may be responsive to a short course of oral corticosteroids, either prednisone or budesonide. Case reports attest to the efficacy of infliximab, but this cannot be routinely advised.[74] It is possible that immunosuppressants may be used to control the inflammatory response underlying the small-bowel mucosal damage, but this should only be done in a closely observed clinical setting, with monitoring for infections. At all times the patient's nutritional status should be maintained, initially using the oral route but administering parenteral nutrition to compliment oral nutrition and when oral feedings are not adequately absorbed and exacerbating the diarrhea.

Portal hypertension may complicate CVID, typically from the nodular regenerative hyperplasia of the liver; however, little fibrosis occurs. It needs to be clarified in individual patients with splenomegaly (≈20%) that portal hypertension is not caused by excessive splenic vein flow associated with splenomegaly (induced by antibody-mediated autoimmune cytopenias) that might be ameliorated by splenectomy. In any scenario, the management of such late complications requires an especially experienced team of internists, surgeons, and nutritionists.

KEY CONCEPTS

Common Variable Immunodeficiency (CVID)

- The majority of gastrointestinal (GI) complications of CVID are infectious and generally do not respond to intravenous or subcutaneous immunoglobulin (IV/SCIG) therapy (compared with sinopulmonary suppurative infections).
- CVID enteropathy is a rare immune-mediated complication of CVID that also does not respond to IV/SCIG.
- CVID enteropathy is often confused with celiac disease because of similar villus damage on biopsy, but additional features (lack of plasma cells, increased epithelial apoptosis, absence of celiac gene risk alleles) can help differentiate them.
- CVID enteropathy has no established therapy though judicious use of short courses of oral steroids, or conventional immunosuppression may relieve the malabsorption and diarrhea temporarily.

Chronic Granulomatous Disease

CGD results from defects in the nicotinamide adenine dinucleotide phosphate (NADPH)–oxidase complex that impair the ability of phagocytic cells to produce the reactive oxygen species required to kill bacteria and fungi within intracellular phagolysosomes (Chapter 22). Patients with CGD have recurrent infections of the skin, lungs, liver, and bone, and nearly half develop gut inflammatory complications affecting areas anywhere from the mouth to the anus. Interestingly, the frequency of GI disease in CGD is higher in the X-linked gp91^{phox} defect, but the most recently described p40^{phox} defect occurred in a young male patient who presented with granulomatous colitis alone.[75-77]

Presentation

The most typical GI complaints in patients with CGD are abdominal pain and diarrhea (with or without rectal bleeding). GI symptoms usually begin in the first decade of life, sometimes preceding the diagnosis of CGD. Although infectious diarrhea (especially *Salmonella* spp. and *C. difficile*) occurs in CGD, an idiopathic CGD-associated IBD also develops: In the mouth, granulomatous stomatitis and dental abscesses cause pain and difficulty eating; in the esophagus, dysphagia, chest pain, and vomiting may result from narrowing by strictures or stenosis and dysmotility related to granulomatous inflammation and fibrosis; in the stomach, loss of motility and capacity caused by thickened walls and narrowed lumen leads to vomiting, epigastric pain, and weight loss as a result of feeding difficulty; in the small and large intestines, diarrhea (including protein losing enteropathy), bowel obstruction (large granulomata compromising the size of the lumen), rectal bleeding, and tenesmus may result from active colitis/enteritis with mucosal ulceration, anal fissures, and perianal abscesses. In addition, because of the transmural nature of the granulomatous inflammation, penetrating complications, such as fistulae and abscesses, can occur. Feeding difficulties and the chronic inflammatory state itself predispose to growth delay that often affects pediatric patients with CGD.

Hepatic abscesses represent another frequent complication in CGD,[78] occurring in up to 45% of patients. These patients most often present with fever as well as abdominal pain, fatigue, and, less often, abdominal tenderness and hepatomegaly on examination. The erythrocyte sedimentation rate (ESR) and the alkaline phosphatase level are elevated in half the affected individuals. However, a high level of suspicion, especially in the setting of fever with or without abdominal pain, should instigate a search for hepatic abscesses.

Immune Pathophysiology

Given the defects in ability to kill intracellular bacteria and fungi, pathogens, and possibly commensals alike, it is thought that the exuberant granulomatous response is caused by delayed antigenic clearance or persistent infection. In this way, granulomata continue to multiply and grow while other inflammatory pathways that normally deal with the microbes or are induced by cytokines are activated. The end results of granulomatous inflammation are most evident in tissues rich in macrophages and reticuloendothelial cells, such as the gut lamina propria, liver, lymph nodes, and spleen.

Diagnosis

The symptoms and signs will dictate the initial diagnostic examinations. For diarrheal complaints, stool culture and examination for *C. difficile* toxin are required; in the setting of hypoalbuminemia, fractional fecal excretion of α_1 antitrypsin can detect protein-losing enteropathy (>50 mg/24 hour) as a result of either diffuse mucosal inflammation or lymphangiectasia. For complaints of dysphagia, vomiting, or epigastric pain, upper endoscopy can help document macroscopic and microscopic involvement with granulomatous inflammation. Radiological studies using oral contrast may be helpful in showing a narrowed lumen, stricturing, and motility and mucosal abnormalities of the esophagus and stomach but cannot provide histological confirmation. However, radiological imaging studies may be the primary diagnostic tools to evaluate obstructive symptoms from the small intestine, including barium small-bowel studies and CT or magnetic resonance enterography. Images can show lumen narrowing; bowel wall thickening; mucosal abnormalities, including ulceration; and penetrating complications, such as fistulization. Finally, to evaluate lower abdominal and perianal pain and rectal bleeding, anoscopy, colonoscopy, and pelvic CT or magnetic resonance imaging (MRI) will help diagnose granulomatous inflammation of the colon and its complications, including perirectal and perianal abscesses. In the case of hepatic abscess detection (generally 1–6 cm), CT, MRI, and ultrasonography have similar sensitivity (≈60%). Active abscesses appear to be solid, hypoechoic lesions on ultrasonography or postcontrast ring–enhancing lesions on CT and MRI.

The GI histological diagnosis hinges on the presence of noncaseating granulomata, both gross and microscopic. These granulomata are often seen against a background of acute inflammation (acute focal colitis, crypt abscesses, cryptitis) as well as chronic inflammation (lymphocytic infiltrate, Paneth cell metaplasia in the colon) that can be mild to severe. The histological picture may resemble Crohn disease except that the granulomata of CGD are well defined, often large collections of epithelioid histiocytes that can expand the mucosa (and even deform the overlying epithelium to make it look flattened in the case of the villous mucosa of the small intestine). Like Crohn disease, the inflammation can affect the three layers of the gut wall, but unlike in Crohn disease, CGD biopsy specimens also show prominent lipid-laden macrophages that have periodic acid-Schiff (PAS)–positive cytoplasmic granules.

Treatment

Current clinical practice for CGD includes using prophylactic antimicrobials to prevent infections, typically trimethoprim-sulfamethoxazole for bacterial and itraconazole for fungal infections. Some clinicians also use IFN-γ (subcutaneous, three-times-weekly dosing) to prevent infections, although this is not a universal practice.[79] Obviously, discovery of infectious etiologies of diarrhea should be treated appropriately.

Once infections are ruled out and granulomatous inflammation of the GI tract is established, with or without complications, such as stricturing of the bowel, treatment with corticosteroids is indicated; beginning doses up to 1 mg/kg/day tapering over 12–20 weeks to maintenance doses of 2.5–5 mg every other day has been reported to induce rapid alleviation of symptoms. Use of sulfasalazine for colitis may have limited benefit in some patients. Isolated reports of successful use of cyclosporine and infliximab have indicated that these agents are best reserved for refractory cases because of the potential for infectious side effects. Similarly both granulocyte–colony-stimulating factor (G-CSF) and granulocyte macrophage–colony-stimulating factor (GM-CSF) have been used with therapeutic benefit in GI complications of CGD because of their success in the granulomatous colitis of glycogen-1-β storage disease.[80] Although there are no data to suggest that IFN-γ worsens established disease, it also may not prevent it as >40% of one cohort developed GI manifestations of granulomatous inflammation after starting it; in contrast, isolated reports have attributed alleviation of GI inflammation to IFN-γ.[81] IL-1 receptor blockade, based on preclinical efficacy data in CGD, may offer limited success.[82]

Finally, it cannot be overstated that surgical drainage of complicating abscesses and resection of fibrotic or refractory strictures need to be pursued, when indicated. Although there can be considerable postoperative complications because of ongoing fistula formation and wound breakdown, these problems can be managed by administration of corticosteroids. In addition, judicious use of endoscopic therapy to dilate narrowed esophageal or pyloric regions is an option for symptomatic strictures.

Treatment of hepatic abscesses has been shown to be amenable to nonsurgical management by adding corticosteroids to the antibiotic regimen.[78] Empiric antibiotic coverage (at least to cover the most frequent isolate, *S. aureus*) is generally appropriate although percutaneous aspiration of hepatic abscesses maybe helpful.[83]

KEY CONCEPTS

Chronic Granulomatous Disease (CGD)

- Half of patients with CGD develop gastrointestinal (GI) involvement with granulomatous inflammation.
- Symptoms are caused by both obstructive complications of stricture formation and mucosal inflammation.
- A resemblance to Crohn disease is suggested by granulomatous inflammation, transmural bowel involvement, and underlying innate immune defect, but unlike in Crohn disease, there are also prominent skin and lung infections and periodic acid-Schiff–positive lipid-laden macrophages in the lamina propria.
- Complicated hepatic abscesses require specialized medical intervention.

GI Complications Occurring in Other Primary Immunodeficiency States

Severe combined immunodeficiency (SCID) covers a wide phenotype of low-to-absent T cells, NK cells, and dysfunctional B cells reflecting the mechanisms of the genetic defect (Chapter 35). Recurrent infectious diarrhea and thrush are typical GI conditions before a diagnosis of SCID in newborns and neonates.[84] After allogeneic hematopoietic stem cell transplantation (HSCT), GI graft-versus-host disease (GvHD) is a possibility. For many of the monogenic immunodeficiency diseases that confer susceptibility to GI inflammation, the success of HSCT to correct the secondary effect on the GI tract will depend entirely on whether the gene defect affects the myeloid or lymphoid cells primarily. If the gut stromal or epithelial cells also depend on normal function of the defective gene, then bone marrow transplantation will not likely relieve the GI effects.

The rare X-linked recessive Wiskott-Aldrich syndrome (Chapter 35) results from Wiskott-Aldrich syndrome (*WAS*) gene mutations, a signaling protein largely restricted to hematopoietic cells, and leads to a syndrome of eczema, thrombocytopenia, and infections related to combined immunodeficiency (impaired antibody responses and T-cell function). An interesting aspect of the gut complication in this immunodeficiency is that patients may develop a noninfectious colitis resembling UC (confluent mucosal inflammation with ulceration, crypt abscesses, and no granulomata) with the bleeding exacerbated by the thrombocytopenia.[85] Patients may respond to mesalamine drugs, but use of steroids and immunosuppressive must be done cautiously because of exacerbating infection risk. Successful bone marrow transplantation can treat the colitis as well.

Early-onset IBD is a severe complication of rare mutations that affect IL-10 production or the IL-10 receptor. Patients present with colitis within weeks of birth that has features of Crohn disease. Diagnosis is made by genetic testing as well as by assay of peripheral blood mononuclear cells for absence of IL-10 production or absent IL-10 signaling demonstrated by decreased/absent induction of phospho-STAT3 following the addition of IL-10.[86] The only cure is with HSCT.

The hyper-IgM syndrome (type 1) (Chapter 34) is an X-linked condition resulting from mutations in CD40 ligand leading to defective antibody class switching (hence high [or normal] IgM with low IgA and IgG) and NK- and T-cell cytotoxicity.

Cryptosporidium infectious diarrhea is seen, and sclerosing cholangitis, cirrhosis, and neoplasms of the hepatobiliary system complicate the course in teenagers and young adults.[87] Rarely, noninfectious colitis has been reported. HSCT, preferably done before chronic complications occur, is the only approach with a potential for cure. X-linked lymphoproliferative syndrome 2 (*XIAP* deficiency) can present with very early-onset fistulizing perianal disease in up to 20% of patients.[88]

Last, IPEX (immune dysfunction/polyendocrinopathy/enteropathy/X-linked) syndrome is interesting in that it results from a defect in FOXP3 expression and function, thereby inhibiting the generation of Tregs (Chapter 18). There is a lack of immune regulation, and enhanced responses, including autoimmunity, occur. The gut mucosa is a major site affected by inflammation, and this complication presents with a watery, sometimes bloody, diarrhea and malabsorption. Biopsies show villus blunting and atrophy in the small bowel and increased intraepithelial lymphocytes, lamina propria infiltration with T cells, eosinophils, and neutrophils elsewhere in the gut.[89] Treatment requires HSCT, but the gut disease may be managed temporarily with corticosteroids and immunosuppressants. Of note, one of the side effects of the anti– cytotoxic T lymphocyte antigen-4 (CTLA-4) antibody ipilimumab, given to augment a T-cell response against solid tumors, is a colitis that can have penetrating ulcers; this unexpected side effect points to the importance of regulatory mechanisms for gut immune homeostasis.[90]

ON THE HORIZON

Primary Immunodeficiency Diseases

- Translation of monogenic immunodeficiency disease mechanisms, especially in very-early-onset inflammatory bowel disease (IBD), into knowledge of pathophysiology (genetic and acquired) of Crohn disease and ulcerative colitis for novel drug development

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Neumann WL, Coss E, Rugge M, et al. Autoimmune atrophic gastritis—pathogenesis, pathology and management. Nat Rev Gastroenterol Hepatol 2013;10:529–41.
2. Lahner E, Norman GL, Severi C, et al. Reassessment of intrinsic factor and parietal cell autoantibodies in atrophic gastritis with respect to cobalamin deficiency. Am J Gastroenterol 2009;104:2071–9.
3. Negrini R, Savio A, Poiesi C, et al. Antigenic mimicry between *Helicobacter pylori* and gastric mucosa in the pathogenesis of body atrophic gastritis. Gastroenterology 1996;111:655–65.
4. Amedei A, Bergman MP, Appelmelk BJ, et al. Molecular mimicry between *Helicobacter pylori* antigens and H+, K+ –adenosine triphosphatase in human gastric autoimmunity. J Exp Med 2003;198:1147–56.
5. Ernst PB, Peura DA, Crowe SE. The translation of *Helicobacter pylori* basic research to patient care. Gastroenterology 2006;130:188–206, quiz 12–3.
6. Egan BJ, Holmes K, O'Connor HJ, et al. *Helicobacter pylori* gastritis, the unifying concept for gastric diseases. Helicobacter 2007;12:39–44.
7. Sepulveda AR, Patil M. Practical approach to the pathologic diagnosis of gastritis. Arch Pathol Lab Med 2008;132:1586–93.
8. Hunt RH, Camilleri M, Crowe SE, et al. The stomach in health and disease. Gut 2015;64:1650–68.
9. Kivrak Salim D, Sahin M, Koksoy S, et al. Local Immune Response in *Helicobacter pylori* Infection. Medicine (Baltimore) 2016;95:e3713.
10. Das S, Sarkar A, Ryan KA, et al. Brain angiogenesis inhibitor 1 is expressed by gastric phagocytes during infection with *Helicobacter pylori* and mediates the recognition and engulfment of human apoptotic gastric epithelial cells. FASEB J 2014;28:2214–24.
11. Kronsteiner B, Bassaganya-Riera J, Philipson C, et al. Systems-wide analyses of mucosal immune responses to Helico*bacter pylori* at the interface between pathogenicity and symbiosis. Gut Microbes 2016;7:3–21.
12. Chey WD, Wong BC. American College of Gastroenterology guideline on the management of *Helicobacter pylori* infection. Am J Gastroenterol 2007;102:1808–25.
13. Marcus EA, Sachs G, Scott DR. Eradication of *Helicobacter pylori* Infection. Curr Gastroenterol Rep 2016;18:33.
14. Neumeister P, Troppan K, Raderer M. Management of gastric mucosa-associated lymphoid tissue lymphoma. Dig Dis 2015;33:11–18.
15. Fasano A. Systemic autoimmune disorders in celiac disease. Curr Opin Gastroenterol 2006;22:674–9.
16. Abadie V, Sollid LM, Barreiro LB, et al. Integration of genetic and immunological insights into a model of celiac disease pathogenesis. Annu Rev Immunol 2011;29:493–525.
17. Caminero A, Galipeau HJ, McCarville JL, et al. Duodenal bacteria from patients with celiac disease and healthy subjects distinctly affect gluten breakdown and immunogenicity. Gastroenterology 2016;151:670–83.
18. Tang F, Sally B, Lesko K, et al. Cysteinyl leukotrienes mediate lymphokine killer activity induced by NKG2D and IL-15 in cytotoxic T cells during celiac disease. J Exp Med 2015;212:1487–95.
19. Castellanos-Rubio A, Fernandez-Jimenez N, Kratchmarov R, et al. A long noncoding RNA associated with susceptibility to celiac disease. Science 2016;352:91–5.
20. Castillo NE, Vanga RR, Theethira TG, et al. Prevalence of abnormal liver function tests in celiac disease and the effect of a gluten-free diet in the US population. Am J Gastroenterol 2015;110:1216–22.
21. Adriaanse M, Leffler DA. Serum markers in the clinical management of celiac disease. Dig Dis 2015;33:236–43.
22. Troncone R, Jabri B. Coeliac disease and gluten sensitivity. J Intern Med 2011;269:582–90.
23. Uhde M, Ajamian M, Caio G, et al. Intestinal cell damage and systemic immune activation in individuals reporting sensitivity to wheat in the absence of coeliac disease. Gut 2016;65:1930–7.
24. van Gils T, Nijeboer P, van Wanrooij RL, et al. Mechanisms and management of refractory coeliac disease. Nat Rev Gastroenterol Hepatol 2015;12:572–9.
25. Fiocchi C. Inflammatory bowel disease pathogenesis: where are we? J Gastroenterol Hepatol 2015;30:12–18.
26. Hou JK, El-Serag H, Thirumurthi S. Distribution and manifestations of inflammatory bowel disease in Asians, Hispanics, and African Americans: a systematic review. Am J Gastroenterol 2009;104:2100–9.
27. Mahid SS, Mulhall AM, Gholson RD, et al. Inflammatory bowel disease and African Americans: a systematic review. Inflamm Bowel Dis 2008;14:960–7.
28. Cho JH, Brant SR. Recent insights into the genetics of inflammatory bowel disease. Gastroenterology 2011;140:1704–12.
29. Cosnes J, Cattan S, Blain A, et al. Long-term evolution of disease behavior of Crohn's disease. Inflamm Bowel Dis 2002;8:244–50.
30. Uhlig HH, McKenzie BS, Hue S, et al. Differential activity of IL-12 and IL-23 in mucosal and systemic innate immune pathology. Immunity 2006;25:309–18.
31. Yen D, Cheung J, Scheerens H, et al. IL-23 is essential for T cell-mediated colitis and promotes inflammation via IL-17 and IL-6. J Clin Invest 2006;116:1310–16.
32. Powrie F, Leach MW, Mauze S, et al. Inhibition of Th1 responses prevents inflammatory bowel disease in SCID mice reconstituted with CD45RBhi CD4+ T cells. Immunity 1994;1:553–62.
33. Fuss IJ, Becker C, Yang Z, et al. Both IL-12p70 and IL-23 are synthesized during active Crohn's disease and are down-regulated by treatment with anti-IL-12 p40 monoclonal antibody. Inflamm Bowel Dis 2006;12:9–15.

34. Reinisch W, de Villiers W, Bene L, et al. Fontolizumab in moderate to severe Crohn's disease: a phase 2, randomized, double-blind, placebo-controlled, multiple-dose study. Inflamm Bowel Dis 2010;16:233–42.
35. Mannon PJ, Fuss IJ, Mayer L, et al. Anti-interleukin-12 antibody for active Crohn's disease. N Engl J Med 2004;351:2069–79.
36. Sandborn WJ, Gasink C, Gao LL, et al. Ustekinumab induction and maintenance therapy in refractory Crohn's disease. N Engl J Med 2012;367:1519–28.
37. Hueber W, Sands BE, Lewitzky S, et al. Secukinumab, a human anti-IL-17A monoclonal antibody, for moderate to severe Crohn's disease: unexpected results of a randomised, double-blind placebo-controlled trial. Gut 2012;61:1693–700.
38. McGovern DP, Kugathasan S, Cho JH. Genetics of Inflammatory Bowel Diseases. Gastroenterology 2015;149:1163–76, e2.
39. Kuballa P, Huett A, Rioux JD, et al. Impaired autophagy of an intracellular pathogen induced by a Crohn's disease associated ATG16L1 variant. PLoS ONE 2008;3:e3391.
40. Murthy A, Li Y, Peng I, et al. A Crohn's disease variant in Atg16l1 enhances its degradation by caspase 3. Nature 2014;506:456–62.
41. Sarin R, Wu X, Abraham C. Inflammatory disease protective R381Q IL23 receptor polymorphism results in decreased primary CD4+ and CD8+ human T-cell functional responses. Proc Natl Acad Sci USA 2011;108:9560–5.
42. Erickson AR, Cantarel BL, Lamendella R, et al. Integrated metagenomics/metaproteomics reveals human host-microbiota signatures of Crohn's disease. PLoS ONE 2012;7:e49138.
43. Reese GE, Constantinides VA, Simillis C, et al. Diagnostic precision of anti-Saccharomyces cerevisiae antibodies and perinuclear antineutrophil cytoplasmic antibodies in inflammatory bowel disease. Am J Gastroenterol 2006;101:2410–22.
44. Baumgart DC, Sandborn WJ. Inflammatory bowel disease: clinical aspects and established and evolving therapies. Lancet 2007;369:1641–57.
45. Colombel JF, Sandborn WJ, Reinisch W, et al. Infliximab, azathioprine, or combination therapy for Crohn's disease. N Engl J Med 2010;362:1383–95.
46. Triantafillidis JK, Nasioulas G, Kosmidis PA. Colorectal cancer and inflammatory bowel disease: epidemiology, risk factors, mechanisms of carcinogenesis and prevention strategies. Anticancer Res 2009;29:2727–37.
47. Farmer RG, Easley KA, Rankin GB. Clinical patterns, natural history, and progression of ulcerative colitis. A long-term follow-up of 1116 patients. Dig Dis Sci 1993;38:1137–46.
48. Loftus EV Jr, Harewood GC, Loftus CG, et al. PSC-IBD: a unique form of inflammatory bowel disease associated with primary sclerosing cholangitis. Gut 2005;54:91–6.
49. Fuss IJ, Strober W. The role of IL-13 and NK T cells in experimental and human ulcerative colitis. Mucosal Immunol 2008;1:S31–3.
50. Heller F, Fuss IJ, Nieuwenhuis EE, et al. Oxazolone colitis, a Th2 colitis model resembling ulcerative colitis, is mediated by IL-13-producing NK-T cells. Immunity 2002;17:629–38.
51. Fuss IJ, Heller F, Boirivant M, et al. Nonclassical CD1d-restricted NK T cells that produce IL-13 characterize an atypical Th2 response in ulcerative colitis. J Clin Invest 2004;113:1490–7.
52. Heller F, Florian P, Bojarski C, et al. Interleukin-13 is the key effector Th2 cytokine in ulcerative colitis that affects epithelial tight junctions, apoptosis, and cell restitution. Gastroenterology 2005;129:550–64.
53. Reinisch W, Panes J, Khurana S, et al. Anrukinzumab, an anti-interleukin 13 monoclonal antibody, in active UC: efficacy and safety from a phase IIa randomised multicentre study. Gut 2015;64:894–900.
54. Mokry M, Middendorp S, Wiegerinck CL, et al. Many inflammatory bowel disease risk loci include regions that regulate gene expression in immune cells and the intestinal epithelium. Gastroenterology 2014;146:1040–7.
55. Kornbluth A, Sachar DB. Ulcerative colitis practice guidelines in adults: American College Of Gastroenterology, Practice Parameters Committee. Am J Gastroenterol 2010;105:501–23, quiz 524.
56. Kao KT, Pedraza BA, McClune AC, et al. Microscopic colitis: a large retrospective analysis from a health maintenance organization experience. World J Gastroenterol 2009;15:3122–7.
57. Beaugerie L, Pardi DS. Review article: drug-induced microscopic colitis - proposal for a scoring system and review of the literature. Aliment Pharmacol Ther 2005;22:277–84.
58. Pardi DS, Kelly CP. Microscopic colitis. Gastroenterology 2011;140:1155–65.
59. Stewart MJ, Seow CH, Storr MA. Prednisolone and budesonide for short- and long-term treatment of microscopic colitis: systematic review and meta-analysis. Clin Gastroenterol Hepatol 2011;9:881–90.
60. Esteve M, Mahadevan U, Sainz E, et al. Efficacy of anti-TNF therapies in refractory severe microscopic colitis. J Crohns Colitis 2011;5:612–18.
61. Furuta GT, Liacouras CA, Collins MH, et al. Eosinophilic esophagitis in children and adults: a systematic review and consensus recommendations for diagnosis and treatment. Gastroenterology 2007;133:1342–63.
62. Noel RJ, Putnam PE, Rothenberg ME. Eosinophilic esophagitis. N Engl J Med 2004;351:940–1.
63. Cheng E, Zhang X, Huo X, et al. Omeprazole blocks eotaxin-3 expression by oesophageal squamous cells from patients with eosinophilic oesophagitis and GORD. Gut 2013;62:824–32.
64. Nurko S, Rosen R. Esophageal dysmotility in patients who have eosinophilic esophagitis. Gastrointest Endosc Clin N Am 2008;18:73–89, ix.
65. Sherrill JD, Rothenberg ME. Genetic dissection of eosinophilic esophagitis provides insight into disease pathogenesis and treatment strategies. J Allergy Clin Immunol 2011;128:23–32, quiz 33–4.
66. Travers J, Rothenberg ME. Eosinophils in mucosal immune responses. Mucosal Immunol 2015;8:464–75.
67. Mehta P, Furuta GT. Eosinophils in Gastrointestinal Disorders: Eosinophilic Gastrointestinal Diseases, Celiac Disease, Inflammatory Bowel Diseases, and Parasitic Infections. Immunol Allergy Clin North Am 2015;35:413–37.
68. Uhlig HH. Monogenic diseases associated with intestinal inflammation: implications for the understanding of inflammatory bowel disease. Gut 2013;62:1795–805.
69. Malamut G, Verkarre V, Suarez F, et al. The enteropathy associated with common variable immunodeficiency: the delineated frontiers with celiac disease. Am J Gastroenterol 2010;105:2262–75.
70. Mannon PJ, Fuss IJ, Dill S, et al. Excess IL-12 but not IL-23 accompanies the inflammatory bowel disease associated with common variable immunodeficiency. Gastroenterology 2006;131:748–56.
71. Agarwal S, Mayer L. Diagnosis and treatment of gastrointestinal disorders in patients with primary immunodeficiency. Clin Gastroenterol Hepatol 2013;11:1050–63.
72. Washington K, Stenzel TT, Buckley RH, et al. Gastrointestinal pathology in patients with common variable immunodeficiency and X-linked agammaglobulinemia. Am J Surg Pathol 1996;20:1240–52.
73. Sanges M, Spadaro G, Miniero M, et al. Efficacy of subcutaneous immunoglobulins in primary immunodeficiency with Crohn's-like phenotype: report of a case. Eur Rev Med Pharmacol Sci 2015;19:2641–5.
74. Akazawa Y, Takeshima F, Yajima H, et al. Infliximab therapy for Crohn's-like disease in common variable immunodeficiency complicated by massive intestinal hemorrhage: a case report. BMC Res Notes 2014;7:382.
75. Matute JD, Arias AA, Wright NA, et al. A new genetic subgroup of chronic granulomatous disease with autosomal recessive mutations in p40 phox and selective defects in neutrophil NADPH oxidase activity. Blood 2009;114:3309–15.
76. Barton LL, Moussa SL, Villar RG, et al. Gastrointestinal complications of chronic granulomatous disease: case report and literature review. Clin Pediatr (Phila) 1998;37:231–6.
77. Marciano BE, Rosenzweig SD, Kleiner DE, et al. Gastrointestinal involvement in chronic granulomatous disease. Pediatrics 2004;114:462–8.
78. Leiding JW, Freeman AF, Marciano BE, et al. Corticosteroid therapy for liver abscess in chronic granulomatous disease. Clin Infect Dis 2012;54:694–700.
79. Marciano BE, Wesley R, De Carlo ES, et al. Long-term interferon-gamma therapy for patients with chronic granulomatous disease. Clin Infect Dis 2004;39:692–9.

80. Wang J, Mayer L, Cunningham-Rundles C. Use of GM-CSF in the treatment of colitis associated with chronic granulomatous disease. J Allergy Clin Immunol 2005;115:1092–4.
81. Born M, Willinek WA, Hassan C. [Gastric outlet obstruction in chronic granulomatous disease]. Z Gastroenterol 2002;40:511–16.
82. Hahn KJ, Ho N, Yockey L, et al. Treatment with anakinra, a recombinant IL-1 Receptor antagonist, unlikely to induce lasting remission in patients with CGD colitis. Am J Gastroenterol 2015;110:938–9.
83. Lublin M, Bartlett DL, Danforth DN, et al. Hepatic abscess in patients with chronic granulomatous disease. Ann Surg 2002;235:383–91.
84. Buckley RH. Primary cellular immunodeficiencies. J Allergy Clin Immunol 2002;109:747–57.
85. Hsieh KH, Chang MH, Lee CY, et al. Wiskott-Aldrich syndrome and inflammatory bowel disease. Ann Allergy 1988;60:429–31.
86. Glocker EO, Kotlarz D, Boztug K, et al. Inflammatory bowel disease and mutations affecting the interleukin-10 receptor. N Engl J Med 2009;361:2033–45.
87. Winkelstein JA, Marino MC, Ochs H, et al. The X-linked hyper-IgM syndrome: clinical and immunologic features of 79 patients. Medicine (Baltimore) 2003;82:373–84.
88. Pachlopnik Schmid J, Canioni D, Moshous D, et al. Clinical similarities and differences of patients with X-linked lymphoproliferative syndrome type 1 (XLP-1/SAP deficiency) versus type 2 (XLP-2/XIAP deficiency). Blood 2011;117:1522–9.
89. Patey-Mariaud de Serre N, Canioni D, Ganousse S, et al. Digestive histopathological presentation of IPEX syndrome. Mod Pathol 2009;22:95–102.
90. Friedman CF, Proverbs-Singh TA, Postow MA. Treatment of the Immune-related adverse effects of immune checkpoint inhibitors: a review. JAMA Oncol 2016;2:1346–53.

MULTIPLE-CHOICE QUESTIONS

1. A 73-year-old Caucasian female who is undergoing workup for chronic diarrhea and weight loss has been incidentally found to have a gastrin level of 2500 pg/mL and have serum immunoglobulin G (IgG) antibody positive for *Helicobacter pylori*. There is no history of peptic ulcer disease, pyrosis, or kidney stones, and yet she has been referred for evaluation for gastrinoma. The first test you order is:
   A. Esophagogastroduodenoscopy with biopsy
   B. Abdominal computed tomography (CT)
   C. Serologic testing for celiac disease
   D. Secretin stimulation test
   E. Fecal testing for *H. pylori* antigen

2. A 22-year-old Caucasian female is referred for chronic abdominal pain and bloating. She reports that a cousin has celiac disease, and she says that "every time that I eat bread, I feel bad and get diarrhea." She has maintained her weight and has no other specific symptoms. She is on no medications and reports a history of irritable bowel syndrome (IBD). She says that her primary physician performed a "celiac test but it was negative." Your first choice of tests is:
   A. Esophagogastroduodenoscopy with biopsies of the duodenum and duodenal bulb
   B. Genetic human leukocyte antigen (HLA) testing for DQ2 and DQ8 risk alleles
   C. Repeat testing for serum IgA antibodies to tissue transglutaminase, including total serum IgA level
   D. Monitor anti–tissue transglutaminase levels while the patient is on a gluten-free diet

3. Which of the following is characteristic of ulcerative colitis?
   A. Presence of antiflagellin antibodies
   B. Inflammation of the full thickness of the bowel wall
   C. Tobacco use increases the frequency
   D. Surgery can be curative

Inflammatory Hepatobiliary Diseases

Carlo Selmi, Michael P. Manns, M. Eric Gershwin

Inflammatory hepatobiliary disease generically refers to chronic autoimmune diseases associated with predominant hepatic or biliary clinical manifestations, with no evidence of infectious disease. Based on specific cellular targets, we can distinguish autoimmune hepatitis (AIH) targeting hepatocytes from primary biliary cholangitis (PBC) (also known as primary biliary cirrhosis) and primary sclerosing cholangitis (PSC), both of which target the biliary tract. Cirrhosis is the common evolution of inflammatory hepatobiliary diseases, regardless of the target tissue, eventually leading to liver failure. However, pathogenesis and therapeutics may vary in between the spectrum of inflammatory hepatobiliary diseases.[1,2] This chapter describes the main characteristics of AIH, PBC, and PSC, with particular attention to clinical manifestations, pathogenesis, autoantibodies, and therapeutic options.

AUTOIMMUNE HEPATITIS

Epidemiology

Autoimmune hepatitis (AIH) is a severe liver disease affecting almost all age groups worldwide. Although an accurate estimate of its incidence is not available, it is believed be around 1 per 100 000 person-years, supposedly with higher incidence in Scandinavia.[3] AIH preferentially affects females, with a male:female ratio of 1:4 and a two-peak incidence during adolescence as well as in the age range 30–45 years.[2]

Pathogenesis

AIH is a chronic inflammatory disease of unknown etiology resulting from the immune-mediated destruction of hepatocytes secondary to a breakdown of immune tolerance against liver tissues.[4] Both genetics and environmental factors contribute to the development of the disease. Genetics certainly represents a strong risk factor for the development of AIH, as susceptibility to AIH is strongly influenced by the large and highly polymorphic human leukocyte antigen (HLA) region, which contains several genes related to the adaptive immune system[5] (Chapter 5). AIH has been associated with polymorphisms in the genes encoding the HLA class II DRB1 alleles, which have also been confirmed in genome-wide association studies (GWAS). In particular, expression of *HLA-DRB1*03:01* and *HLA-DRB1*04:01* alleles confers susceptibility to AIH-1 in European and North American populations, whereas the *DRB1*04:05* and *DRB1*04:04* alleles are linked to AIH susceptibility in Japan, Argentina, and Mexico.[6] Interestingly, *DRB1*13:02* and *DQB1*03:01* alleles were found to be protective.[2] Moreover, certain HLA alleles have been associated with AIH clinical manifestations, response to treatment, and prognosis.

Environmental factors have also been linked to AIH development. Specifically, infections such as hepatitis C virus (HCV) may trigger a process of molecular mimicry leading to exaggerated immune responses against self-components with structural homology. In support of this hypothesis, it has been demonstrated that HCV shares high amino-acid sequence homology with the autoantigenic target of antiliver/kidney microsomal-1 (LKM-1) autoantibodies, and that 10% of HCV patients are seropositive for anti-LKM-1. Nonviral environmental factors include medications, in particular antibiotics (nitrofurantoin and minocycline), statins, and anti-tumor necrosis factor (TNF)-α agents (adalimumab and infliximab). However, the disease is slightly different from classical AIH and usually does not require long-term immunosuppressive treatment.[2]

From an immunological standpoint, liver damage is likely orchestrated by CD4 lymphocytes, which can differentiate into T-helper cell-1 (Th1), Th2, and Th17 cells (Chapter 16). These effector cells initiate a cascade of immune reactions largely determined by the cytokines they produce: (i) Th1 cells secrete mainly interleukin (IL)-2 and interferon (IFN)-γ; IFN-γ is considered the main orchestrator of tissue damage since it not only stimulates CD8 cells but also enhances the expression of HLA class I, induces the expression of HLA class II molecules on hepatocytes, and activates monocytes/macrophages, which in turn release IL-1 and TNF-α; (ii) Th2 cells produce IL-4, IL-10, and IL-13, cytokines that induce the maturation of B cells into plasma cells, with consequent production of autoantibodies; (iii) Th17 cells, which arise in the presence of transforming growth factor (TGF)-β and IL-6 and produce IL-17, IL-22, TNF-α, and the chemokine CCL- 20.[7] Th17 cells, which are crucial in PBC, are being evaluated also in AIH, as an increased number of Th17 cells has been reported in the peripheral blood and liver of patients with AIH compared with healthy controls.[8]

Clinical Features and Diagnosis

AIH is a complex disease with various manifestations, predominantly in the liver, but it is not limited to this organ. The onset is most frequently insidious, whereas 20–30% of patients present with an acute icteric hepatitis, which is consistently associated with peripheral hypergammaglobulinemia. Two types of AIH are distinguished primarily based on the autoantibody patterns: *i.e.,* AIH type 1, with antinuclear and/or antismooth muscle antibodies; and AIH type 2, with antiliver/kidney/microsomal type 1 antibody and/or antiliver cytosol type 1 antibody. Type 1 AIH (AIH-1) can affect people of any age and sex, and it is associated with expression of *HLA-DRB1*03:01*. AIH-1 patients are more likely to be male, to present with high immunoglobulin G (IgG) levels, and to be positive for antinuclear antibodies/

anti-SMA. Also, AIH-1 may deteriorate despite corticosteroid treatment and is more likely to progress to liver transplantation.[2] Type 2 AIH (AIH-2) affects primarily girls and young women, and it has been linked to alleles encoding the DR3 (*DRB1*0301*) and DR7 (*DRB1*07:01*) molecules, and to positivity for anti-LMK antibodies.[2]

Blood tests usually show marked elevation of aminotransferases, usually with an increase in the cholestatic pattern (increased alkaline phosphatase, bilirubin). Furthermore, serum globulins, particularly of the γ type, are commonly increased in AIH cases, regardless of the histological stage.

Lastly and most importantly, serum autoantibodies are invariably positive (defined as titer >1 : 80) in patients with AIH. Antinuclear antibodies (ANAs) and/or antismooth-muscle (SMAs, directed against actin) autoantibodies are more frequently detected in AIH-1, whereas anti-LKM-1 antibodies (directed at cytochrome P450 2D6 or UDP-glucuronosyltransferases) predominate in AIH-2.

Without treatment, the clinical course of AIH is characterized by high mortality, with 5- and 10-year survival rates estimated as 50% and 10%, respectively. However, the use of corticosteroids has dramatically improved the course of the disease, with a 10-year survival rate exceeding 90%.[2] Complications associated with AIH are similar to those of other progressive liver diseases, as chronic hepatitis can evolve to cirrhosis and ultimately to hepatocellular carcinoma (HCC), despite the use of immunosuppressive therapy. At the time of diagnosis, approximately 30% of adult patients have histological evidence of cirrhosis; when appropriately treated, however, only a small number of patients develop cirrhosis during follow-up if inflammation is resolved. The occurrence of HCC in patients with AIH is a rare event and develops only in long-standing cirrhosis. In the absence of solid data on large numbers of cases, the incidence of primary liver neoplasia should be regarded as similar to other nonviral cases of cirrhosis.

There is no single diagnostic test for AIH; therefore diagnosis is based upon several indicative clinical, serological, biochemical, and histological findings. The presence of other causes of liver disease must also be excluded. A set of diagnostic criteria has been established, and subsequently updated, with high sensitivity and specificity (Table 76.1). Factors taken into account in this system include sex, plasma biochemical variables, serum autoantibodies, liver histology, possible cofactors (drugs, alcohol, viruses), and response to medical treatment.[9]

TABLE 76.1 Diagnostic Criteria for Autoimmune Hepatitis

Criteria	Points
Sex	
Male	+2
Female	0
Ratio of alkaline phosphatase vs. AST/ALT	
>3.0	+3
1.5–2.0	+2
1.0–1.5	+1
<1.0	0
Autoantibodies (ANA, SMA, LKM-1) titer	
>1 : 80	+3
1 : 80	+2
1 : 40	+1
<1 : 40	0
AMA	
Positive	–4
Negative	0
Seropositivity for other autoantibodies	+2
Viral hepatitis markers	
Negative	+3
Positive	–3
History of drug use	
Yes	–4
No	+1
Average alcohol consumption (g/day)	
<25	+2
>60	–2
Presence of genetic factors (HLA, DR3 or DR4)	+1
Presence of other autoimmune disorders	+2
Liver histology	
Interface hepatitis	+3
Predominant lymphocytic infiltrate	+1
Rosetting of liver cells	+1
None of the above	–5
Biliary changes	–3
Other changes	–3
Response to therapy	
Complete	+2
Relapse	+3

A score >15 or >17 indicates a definite diagnosis of autoimmune hepatitis pre- or posttreatment, respectively. On the other hand, scores of 10–15 and 12–17 indicate a probable diagnosis, pre- or posttherapy, respectively. AMA, antimitochondrial autoantibodies; ANA, antinuclear antibodies; HLA, human leukocyte antigen; LKM-1, antiliver–kidney microsomal antibodies; SMA, antismooth-muscle antibodies.

Serum Autoantibodies

Autoantibodies represent a critical feature of the disease and may guide the diagnosis. In 2004 the International Autoimmune Hepatitis Group established procedures and reference guidelines for more reliable serum autoantibody testing to overcome the lack of standardization.[10] Not only serum ANA, SMA, and LKM should be evaluated, but other sets of autoantibodies should also be tested in suspected cases, including those against liver-cytosol type 1 (LC1), perinuclear antineutrophil cytoplasmic antibodies (pANCA), soluble liver antigen/liver-pancreas antigen (SLA/LP), and the asialoglycoprotein receptor.[11] Finally, other less specific autoantibodies are also detected in a subgroup of patients, *i.e.*, anticardiolipin, chromatin, and *Saccharomyces cerevisiae*, but these are of limited clinical significance.

ANAs were the first autoantibodies observed in AIH sera more than 50 years ago and remain the most sensitive marker of AIH, producing most frequently a homogeneous or speckled pattern. However, the test is not specific for AIH, since ANA positivity is not uncommon in sera from patients with viral or other autoimmune liver diseases, as well as in as many as 15% of healthy subjects, especially in older age groups.[12] Similarly to most autoantibodies, ANA positivity, pattern, or titer does not reflect various AIH phenotypes; nor can these factors predict the disease's natural history.

Serum SMAs are autoantibodies reacting with different proteins (actin, tubulin, vimentin, desmin, cytokeratins) of the cytoskeletal components (microfilaments, microtubuli, intermediate filaments). Their presence characterizes both autoimmune (AIH-1, celiac disease) and viral diseases (chronic hepatitis C, infectious mononucleosis). When detected at high titers (>1 : 80), they are considered a sensitive marker for AIH-1, being found in up to 80% of cases. A recent study showed that anti-SMA-T/G (peritubular/glomerular staining) positive subjects with normal liver function are at low risk of progression to AIH, whereas positive SMA and raised ALT (>55 IU/L) indicate higher risk, with a positive predictive value of 22%.[13]

Serum autoantibodies against LKM-1 are the main serological markers of AIH-2 and recognize the proximal renal tubule and hepatocellular cytoplasm. The 50-kDa autoantigen was identified as the cytochrome P450 2D6 (*CYP2D6*). Interestingly, the sequence between amino acids 316 and 327, which is most likely exposed on the surface of the molecule, appears to be a region capable of differentiating LKM-1 activity in AIH and HCV and may represent a key target for autoimmunity. The mechanisms of onset remain enigmatic, and solid evidence of a causative role of hepatitis C virus cross-reactivity is still awaited. Similarly, the pathogenic role of anti-LKM-1 antibodies and their prognostic significance are debated, despite the development of AIH-2 in animal models after immunization with human CYP2D6 or with adenoviruses in mice transgenic for human *CYP2D6.* Finally, two other types of serum anti-LKM have been described in patients with ticrynafen-associated hepatitis (anti-LKM-2, directed against CYP2C9) and in 10% of patients with type 2 AIH (anti-LKM-3, directed against uridine-diphosphate glucuronosyl transferase 1A [UGT1A]), either alone or in combination with LKM-1 antibodies.

Anti-SLA/LP antibodies are detectable by radioimmunoassay and enzyme-linked immunosorbent assay, but not by immunofluorescence, and are directed against various epitopes of a UGA tRNA suppressor. Serum anti-SLA/LP antibodies are occasionally found in patients with AIH who are negative for ANA, SMA, and anti-LKM and are cumulatively detected in 10–30% of cases of AIH-1 and -2. Recent data indicate that anti-SLA/LP antibodies are also detectable in subgroups of pediatric patients with autoimmune cholangitis or in adult patients with HCV infection when tested with sensitive methods.

Anti-LC1 antibodies are detected by indirect immunofluorescence in sera from up to 50% of patients with type 2 AIH and less frequently in type 1 AIH or chronic hepatitis C. Importantly, however, anti-LC1s are the only detectable markers in 10% of AIH cases. The LC1 autoantigen is the liver formiminotransferase cyclodeaminase, an enzyme involved in folate metabolism. Interestingly, serum anti-LC1 antibodies correlate with AIH severity and progression.

Antibodies to the asialoglycoprotein receptor are observed in up to 90% of all patients with AIH and often coexist with other autoantibodies while lacking specificity for the disease. Similar to anti-LC1, however, antiasialoglycoprotein titers are associated with more florid inflammatory disease activity and with less effective treatment responses.

Finally, antibodies to neutrophil cytoplasmic antigens (pANCA) can be detected by indirect immunofluorescence in sera from patients with AIH-1; they also can be detected in a subgroup of patients with PSC or chronic viral hepatitis.

Histology

The role of liver histology in the management of AIH remains critical, and all patients with suspected cases should undergo a liver biopsy. In fact, although no typical feature can prove the diagnosis, histology remains the gold standard for grading and staging, particularly to determine the response to therapy. Common findings include periportal hepatitis with lymphocyte and plasma cell infiltrate and piecemeal necrosis. Fibrosis usually is observed, and bridging necrosis ultimately indicates advanced disease evolving into frank cirrhosis. Importantly, the presence of granulomas, bile duct damage, or iron or copper accumulation should not be overlooked, since these signs point toward other diagnoses. On the other hand, steatosis is a nonspecific finding that does not rule out AIH.[9]

Therapy

In contrast to other autoimmune liver diseases, immunosuppressants are the treatment of choice for AIH, based on the good response in terms of biochemistry, histology, and survival.[14] Corticosteroids, in particular prednisone, represent the first-line of treatment in monotherapy or in combination with azathioprine, inducing remission (*i.e.,* normal transaminases and IgG levels) in over 80% of patients, regardless of the presence of cirrhosis.[9] Once achieved, remission can be maintained with azathioprine alone. Relapses after therapy discontinuation are common since only 20% of patients remain in sustained remission. It should be noted, however, that subgroups of patients manifest disease progression (approximately 10%) or are intolerant to standard therapy (13%). In such patients, other drugs have been anecdotally tried, including methotrexate, cyclophosphamide, tacrolimus, ursodeoxycholic acid, cyclosporine, and mycophenolate mofetil, the latter two being those most frequently reported as alternative medications. Liver transplantation is the ultimate treatment for AIH patients presenting with acute liver failure or end-stage chronic liver disease and for those with hepatocellular carcinoma that meet the transplant criteria. Although liver transplantation for AIH is very successful, it should be noted that AIH may recur after transplant.[9] Patients with AIH undergoing liver transplant have overall 5- and 10-year survival rates of 90% and 75%, respectively, although infectious complications and disease recurrence are common.[15]

KEY CONCEPTS

- Autoimmune hepatitis (AIH) is a severe autoimmune disease associated with high morbidity and mortality, especially due to the development of cirrhosis and possibly hepatocellular carcinoma.
- Autoantibodies represent a distinctive feature of AIH and can also aid in the definition of AIH subsets.
- Liver biopsy is helpful, and histology remains the gold standard for grading and staging, particularly to determine the response to therapy.

CLINICAL PEARLS

- Autoimmune hepatitis (AIH) may cause an increase in aminotransaminases and in cholestatic markers.
- Autoantibody testing is helpful. Antinuclear antibodies (ANAs), SMAs, and antiliver/kidney microsomal (LKM) are the most important, but other sets of autoantibodies also should be tested in suspected cases, in particular anti-LC1, perinuclear antineutrophil cytoplasmic antibodies (pANCA), soluble liver antigen/liver-pancreas antigen (SLA/LP), and the asialoglycoprotein receptor antibody.
- Diagnostic criteria are available and provide good sensitivity and specificity.

THERAPEUTIC PRINCIPLES

- Glucocorticoids represent the cornerstone of autoimmune hepatitis (AIH) therapy as monotherapy or in combination with azathioprine.
- Other immunosuppressants, *i.e.,* methotrexate, tacrolimus, cyclosporine, and mycophenolate mofetil, have been used in patients who are intolerant or refractory to standard therapy.
- Liver transplantation is very successful; however, AIH may recur after transplant.

PRIMARY BILIARY CHOLANGITIS

Nomenclature

Recently the nomenclature for primary biliary cirrhosis has been shifted toward primary biliary cholangitis (PBC) to correct the inaccurate description and remove the cirrhosis stigma as well as all the misunderstandings, disadvantages, and discriminations emanating from this misnomer. This change was necessary after the dramatic improvement in PBC diagnosis, prognosis, and treatment. Currently, an early diagnosis of PBC can be established with the more accurate measurements of markers of cholestasis and with improvement in the detection of the classic serological hallmark, antimitochondrial antibodies (AMAs). In parallel, treatment was improved with the introduction of orthotopic liver transplantation and ursodeoxycholic acid (UDCA, 13–15 mg/kg daily) therapy. At present, two of three patients diagnosed with primary biliary cirrhosis and treated with UDCA have an expected survival equal to that of the general population, and only a minority will ever develop cirrhosis.[16]

Epidemiology

PBC is a chronic autoimmune cholestatic disease characterized by high-titer serum AMA and immune-mediated destruction of the small- and medium-sized intrahepatic bile ducts.[1] It affects women more frequently, with a female:male ratio of 9 : 1; the average age at diagnosis is within the fifth and sixth decades of life. Geographically, it presents with widely variable prevalence; the highest rate has been described in Minnesota, with a point prevalence of 402 per million in the general population.[17,18]

Pathogenesis

As with essentially every autoimmune disease, genetics and environmental factors interact to lead to PBC. Genetic factors play a significant role in PBC susceptibility, as suggested by the fact that it is more frequent in relatives of affected individuals; the term "familial PBC" has been coined to indicate families that have more than one case. In our experience, 6% of cases have a first-degree relative that is also affected,[19] while AMA may be positive in first-degree relatives and offspring of patients with PBC without any sign of disease, thus indirectly suggesting the existence of a strong genetic predisposition.[20] More importantly, the concordance rate observed among monozygotic twins for PBC is 63%, among the highest reported in autoimmunity, reinforcing the idea that genetic risk factors play an important role.[21] As for other autoimmune disorders, genetic factors are not limited to a single gene but to a complex multigene trait. The relevance of the multifactorial genetic basis in PBC has been reinforced by recent observations including the high concordance rate among monozygotic (identical) twins[21] as well as by reports that lymphocytes from women with PBC preferentially use one of their X chromosomes, rather than the usual random X chromosome inactivation observed in healthy women.[22,23] PBC has been associated with the *HLA-DRB1*08* allele but also with two protective alleles, *HLA-DRB1*11* and *-DRB1*13*. Polymorphisms in HLA class II loci were confirmed to play a key role by GWAS. Notably, this study demonstrated that genes coding for interleukin-12 and its receptor also confer risk for PBC.[24]

However, it can be hypothesized that other factors may contribute to genetics as complementary modes of pathogenetic import, including epigenetics, exposure to environmental factors, etc. Infections and xenobiotics play a key role in increasing the susceptibility to PBC. Urinary tract and vaginal infections are more frequent in PBC than in controls, with *E. coli* being the main etiological agent. The mechanisms linking infections and autoimmune biliary tract damage are suspected to involve molecular mimicry. Besides *E. coli*, a number of other bacteria have been examined as cross-reactive agents in PBC, including *Proteus mirabilis, Klebsiella pneumoniae, Staphylococcus aureus, Neisseria meningitidis, Salmonella minnesota, Mycobacterium gordonae, and Trypanosoma brucei. Novosphingobium aromaticivorans,* a ubiquitous xenobiotic-metabolizing gram-negative bacterium, has been proposed to be the ideal candidate for PBC development because it contains proteins with the highest degree of homology with the major epitope of PDC-E2 while having the ability to metabolize organic compounds and estrogens. Moreover, *N. aromaticivorans* can elicit antibody reactivity in PBC patients but not in controls. Regarding viral infections, a novel human beta-retrovirus was found in perihepatic lymph nodes and other biological samples from patients with PBC; however, the data have remained inconclusive.[24] Xenobiotics are foreign compounds that may trigger an autoimmune response to self-proteins, which may modify their molecular structures or complex to self or nonself proteins to generate neoantigens. In this view, it was shown that 2-nonynoic acid, an organic compound found in nail polish, was able to elicit AMA production in patients with PBC once they were attached to the major mitochondrial epitope backbone.[25]

The complex pathogenesis of PBC, however, is not fully explained by genetics and environmental factors; as suggested for most other autoimmune diseases, epigenetic modifications could represent the link between genetic and environmental factors influencing the onset and evolution of PBC. Epigenetic dysregulation results in the overexpression of certain genes in several key immune cells.[26] However, conclusive data implicating epigenetic modifications in the pathogenesis of PBC are not yet available.

Clinical Features and Diagnosis

In the early phase of disease, symptoms of PBC are mainly fatigue and pruritus, while physical findings may include skin hyperpigmentation, hepatosplenomegaly, and (rarely) xanthelasmas. End-stage symptoms are indistinguishable from those of other types of liver cirrhosis and include ascites, jaundice, hepatic encephalopathy, and upper digestive bleeding. Fatigue is a nonspecific symptom present in 70% of PBC patients and is often overlooked, particularly in middle-aged women. The severity of fatigue is independent of the stage of PBC or its other features (pruritus or severe cholestasis), and it appears not to depend on psychiatric factors. Pruritus is similarly frequent, being present in 70% of patients with PBC and jaundice. Pruritus might long precede jaundice onset and typically worsens at night, after contact with wool, or in warm climates. Portal hypertension is frequently found in patients with PBC and, importantly, does not imply the presence of liver cirrhosis. Over half of untreated patients eventually develop portal hypertension over a 4-year period. Prevention and treatment for PBC-associated portal hypertension are the same as for other chronic liver diseases. Metabolic bone disease is present in PBC when comparing sex- and age-matched healthy individuals. Hyperlipidemia is common in up to 85% of patients with PBC; both elevated serum cholesterol and triglyceride levels are encountered. Interestingly, however, such alterations are not accompanied by a proportionally increased incidence of cardiovascular events or atherosclerosis and do not correlate with disease stage. Autoimmune diseases may overlap with PBC,

TABLE 76.2 **Diagnostic Criteria for Primary Biliary Cholangitis**

Parameters
Elevated ALP > 2 × ULN or GGT > 5 × ULN
AMA positivity
Chronic granulomatous cholangitis on liver biopsy

Diagnosis defined in the presence of at least 2 out of 3 criteria.
ALP, alkaline phosphatase; AMA, antimitochondrial antibodies; ULN, upper limit of normal; GGT, γ-glutamyltransferase.

most commonly Sjögren syndrome, Raynaud phenomenon, autoimmune thyroid disease, scleroderma, and systemic lupus erythematosus, whereas the prevalence of rheumatoid arthritis does not differ when PBC cases are compared with healthy controls. Patients affected by both PBC and scleroderma have a less aggressive liver disease, which suggests an active interaction between the two conditions. As with other types of cirrhosis, end-stage PBC can be complicated by the occurrence of HCC, and patients with intense nodular liver structure on ultrasound should be monitored by computed tomography. Importantly, PBC is not associated with cholangiocarcinoma (CCA) or breast cancer.[1]

The diagnosis of PBC is based on the presence of two out of three internationally accepted criteria, *i.e.*, detectable serum AMA (titer >1 : 40), increased enzymes indicating cholestasis (*i.e.*, alkaline phosphatase) for longer than 6 months, and a compatible or diagnostic liver histology (Table 76.2).[27] In a large number of cases (20–60%), the diagnosis of PBC is established in the absence of symptoms indicating a liver condition or cholestasis, and the proportion of asymptomatic cases at diagnosis has been steadily increasing over the past decade. At presentation, PBC is suspected if a biochemical cholestatic pattern (increased plasma alkaline phosphatase or γ-glutamyltransferase) is present with no similar increase in plasma aminotransferase levels. Serum IgMs are typically elevated in PBC cases, with no correlation with AMA titers or levels of other Ig subtypes. Once cirrhosis has developed, biochemical alterations are similar to those with other types of cirrhosis. The progression of PBC varies widely, and the factors influencing the severity and progression of the disease are largely unknown. Having symptoms at presentation is considered the major factor determining survival rates of patients with PBC. In fact, asymptomatic PBC is accompanied by 10-year survival rates similar to those of the general population. On the other hand, 67% of precirrhotic patients will develop liver cirrhosis over a 7-year observation period, while 70% of asymptomatic patients will develop symptoms. Patients with symptomatic PBC show a more rapid progression to late-stage disease and a worse prognosis than their asymptomatic counterparts, with survival times among symptomatic subjects within 6–10 years.[28] Older age at diagnosis and signs of advanced disease (clinical, histological, or biochemical) are associated with a worse prognosis. The establishment of accurate prognostic models to predict survival in patients with PBC is of obvious importance in clinical practice. The model based on the Mayo score is the only one validated and is the most widely utilized[29]; it is calculated based on clinical (age, presence of ascites) and biochemical variables as represented by cholestasis (bilirubin levels) and liver function (prothrombin time, albumin). However, this model has its limitations, as it is a static representation of a dynamic entity and has a lower accuracy for patients with early disease.

Serum Autoantibodies

AMAs are highly specific for PBC and can be detected in nearly 100% of patients when sensitive diagnostic methodologies based on recombinant antigens are used.[24] In most clinical settings, however, indirect immunofluorescence techniques are used for initial screening of cases and might provide falsely positive or negative results. AMAs are directed against components of the 2-oxoacid dehydrogenase (2-OADC) family of enzymes within the mitochondrial respiratory chain, most frequently the E2- and E3-binding protein (E3BP) components of the pyruvate dehydrogenase complex and the E2 components of the 2-oxo glutarate dehydrogenase and branched-chain 2-oxo acid dehydrogenase complexes. For all three antigens, epitopes contain the motif DKA, with lipoic acid covalently bound to the lysine (K) residue. The role of lipoic acid in epitope recognition by AMA is unclear. A direct pathogenic role of AMA is debatable, since no clinical correlation can be found and animal models developing serum AMA do not develop PBC-like liver lesions. Autoantibodies other than AMA can be found in 76% of PBC patients.[30] ANA can be found in 50% of PBC patients, with the most common patterns being "nuclear rim" or "multiple nuclear dots." The pattern is based on the recognition by the autoantibodies of gp210 and nucleoporin 62 (within the nuclear pore complex) and Sp100 and promyelocytic leukemia protein (PML) (possibly also cross-reacting with small ubiquitin-like modifiers, SUMO), respectively.[31] ANA-positive patients are more frequently AMA-negative, possibly because of the lack of a masking effect of these latter antibodies. Similarly, the pathogenic role of ANA in PBC remains enigmatic, although cross-sectional and longitudinal data demonstrate an association between ANA positivity and a worse prognosis.[32] Finally, patients with PBC and limited systemic sclerosis have detectable serum anticentromere antibodies in 10–15% of cases. Monitoring autoantibody titers does not correlate with PBC severity or clinical outcomes, whereas a change in the sp100 autoantibody level may have prognostic utility with respect to the development of fibrosis on liver biopsy.

Histology

PBC histology is classified into four stages (Fig. 76.1). Stage I manifests with portal tract inflammation with predominantly lymphoplasmacytic infiltrates, resulting in vanishing septal and interlobular bile ducts (diameter < 100 μm). At this stage, bile duct obliteration and granulomas (possibly found at all stages) are strongly suggestive of PBC. In stage II a periportal inflammatory infiltrate is observed, and signs of cholangitis, granulomas, and florid proliferation of ductules are typical. Stage III demonstrates septal or bridging fibrosis, with ductopenia (over half of the visible interlobular bile ducts having vanished) and copper deposition in periportal and paraseptal hepatocytes visible. Stage IV corresponds to frank cirrhosis. The observation of eosinophils in the portal tract is considered a specific finding in PBC histology.

Novel noninvasive biomarkers are under evaluation for predicting liver histology, and in this view the aminotransferase to platelet ratio index (APRI) and fibrosis index on the basis of the four factors (FIB-4) scores are statistically different between groups and may be a predictor of advanced disease in PBC.[33]

Therapy

PBC treatment is currently based on UDCA, which is the only approved drug for this rare disease. Its mechanism of action is

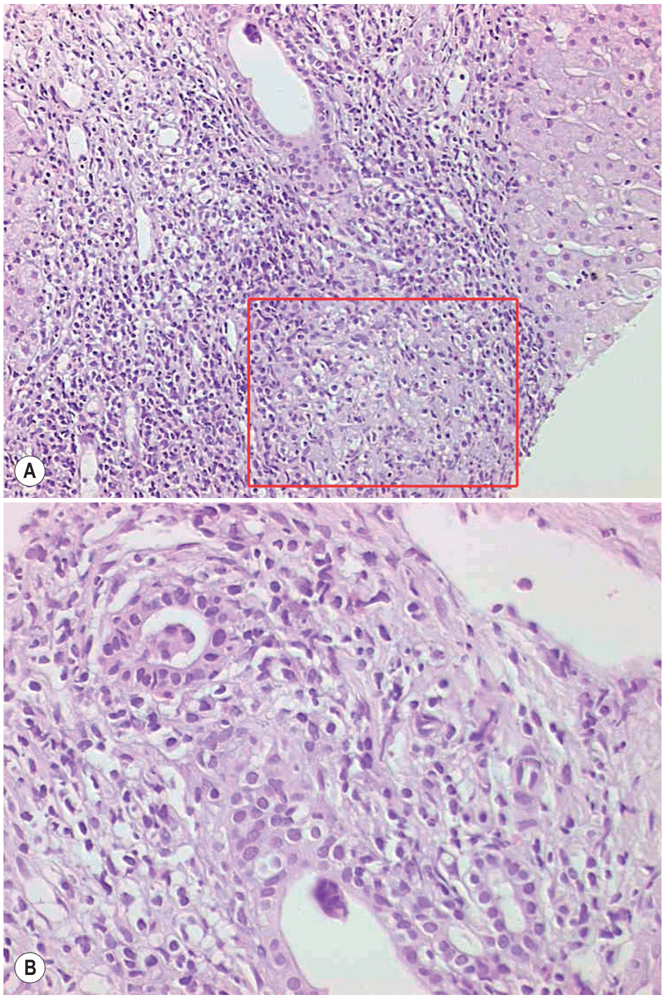


FIG 76.1 Histological Findings in Early Stages of Primary Biliary Cirrhosis, i.e., Nonsuppurative Destructive Cholangitis, Following Hematoxylin and Eosin Staining. (A) Mixed lymphocytic and plasma cell periductular inflammation with bile duct infiltration and granulomatous reaction (square). Magnification × 200. (B) Detail of bile duct disruption with lymphocytic and plasmacellular periductular and intraepithelial infiltration. Magnification × 400.

incompletely understood. Depending on the various phases of the disease, various therapies might be effective.

During the early phase of the disease, glucocorticoids might be effective; however, safety concerns about the long-term use of such medications arise. Budesonide, due to its high first-pass metabolism, has minimal systemic adverse effects; at 6–9 mg daily, it has been demonstrated superior to UDCA both in terms of histology and biochemical markers. Other immunosuppressants, such as methotrexate and azathioprine, have also been suggested, and there is evidence supporting the use of the latter in PBC with autoimmune hepatitis overlap syndrome. In recent years, because of better understanding of PBC pathogenesis, new targeted therapies have been tested. Considering a possible role of the IL-17/23 axis, ustekinumab, a monoclonal antibody against the p40 subunit of IL-23, has been tested. Although associated with a modest decrease in alkaline phosphatase after 28 weeks of therapy, ustekinumab did not otherwise change alkaline phosphatase.[34] Other therapies targeting T-cell costimulation, *e.g.*, cytotoxic T lymphocyte antigen-4 (CTLA-4) Ig (abatacept) or the CD40-antagonist FFP104, are under investigation.[35] When the disease has already progressed and bile is accumulating, obeticholic acid (OCA), an analog of CDCA with a much higher affinity to the farnesoid X receptor, has been shown to decrease bile synthesis, promote secretion, and induce liver regeneration in animal models. In a phase III trial, OCA administered with UDCA or as monotherapy for 12 months resulted in a decrease in alkaline phosphatase and total bilirubin levels that differed significantly from the changes observed with placebo.[35]

Ultimately, UDCA represents the cornerstone therapy of PBC. Doses ranging from 13 to 15 mg/kg are currently used for optimum bile enrichment and in 50% of patients normalize alkaline phosphatase. Other immunosuppressive treatments should be started only in combination with UDCA.

Liver transplantation is the ultimate treatment for end-stage PBC, with survival rates of 92% and 85% at 1 and 5 years after transplant, respectively. Recurrence is common, and its rates seem to be influenced by certain immunosuppressive regimens. The use of UDCA in cases of recurrence is safe and recommended.

KEY CONCEPTS

- Primary biliary cholangitis (PBC) is a chronic autoimmune cholestatic disease affecting preferentially women in the fifth to sixth decades of life.
- Genetics plays a strong role in PBC pathogenesis, as suggested by familial clustering and the high concordance in monozygotic twins.
- Antimitochondrial antibodies (AMAs) are highly specific for PBC and can be detected in nearly 100% of patients.

CLINICAL PEARLS

- Fatigue and pruritus represent the most frequently observed symptoms, which may be disabling for patients.
- Primary biliary cholangitis (PBC) can be frequently associated with other autoimmune diseases, *e.g.*, Sjögren syndrome and systemic sclerosis.
- PBC patients should be observed for metabolic bone disease.

THERAPEUTIC PRINCIPLES

- Ursodeoxycholic acid (UDCA) (13–15 mg/kg) remains the cornerstone therapy for primary biliary cholangitis and is the only approved drug.
- Glucocorticoids may be effective in early phases of the disease.
- Biological therapies are under investigation, *e.g.*, ustekinumab and abatacept.
- Obeticholic acid appears to be effective in reducing cholestatic biomarkers.

PRIMARY SCLEROSING CHOLANGITIS

Primary sclerosing cholangitis (PSC) is a progressive cholestatic liver disease of unknown etiology presenting with autoimmune features and associated with significant morbidity and mortality. Unlike PBC, PSC can affect all tracts of the biliary tree, including the extrahepatic bile ducts.

Epidemiology

The prevalence of PSC is approximately 10 in 100 000 in northern Europe[4,5] and the United States, while it is far less common in southern Europe and Asia. Recent data from Olmsted County, Minnesota, reported a disease prevalence of 20.9 per 100 000 men and 6.3 per 100 000 women.[36] Epidemiological data indicate that annual incidence rates are not increasing over time despite earlier ages at diagnosis, and survival is possibly longer, similar to what was observed in PBC.

In contrast to the vast majority of autoimmune diseases, PSC is more commonly diagnosed in men, with a female:male ratio estimated as 1 : 2, with a preference for 30- to 40-year-olds. PSC has a strong association with inflammatory bowel disease. Approximately 60–80% of patients with PSC present with inflammatory bowel disease, of which 87% have ulcerative colitis and 13% have Crohn disease.[37]

Pathogenesis

The etiopathogenesis of PSC is unknown, but evidence is growing that (auto) immune-mediated mechanisms play a role. The notion of an autoimmune pathogenesis is supported by the frequent association with inflammatory bowel disease (IBD), the presence of serum autoantibodies, and the reported HLA associations. Genetic risk determines PSC susceptibility, including HLA haplotypes on chromosome 6p21, the inhibitory HLA-C2 killer-immunoglobulin receptor (KIR) ligand, and HLA-C1 homozygosity. Also, a positive and negative association with *HLA-HLDRB1*15* and *-DRB1*07,* respectively, has been reported. A recent study on Italian patients with PBC did not find any association with *DRB1*03, *04* or **13:01* alleles typically detected in Northern Europe.[38] Family and twin studies are not available for PSC; however, a GWAS has reported the HLA region as the only major association, albeit in a minority of patients and sometimes overlapping with IBD-associated genes. The innate immune system and microorganisms (possibly derived from an IBD-affected gut) may participate in the onset and/or perpetuation of disease. It has been proposed that cholangiocytes are first activated by bacterial stimuli in the presence of gut-specific chemokines and endothelial cell adhesion molecules in the tissue microenvironment. Also, gut-primed T cells can migrate into the portal tracts and peribiliary spaces to form focal lesions. Finally, chronic inflammation and progressive fibrosis of the biliary epithelium lead to chronic cholestasis secondary to vanishing bile ducts, ultimately causing biliary cirrhosis.

Clinical Features and Diagnosis

PSC symptoms generally are nonspecific and include abdominal pain, jaundice, and fever in the case of bacterial cholangitis; at more advanced stages, symptoms include the characteristics of decompensated cirrhosis or neoplasia. Commonly, PSC is further complicated by episodic bacterial cholangitis, especially in the setting of biliary strictures. Lastly, subgroups of patients manifest the "small duct" variant or overlap syndrome.

The median time span from diagnosis to liver-related death or liver transplantation can be estimated at 18 years,[39] and the prognosis is influenced by the possible onset of CCA, which may be difficult to detect because biliary structures are already altered. It is important to distinguish small-duct PSC, as the natural history is relatively benign and only a minority (12%) of patients develop classical PSC.[40]

Due to the nonspecific symptoms, PSC currently is diagnosed most commonly in the absence of symptoms and during routine blood tests in healthy individuals or patients with IBD. In laboratory testing, PSC is characteristically accompanied by a biochemical cholestatic pattern, as indicated by elevated serum alkaline phosphatase and γ-glutamyltransferase. Liver function tests are normal until late stages. Autoantibodies are of limited use in the diagnosis of PSC due to low sensitivity and specificity. Only a limited percentage of patients (33%) has positive pANCA, which usually is found in IBD patients without PSC.

Imaging represents a useful diagnostic tool, as it may find the unique strictured and dilated tracts within the intrahepatic or extrahepatic bile ducts. Among the imaging techniques, endoscopic retrograde cholangiopancreatography (ERCP) and magnetic resonance cholangiopancreatography (MRCP) are currently considered equal for sensitivity, but their results are influenced by the operator's skill and experience.

Histology

Although not necessary for establishing the diagnosis, liver histology is essential when staging PSC or when the small-duct variant, an overlap syndrome, or CCA is suspected. The histological picture varies widely, from minimal alterations to cirrhosis with portal inflammation, concentric "onion skin" periductal fibrosis, and periportal fibrosis developing into septal and bridging necrosis. Fig. 76.2 illustrates the histological findings in two representative cases of early and advanced PSC.

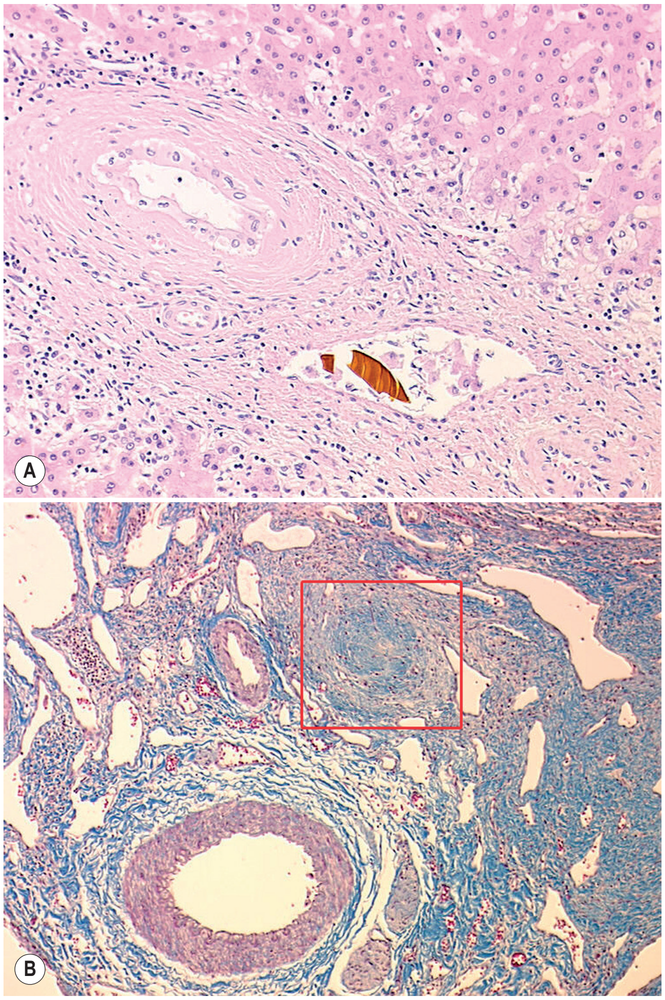


FIG 76.2 Histological Findings in Primary Sclerosing Cholangitis. (A) Early disease and periductular fibrosis. Magnification x 200, hematoxylin and eosin staining. (B) Advanced disease with cirrhosis and bile duct substitution by fibrous scar (square). Magnification × 200, Masson staining.

Therapy

The treatment of PSC includes medical and endoscopic measures as well as liver transplantation. UDCA has been investigated in several clinical trials, with conflicting results. Overall, the available evidence suggests that UDCA fails to induce a substantial change in the course of PSC, despite remaining the most prescribed drug. However, it appears that high-dose UDCA (20 mg/kg/day) might reduce the rate of progression and possibly prevent the development of colon cancer in patients with PSC and Ulcerative colitis (UC). Endoscopic measures are indicated to treat complicated PSC through the opening of short- and long-segment stenosis of the common bile duct and short-segment stenosis of the hepatic ducts near to the bifurcation. The treatment can be repeated over time once restenosis ensues, and resulting survival rates are higher than those of patients not treated endoscopically. Finally, PSC represents an important indication for liver transplantation since patients are commonly younger than those with other autoimmune liver diseases. Recurrence of disease is common and affects 20–40% of transplanted patients during prolonged follow-up. It is still debated whether medical treatment with UDCA in fact produces a longer survival.

KEY CONCEPTS

- Primary sclerosing cholangitis (PSC) is a chronic autoimmune cholestatic disease that can affect all tracts of the biliary tree, including the extrahepatic bile ducts.
- Unlike other patients with autoimmune diseases, the majority are male.
- PSC can be complicated by cholangiocarcinoma.

CLINICAL PEARLS

- Primary sclerosing cholangitis (PSC) is frequently associated with IBD (60–80% of patients).
- PSC is not associated with any specific autoantibodies; however, perinuclear antineutrophil cytoplasmic antibodies (pANCA) may be positive.
- Imaging is helpful in detecting PSC, which manifests with strictured and dilated bile ducts, both intra- and extrahepatic.

THERAPEUTIC PRINCIPLES

- Ursodeoxycholic acid (UDCA) (20 mg/kg) may be effective in slowing primary sclerosing cholangitis (PSC) progression.
- Endoscopic procedures are indicated to treat stenoses.
- Liver transplantation is highly effective, and PSC represents an important indication for such procedure, particularly due to the younger age of affected patients.

AUTOIMMUNE CHOLANGITIS

The term *autoimmune cholangitis* was first introduced to indicate AMA-negative PBC, possibly with serum ANA. However, more recently this category has been expanded to include (i) serum ANA and/or SMA positivity and/or hypergammaglobulinemia; (ii) serum AMA negativity by immunofluorescence; (iii) biochemical and/or histological features of cholestatic and hepatocellular injury; and (iv) exclusion of chronic viral, metabolic, or toxic liver disease. This definition possibly includes PBC with atypical presentation, small-duct PSC, idiopathic adulthood ductopenia, AIH with bile duct damage, concurrent AIH and small-duct PSC, and various transitional stages of the classic diseases. Consensus is awaited on this issue, and standardization of diagnostic criteria for overlap syndromes also has not been achieved.

OVERLAP SYNDROMES

As many as 18% of patients with autoimmune liver disease also present with features of a second autoimmune liver disease. An even higher percentage of patients has an additional autoimmune disease, *e.g.,* rheumatoid arthritis, Sjögren syndrome, or systemic sclerosis. These patients are considered to have overlap syndromes. Patients with hepatobiliary overlap syndromes usually present with both hepatocellular and cholangiocellular injury with biochemical and histological features of AIH and PBC or PSC. When not treated, these patients show a progressive course toward liver cirrhosis and failure. More specifically, AIH–PBC overlap syndrome is found in 10% of adults with AIH or PBC, and AIH–PSC overlap syndromes are found in 6–8% of children, adolescents, and young adults with AIH or PSC. Besides the existence of overlaps, in rare cases transitions are also possible from PBC to AIH, from AIH to PBC, or from AIH to PSC. The pathogenesis of the overlap syndromes is poorly understood, and few data are available regarding the clinical characteristics and outcomes of such patients. Thus the clinical management of overlap syndromes is based on single diseases, and medical treatment remains empiric. Thus UDCA is used for chronic cholestasis, and immunosuppressants (steroids and azathioprine) are used for AIH; for end-stage disease, liver transplantation is indicated.

IgG4-RELATED CHOLANGITIS

There have been numerous reports of a new clinical entity coined *autoimmune pancreatitis/IgG4-associated sclerosing cholangitis* with peculiar clinical and therapeutic characteristics. IgG4-related sclerosing cholangitis is more frequent in men than in women, and usually at an older age. The clinical features of the disease appear similar to those with CCA or PSC. IgG4-related sclerosing cholangitis has been associated with autoimmune pancreatitis, and the coexistence of both manifestations is suggestive of IgG4-related disease. The differential diagnosis with PSC may be difficult; however, elevated IgG4 serum concentration (up to 10% of PSC patients may have elevated levels) and longer stenosis found by imaging may support the diagnosis. Histology is important, as it may discover an IgG4-positive lymphoplasmacytic tissue infiltrate. Overall, the prognosis is considered to be good; however, some cases of portal hypertension and liver cirrhosis have been reported.

Corticosteroids induce a remarkable improvement in clinical and biochemical manifestations, and a dosage of 0.6 mg/kg/day of prednisolone should be started for 2–4 weeks and gradually tapered over 2–3 months. Immunosuppressant agents have been used effectively in anecdotal cases. The optimal treatment is still debated.

ON THE HORIZON

Inflammatory hepatobiliary diseases remain an enigmatic and interesting group of tissue-specific autoimmune diseases, especially due to the frequent overlap among the different entities. However, primary biliary cholangitis (PBC), primary sclerosing cholangitis (PSC), and autoimmune hepatitis (AIH) should be considered single diseases because the pathogenesis, clinical manifestations, prognosis, and therapeutic approaches are different. Autoantibodies may help the clinician in the diagnosis of the various diseases. However, it is important to remember that there are limitations in the use of autoantibodies, especially due to variability in testing procedures, nonstandardized methods, and reporting. Treatment is changing along with new approaches to targeted therapies. However, in most cases corticosteroids and ursodeoxycholic acid (UDCA) remain the cornerstones of therapy.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Selmi C, Bowlus CL, Gershwin ME, et al. Primary biliary cirrhosis. Lancet 2011;377:1600–9.
2. Liberal R, Krawitt EL, Vierling JM, et al. Cutting edge issues in autoimmune hepatitis. J Autoimmun 2016;75:6–19.
3. Jepsen P, Gronbaek L, Vilstrup H. Worldwide Incidence of Autoimmune Liver Disease. Dig Dis 2015;33(Suppl. 2):2–12.
4. Vergani D, Mieli-Vergani G. Cutting Edge Issues in Autoimmune Hepatitis. Clin Rev Allergy Immunol 2012;42:309–21.
5. Karlsen TH, Chung BK. Genetic Risk and the Development of Autoimmune Liver Disease. Dig Dis 2015;33(Suppl. 2):13–24.
6. Mells GF, Kaser A, Karlsen TH. Novel insights into autoimmune liver diseases provided by genome-wide association studies. J Autoimmun 2013;46:41–54.
7. Liberal R, Vergani D, Mieli-Vergani G. Update on Autoimmune Hepatitis. J Clin Transl Hepatol 2015;3:42–52.
8. Zhao L, Tang Y, You Z, et al. Interleukin-17 contributes to the pathogenesis of autoimmune hepatitis through inducing hepatic interleukin-6 expression. PLoS ONE 2011;6:e18909.
9. Liberal R, Grant CR, Longhi MS, et al. Diagnostic criteria of autoimmune hepatitis. Autoimmun Rev 2014;13:435–40.
10. Manns MP, Vogel A. Autoimmune hepatitis, from mechanisms to therapy. Hepatology 2006;43:S132–44.
11. Cancado EL, Abrantes-Lemos CP, Terrabuio DR. The importance of autoantibody detection in autoimmune hepatitis. Front Immunol 2015;6:222.
12. Selmi C, Ceribelli A, Generali E, et al. Serum antinuclear and extractable nuclear antigen antibody prevalence and associated morbidity and mortality in the general population over 15 years. Autoimmun Rev 2016;15:162–6.
13. Mullin S, Rabah R, Malas S, et al. Autoimmune Hepatitis Type 2 Associated With Positive Antimitochondrial Antibodies: An Overlap Syndrome? Clin Pediatr (Phila) 2016;55:479–82.
14. Montano Loza AJ, Czaja AJ. Current therapy for autoimmune hepatitis. Nat Clin Pract Gastroenterol Hepatol 2007;4:202–14.
15. Schramm C, Bubenheim M, Adam R, et al. Primary liver transplantation for autoimmune hepatitis: a comparative analysis of the European Liver Transplant Registry. Liver Transpl 2010;16:461–9.
16. Beuers U, Gershwin ME, Gish RG, et al. Changing nomenclature for PBC: from 'cirrhosis' to 'cholangitis. Hepatology 2015;62:1620–2.
17. Selmi C. The worldwide gradient of autoimmune conditions. Autoimmun Rev 2010;9:A247–50.
18. Podda M, Selmi C, Lleo A, et al. The limitations and hidden gems of the epidemiology of primary biliary cirrhosis. J Autoimmun 2013;46:81–7.
19. Gershwin ME, Selmi C, Worman HJ, et al. Risk factors and comorbidities in primary biliary cirrhosis: a controlled interview-based study of 1032 patients. Hepatology 2005;42:1194–202.
20. Lazaridis KN, Talwalkar JA. Clinical epidemiology of primary biliary cirrhosis: incidence, prevalence, and impact of therapy. J Clin Gastroenterol 2007;41:494–500.
21. Selmi C, Mayo MJ, Bach N, et al. Primary biliary cirrhosis in monozygotic and dizygotic twins: genetics, epigenetics, and environment. Gastroenterology 2004;127:485–92.
22. Miozzo M, Selmi C, Gentilin B, et al. Preferential X chromosome loss but random inactivation characterize primary biliary cirrhosis. Hepatology 2007;46:456–62.
23. Invernizzi P, Miozzo M, Battezzati PM, et al. Frequency of monosomy X in women with primary biliary cirrhosis. Lancet 2004;363:533–5.
24. Invernizzi P, Selmi C, Gershwin ME. Update on primary biliary cirrhosis. Dig Liver Dis 2010;42:401–8.
25. Rieger R, Gershwin ME. The X and why of xenobiotics in primary biliary cirrhosis. J Autoimmun 2007;28:76–84.
26. Marzorati S, Lleo A, Carbone M, et al. The epigenetics of PBC: the link between genetic susceptibility and environment. Clin Res Hepatol Gastroenterol 2016;40:650–9.
27. Ali AH, Carey EJ, Lindor KD. Diagnosis and management of primary biliary cirrhosis. Expert Rev Clin Immunol 2014;10:1667–78.
28. Pares A, Rodes J. Natural history of primary biliary cirrhosis. Clin Liver Dis 2003;7:779–94.
29. Grambsch PM, Dickson ER, Kaplan M, et al. Extramural cross-validation of the Mayo primary biliary cirrhosis survival model establishes its generalizability. Hepatology 1989;10:846–50.
30. Agmon-Levin N, Shapira Y, Selmi C, et al. A comprehensive evaluation of serum autoantibodies in primary biliary cirrhosis. J Autoimmun 2010;34:55–8.
31. Selmi C, Zuin M, Bowlus CL, et al. Anti-mitochondrial antibody-negative primary biliary cirrhosis. Clin Liver Dis 2008;12:173–85, ix.
32. Invernizzi P, Selmi C, Ranftler C, et al. Antinuclear antibodies in primary biliary cirrhosis. Semin Liver Dis 2005;25:298–310.
33. Olmez S, Sayar S, Avcioglu U, et al. The relationship between liver histology and noninvasive markers in primary biliary cirrhosis. Eur J Gastroenterol Hepatol 2016;28:773–6.
34. Hirschfield GM, Gershwin ME, Strauss R, et al. Ustekinumab for patients with primary biliary cholangitis who have an inadequate response to ursodeoxycholic acid: a proof-of-concept study. Hepatology 2016;64:189–99.
35. Mousa HS, Carbone M, Malinverno F, et al. Novel therapeutics for primary biliary cholangitis: toward a disease-stage-based approach. Autoimmun Rev 2016;15:870–6.
36. Bowlus CL. Cutting Edge Issues in Primary Sclerosing Cholangitis. Clin Rev Allergy Immunol 2011;139–50.
37. Seo N, Kim SY, Lee SS, et al. Sclerosing Cholangitis: Clinicopathologic Features, Imaging Spectrum, and Systemic Approach to Differential Diagnosis. Korean J Radiol 2016;17:25–38.
38. Hov JR, Lleo A, Selmi C, et al. Genetic associations in Italian primary sclerosing cholangitis: heterogeneity across Europe defines a critical role for HLA-C. J Hepatol 2010;52:712–17.
39. Levy C, Lindor KD. Primary sclerosing cholangitis: epidemiology, natural history, and prognosis. Semin Liver Dis 2006;26:22–30.
40. Bjornsson E, Boberg KM, Cullen S, et al. Patients with small duct primary sclerosing cholangitis have a favourable long term prognosis. Gut 2002;51:731–5.

MULTIPLE CHOICE QUESTIONS

1. Autoimmune liver diseases:
   A. In all cases affects women more than men
   B. Are categorized by the type of cells targeted
   C. Are invariably associated with IBD
   D. Are invariably ANA positive

2. Primary biliary cholangitis:
   A. Is predominant in men
   B. Is predominant in children
   C. Is predominant in women
   D. Affects equally both sexes

3. Primary biliary cholangitis is associated more frequently with:
   A. Rheumatoid arthritis
   B. Fybromialgia
   C. Infertility
   D. Sjogren syndrome

4. Autoimmune hepatitis:
   A. Is associated with serum ANCA
   B. Is associated with serum ANA
   C. Is associated with serum AMA
   D. Is associated with serum hypogammaglobulinemia

5. Primary sclerosing cholangitis:
   A. Is associated with IBD
   B. Is diagnosed most efficiently with liver histology
   C. Is associated with serum ANA
   D. Is associated with serum AMA

77

Immunotherapy of Cancer

Anu Sharma, Matthew Campbell, Cassian Yee, Sangeeta Goswami,
Padmanee Sharma

The ability of the immune system to recognize and eradicate cancer was first postulated in the 19th century; however, proof of principle remained elusive until later. It was noted that some patients with sarcoma underwent spontaneous regression of their tumors upon incidental development of a skin infection with *Streptococcus pyogenes,* which was conjectured to elicit an immune response against the infection as well as against cancer cells. After this observation, a mixture of heat-killed *S. pyogenes* and *Serratia marcescens* (Coley's toxin) was used to treat patients with sarcoma, which resulted in complete tumor regression in some patients.[1] These early studies with Coley's toxin prompted many clinical trials aimed at stimulating the immune response to eradicate cancer, including the use of Bacille Calmette-Guérin (BCG) as a treatment for superficial bladder cancer.[2] However, incomplete understanding of the mechanistic details of immune responses led to failure of many early clinical trials. With recent advances in our understanding of basic principles that guide immune responses, specifically T-cell responses, immunotherapy is now established as one of the pillars of cancer treatment.

One important principle that has reshaped our thinking of cancer immunology is related to immunosurveillance. Cancer immunosurveillance relies on immune cells, including T cells, which can recognize antigens, including mutated proteins that generate novel antigens. T cells recognize these antigens when they are bound to self-MHC (major histocompatibility complex) molecules.

In the early 2000s, the hypothesis of cancer immunosurveillance was redefined to state that the immune system controls both the tumor quality (immunogenicity) and the quantity (tumor load), thus eliminating tumors and sculpting the immunogenic phenotypes of tumors in immunocompetent hosts. The interaction between the host immune system and the tumor cells has been proposed in three phases, including elimination, equilibrium, and escape (the three E's). In the *elimination phase,* cancer cells are detected and destroyed by the adaptive and innate arms of the immune system before becoming clinically apparent. These data helped to frame the importance of the immune system in cancer development and revived the theory of cancer immunosurveillance.[3]

Cancer cells that evade elimination enter the ***equilibrium phase,*** where T cells keep the remaining cancer cells in check without eliminating them. This phase may extend throughout the host's lifetime.[3] Clinical examples of the equilibrium phase include the duration between successful treatment of the primary tumor and relapse, the existence of disease-free state despite micrometastasis, and the development of donor-derived tumors after organ transplant.[4]

KEY CONCEPTS

Three Phases of Immune Interaction With Tumor Cells

- Elimination: Tumor cells are detected and destroyed by the immune system.
- Equilibrium: Tumor cells are not eliminated by the immune system.
- Escape: Tumor cells escape immune control.

The continuous interaction of cancer cells with the immune system may favor the outgrowth of less immunogenic tumors that can escape immune control, termed the ***escape phase.*** Tumor cell escape can occur through different mechanisms including (i) loss of antigens and/or MHC expression; (ii) increased resistance to the cytotoxic effects of immunity through induction of antiapoptotic mechanisms; (iii) development of defects in antigen processing/presentation; (iv) loss of crucial genes, such as genes related to the interferon (IFN) signaling pathway, which makes tumor cells more resistant to immune-mediated killing by IFN; or (v) recruitment of immunosuppressive cells to the tumor microenvironment, which can inhibit antitumor immune responses and contribute to the escape of edited tumor variants.[5]

KEY CONCEPTS

Mechanisms of Tumor Cell Immune Escape

- Loss of antigen or major histocompatibility complex (MHC) expression
- Resistance to cytotoxicity
- Defects in tumor antigen processing/presentation
- Loss of crucial genes involved in the immune response
- Recruitment of immunosuppressive cells to the tumor microenvironment

Besides increasing our understanding of immunosurveillance, we have gained new insights regarding mechanisms that regulate T-cell responses, which have led to the development of an entirely new field known as immune checkpoint therapy. The current chapter will discuss (i) Activation and regulation of T-cell responses, (ii) Immune checkpoint therapy, (iii) Other inhibitory immune checkpoints, (iv) Immune checkpoint therapy with clinical benefits in solid tumors and hematological malignancies, (v) Early and late-phase trials with immune checkpoint therapy in other tumors, (vi) Immune costimulatory molecules,

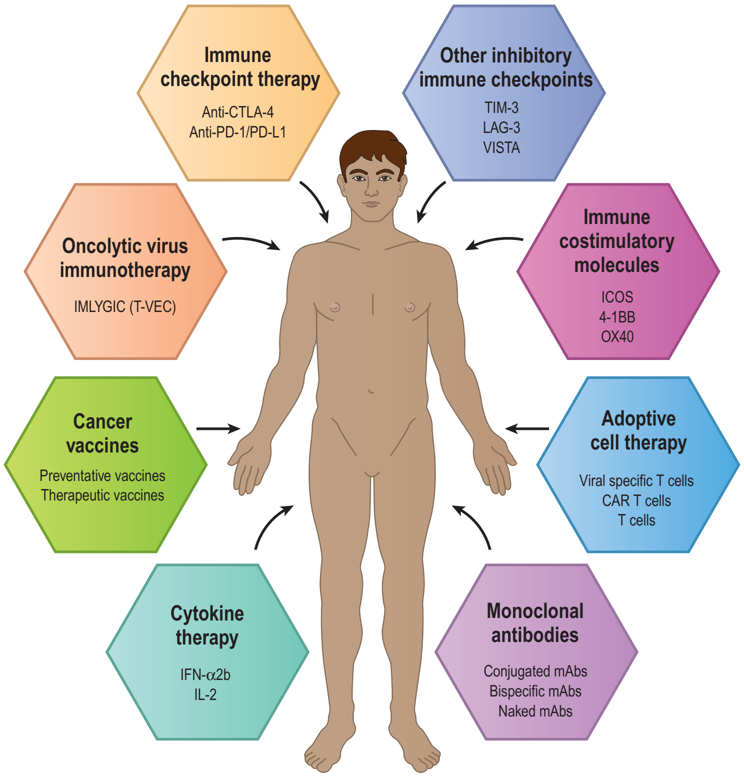

Immune checkpoint therapy
Anti-CTLA-4
Anti-PD-1/PD-L1
Other inhibitory immune checkpoints
TIM-3
LAG-3
VISTA
Oncolytic virus immunotherapy
IMLYGIC (T-VEC)
Immune costimulatory molecules
ICOS
4-1BB
OX40
Cancer vaccines
Preventative vaccines
Therapeutic vaccines
Adoptive cell therapy
Viral specific T cells
CAR T cells
T cells
Cytokine therapy
IFN-α2b
IL-2
Monoclonal antibodies
Conjugated mAbs
Bispecific mAbs
Naked mAbs

FIG 77.1 Anticancer immunotherapeutic strategies.

(vii) Adoptive cell transfer, (viii) Monoclonal antibodies, (ix) Cytokine therapy, (x) Cancer vaccines, (xi) Oncolytic virus immunotherapy, (xii) Clinical challenges in immunotherapy, and (xiii) Perspectives on future developments. Various anticancer immunotherapeutic strategies are illustrated in Fig. 77.1.

ACTIVATION AND REGULATION OF T-CELL RESPONSES

T-cell activation involves complex interactions involving both T-cell receptor (TCR) signaling and CD28 costimulation (Fig. 77.2) (Chapter 12). T-cell receptor interacts with foreign antigen in the context of self-MHC, which provides "signal 1." However, signal 1 by itself is insufficient to enable T-cell activation. "Signal 2" is provided when CD28 receptor, which is constitutively expressed on T cells, binds to B7-1 (CD80) and B7-2 (CD86) molecules that are expressed on antigen-presenting cells (APCs) but not on tumor cells. Therefore tumor cells alone are unable to start the process of T-cell activation. T cells require fragments of tumor cells to be phagocytosed by APCs such as dendritic cells, with eventual antigen processing and presentation by the APCs. T cells then interact with the APCs to receive both signals 1 and 2 for appropriate T-cell activation, which leads to cytokine production and proliferation as well as active killing of tumor cells.

Early research in the mid-1990s demonstrated that T-cell activation was a complex event, which enabled proliferation and functional differentiation but also induced inhibitory pathways that eventually could attenuate and terminate T-cell responses. The first intrinsic T-cell inhibitory molecule to be described was cytotoxic lymphocyte antigen-4 (CTLA-4), which is expressed on T cells only after T-cell activation. CTLA-4 has high homology to CD28 and binds to B7 (CD80 and CD86) molecules with much higher affinity than CD28. Two research groups independently showed that activation of T cells results in induction of CTLA-4, which accumulates at the T cell–APC interface and eventually outcompetes CD28 for binding to B7,

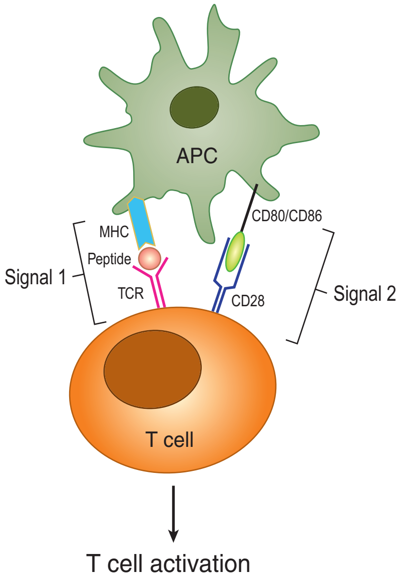

APC
MHC
Peptide
TCR
CD80/CD86
CD28
Signal 1
Signal 2
T cell
T cell activation

FIG 77.2 T cell activation. The "signal 1" for T cell activation, occurs following the recognition of MHC-peptide complex on an antigen-presenting cell (APC) by the T cell receptor (TCR) on a T cell. The "signal 2" for T cell activation is provided by binding of B7 molecules (CD80/86) on the APC to CD28 on the T cells. Following this interaction, T cells are activated and perform various effector functions.

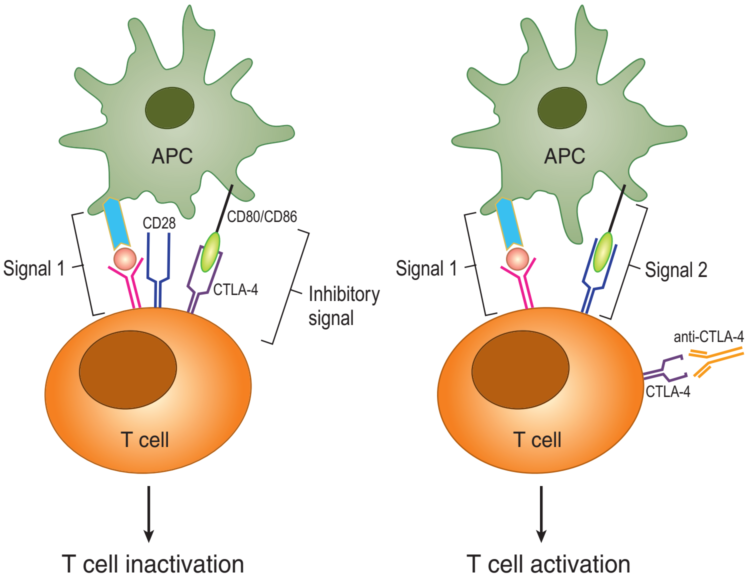


FIG 77.3 Regulation of T cell response. Upon T cell activation, T cells traffic to tumor leading to the cell surface expression of immune checkpoints such as CTLA-4, which bind with a higher affinity to the B7 molecules, thereby downregulating T cell responses. Targeting CTLA-4 with anti-CTLA-4 antibodies blocks this inhibitory interaction and CD28 is again free to interact with the B7 molecules, further amplifying T cell responses against the tumor.

thereby abrogating T-cell responses.[6,7] Based on these data, researchers proposed to enhance antitumor immune responses by blocking the CTLA-4 inhibitory pathway (Fig. 77.3). This approach, termed immune checkpoint therapy, has proven to be a paradigm shift in cancer therapy.

IMMUNE CHECKPOINT THERAPY

Anti-CTLA-4 Therapy

CTLA-4 is an immune checkpoint that is expressed on the surface of activated T cells to restrain T-cell responses. In other words, unrestricted T-cell responses can damage normal tissues and cells; therefore CTLA-4 plays an important role in controlling T-cell responses and in preventing damage to self. This mechanism protects normal cells, but in the setting of cancer, this mechanism also protects cancer cells. Preclinical studies elegantly demonstrated that blocking CTLA-4 enhances antitumor immune responses.[8]

Ipilimumab was the first fully human monoclonal (IgG1) antibody (mAb) against CTLA-4 to enter clinical trials in the late 1990s. Phase I and II trials demonstrated remarkable clinical responses and tumor regression in cancer patients, including patients with melanoma, renal cell carcinoma (RCC), prostate cancer, urothelial carcinoma, and ovarian cancer.[9] In both early and late-phase trials, ipilimumab has shown consistent activity against melanoma. The success of anti-CTLA-4 in eliciting durable clinical responses and in potentially curing some patients with cancer led to the establishment of a new field termed *immune checkpoint therapy,* which has expanded with the identification of additional T-cell inhibitory pathways.[10]

Anti-PD-1/PD-L1 Therapy

Programmed death 1 (PD-1) belongs to the CD28/CTLA-4 family and is expressed on the surface of activated T cells, B cells, monocytes, dendritic cells (DC) and natural killer (NK) cells. Its role in T-cell inhibition was first demonstrated in early 2000. Unlike CTLA-4, PD-1 appears to inhibit T-cell responses by interfering with TCR signaling, which is distinctly different from CTLA-4 and its ability to inhibit T-cell responses by competing with CD28 for binding to B7 molecules. PD-1 interacts with two ligands, PD-L1 and PD-L2, which appear to be equally capable of suppressing T-cell responses.[11]

Antibodies targeting the PD-1/PD-L1 interaction have shown beneficial clinical responses in multiple tumor types. Nivolumab was the first mAb targeting PD-1 to show significant clinical activity in patients with unresectable or metastatic melanoma, non–small-cell lung carcinoma (NSCLC), metastatic RCC, and classical Hodgkin lymphoma.[11]

Given the considerable activity of agents targeting immune checkpoints, numerous early studies have been conducted across several solid tumors. Presented in Table 77.1 is the published literature for agents that have been US Food and Drug Administration (FDA) approved or tested in the phase III setting in melanoma, RCC, and NSCLC.

KEY CONCEPTS

Primary Targets of Immune Checkpoint Therapy

- CTLA-4: interferes with T-cell costimulatory signal
- PD-1: interferes with T-cell receptor signaling

The blockade of CTLA-4 and PD-1/PD-L1 inhibitory interactions has been extensively explored clinically. Nevertheless, these interactions represent only a fraction of the potential targets that can serve to improve antitumor immune responses. Additional inhibitory pathways have been identified and are currently being targeted to improve antitumor immune responses. A few of these are discussed in the sections that follow.

T-Cell Immunoglobulin and Mucin Domain 3

T-cell immunoglobulin and mucin-domain containing-3 (TIM-3) is expressed by dendritic cells, monocytes, and T cells. TIM-3 is postulated to bind to its presumed ligand galectin-9, which is expressed on tumors. Studies in murine models indicate that TIM-3 is coexpressed with PD-1 on T cells. Dual blockade of TIM-3 and PD-1 has been shown to improve antitumor immune responses and reduce tumor growth in preclinical studies.[12] In a phase I/II trial, the safety and efficacy of MBG453, an antibody against TIM-3, is being evaluated as a single agent

TABLE 77.1 FDA Approval of Immune Checkpoint Inhibitors in the Treatment of Cancer

Cancer Type	Drug	Dose, Frequency, Duration	Disease State	Line of Treatment	Approval Year
Melanoma	Ipilimumab	3 mg/kg IV Q3wk × 4 doses until PD/tox	Unresectable/Metastatic	Any	
Melanoma	Pembrolizumab	2 mg/kg IV Q3wk until PD/tox	Unresectable/Metastatic	Refractory	2014
Melanoma	Nivolumab	3 mg/kg IV Q2wk until PD/tox	Unresectable/Metastatic	Post ipilimumab	2014
Melanoma	Pembrolizumab	2 mg/kg IV Q3wk until PD/tox	Unresectable/Metastatic	Initial	2015
Melanoma	Nivolumab	3 mg/kg IV Q2wk until PD/tox	Unresectable/Metastatic	Initial	2015
Melanoma	Ipilimumab plus Nivolumab	Ipi 3 mg/kg IV Q3wk ×4 + Nivo 1 mg/kg IV Q3wk ×4 followed by Nivo Q3wk ×4 until PD/tox	Unresectable/Metastatic BRAF V600 WT	Initial	2015
Melanoma	Ipilimumab	3 mg/kg IV Q3wk × 4 doses 12 until PD/tox	Surgically resected, high risk for recurrence	Adjuvant	2015
Kidney	Nivolumab	3 mg/kg IV Q3wk until PD/tox	Advanced	Post 1 antiangiogenic therapy	2015
NSCLC All PD-L1 +	Pembrolizumab	2 mg/kg IV Q3wk until PD/tox	Advanced	Post platinum	2015
NSCLC All	Nivolumab	3 mg/kg IV Q2wk until PD/tox	Advanced	Post platinum, Post EGFR/ALK for pt with mutation	2015
NSCLC	Atezolizumab	1200 mg/kg IV Q3wk until PD/tox	Advanced	Post platinum, Post EGFR/ALK for pt with mutation	2016
Urothelial	Atezolizumab	1200 mg IV Q3wk until PD/tox	Unresectable/Metastatic	Post platinum	2016
Hodgkin lymphoma	Nivolumab	3 mg/kg IV Q3wk until PD/tox		Post autologous SCT, posttransplant brentuximab vedotin	2016
Head and Neck Squamous	Pembrolizumab	200 mg IV Q3wk until PD/tox	Unresectable/Metastatic	On or post platinum	2016
Head and Neck Squamous	Nivolumab	3 mg/kg IV Q2wk until PD/tox	Unresectable/Metastatic	On or post platinum	2016

ALK, anaplastic lymphoma kinase; EGFR, epidermal growth factor receptor; FDA, US Food and Drug Administration; IV, intravenous; ipi, ipilimumab; nivo, nivolumab; NSCLC, non–small-cell lung carcinoma; PD, disease progression; pt, patients; SCT, stem cell transplant; tox, toxicity; wk, week; WT, wild type.

and in combination with anti-PD-1 in patients with advanced malignancies (ClinicalTrials.gov).

Lymphocyte Activation Gene 3

Lymphocyte activation gene-3 (LAG-3) is expressed on B cells, some T cells, and NK cells. LAG-3 is postulated to bind to MHC class II, which inhibits T-cell responses.[13] Immuntep (IMP321), a recombinant soluble fusion protein of LAG-3, is currently being tested in two or more clinical trials (ClinicalTrials.gov). A phase I study is also testing another anti-LAG-3 (BMS-986016) as monotherapy or in combination with anti-PD-1 (nivolumab) in patients with advanced solid tumors (ClinicalTrials.gov).

V-Domain Ig Suppressor of T-Cell Activation

V-domain Ig suppressor of T-cell activation (VISTA) is a negative immune checkpoint ligand, which shares homology with PD-L1 and is primarily expressed on hematopoietic cells. Preclinical studies with VISTA blockade have shown improved antitumor immune responses, leading to impeded tumor growth and improved survival.[14] Clinical trials with anti-VISTA mAbs are ongoing and include a phase I study with JNJ-61610588, a fully human IgG1 anti-VISTA mAb (ClinicalTrials.gov).

IMMUNE CHECKPOINT THERAPY WITH CLINICAL BENEFIT IN SOLID TUMORS AND HEMATOLOGICAL MALIGNANCIES

Melanoma

In recent years, immunotherapy has dramatically changed the landscape of melanoma treatment: Since 2011, four immunotherapies have been approved by the FDA.

A randomized phase III trial in patients with metastatic melanoma was conducted whereby patients received either anti-CTLA-4 antibody (ipilimumab), with or without a peptide (gp100) vaccine, or vaccine treatment alone. The observed median overall survival (OS) was 10 months for patients who received ipilimumab plus gp100 vaccine, 10.1 months for patients who received ipilimumab alone, and 6.4 months for patients who received only the gp100 vaccine. Importantly, approximately 20% of patients who received ipilimumab had long-term survival of greater than 3 years. These data led to the FDA approval of ipilimumab for metastatic melanoma in 2011.[15]

Another phase III trial in patients with previously untreated metastatic melanoma demonstrated that OS was significantly longer for patients who received ipilimumab plus dacarbazine chemotherapy compared with patients who received dacarbazine alone.[16] Additionally, ipilimumab versus placebo was investigated in patients with high-risk stage III melanoma after complete resection, where ipilimumab was given every 3 weeks for 4 doses, followed by every 3 months for up to 3 years. Median recurrence-free survival was 46.5% in patients treated with ipilimumab versus 34.8% for patients who received placebo, which led to the FDA approval of ipilimumab in the adjuvant setting.[17] Ipilimumab is approved in the metastatic as well as in the adjuvant setting for the treatment of melanoma. Importantly, to provide an accurate estimate of long-term survival following treatment with ipilimumab, a pooled analysis was performed from 10 prospective and two retrospective studies, including the two phase III trials mentioned above. The data demonstrated long-term survival of approximately 20% of patients, with survival of 10 years noted for some patients. The results were independent of ipilimumab dosing or prior therapy.[18]

Antibodies against another immune checkpoint inhibitor, PD-1 (nivolumab), have also been explored for the treatment of patients with metastatic melanoma. A phase III study in patients with advanced melanoma who progressed after ipilimumab therapy demonstrated an objective response rate (ORR) of 31.7% in patients treated with nivolumab compared with an ORR of 10.6% in patients treated with chemotherapy alone.[19]

Another randomized clinical trial in patients with previously untreated metastatic melanoma without BRAF mutation showed an improved OS in patients treated with nivolumab (72.9%) compared with patients treated with dacarbazine chemotherapy (42.1%) at 1 year. Subgroup analysis demonstrated that the nivolumab-treated patients had improved OS regardless of expression of the ligand for PD-1 (PD-L1) on tumor cells.[20] Based on these two phase III trials, nivolumab was FDA approved for the treatment of patients with refractory unresectable or metastatic melanoma and also for the first-line treatment of patients with metastatic melanoma.

Another anti-PD-1 mAs, pembrolizumab, was investigated in a phase II study that randomized patients with metastatic melanoma who were previously treated with ipilimumab and a BRAF inhibitor. Patients received either pembrolizumab 2 mg/kg every 3 weeks, pembrolizumab 10 mg/kg every 3 weeks, or chemotherapy. The 6-month progression-free survival (PFS) was significantly higher in patients treated with pembrolizumab compared with patients treated with chemotherapy (34% in 2 mg/kg cohort, 38% in 10 mg/kg cohort, and 16% in chemotherapy cohort).[21] The FDA approved pembrolizumab at a dose of 2 mg/kg intravenous every 3 weeks for the treatment of refractory melanoma. Pembrolizumab versus ipilimumab was then investigated in another phase III trial, where patients were randomized to two pembrolizumab regimens (10 mg/kg dose every 2 weeks or every 3 weeks) and an ipilimumab regimen (3 mg/kg dose every 3 weeks for 4 cycles), with PFS and OS as the primary endpoints. The 6-month PFS was 47.3% with pembrolizumab treatment every 2 weeks compared with 46.4% with pembrolizumab treatment every 3 weeks, and 26.5% with ipilimumab treatment. One-year estimates of survival were 74.1% for patients receiving pembrolizumab every 2 weeks, 68.4% for those receiving pembrolizumab every 3 weeks, and 58.2% for those receiving ipilimumab, leading to the FDA approval of pembrolizumab as a first-line treatment for patients with metastatic melanoma.[22]

Another phase III trial compared ipilimumab, nivolumab, and the combination in patients with untreated, unresectable, or metastatic melanoma with PFS and OS as coprimary endpoints. Treatment with combination therapy was statistically superior, leading to a median PFS of 11.5 months versus 6.9 months for nivolumab alone versus 2.8 months for ipilimumab alone. However, the data on the OS were not sufficiently mature to present. Based on these findings, the FDA approved combination therapy with ipilimumab plus nivolumab as first-line treatment for patients with previously untreated, unresectable, or metastatic melanoma, regardless of PD-L1 expression on tumor cells.[23]

KEY CONCEPTS

US Food and Drug Administration-Approved Checkpoint Inhibitors in Melanoma

- Ipilimumab: metastatic melanoma, used in the adjuvant setting
- Nivolumab: first-line treatment for metastatic melanoma; treatment of refractory unresectable or metastatic melanoma
- Pembrolizumab: first-line treatment for metastatic melanoma
- Ipilimumab plus nivolumab: first-line treatment for patients with previously untreated, unresectable, or metastatic melanoma

Lung Cancer

Similar to melanoma, NSCLC is also characterized by high mutational load,[24] and immune checkpoint therapy has led to durable antitumor response in NSCLC.

In a phase III study, patients with advanced squamous NSCLC who progressed during or after first-line chemotherapy were randomly assigned to receive nivolumab versus docetaxel chemotherapy. OS and PFS were significantly better with nivolumab compared with docetaxel, regardless of PD-L1 expression on tumor cells. This led to FDA approval of nivolumab for the treatment of patients with squamous NSCLC.[25] This approval was subsequently expanded to all subtypes of NSCLC, based on another phase III study that demonstrated a median OS of 12.2 months for patients with nonsquamous NSCLC who received nivolumab treatment compared with 9.4 months for patients who received docetaxel.[26]

Similarly, another phase III study evaluated pembrolizumab versus docetaxel for the treatment of patients with previously treated advanced NSCLC (both squamous and nonsquamous). Patients enrolled in this trial were selected to have at least 1% of tumor cells positive for the expression of PD-L1. The median OS observed with pembrolizumab was 10.4 months with the 2 mg/kg dose and 12.7 months with the 10 mg/kg dose versus 8.5 months with docetaxel. In patients with at least 50% of tumor cells expressing PD-L1, the OS was 14.9 months with pembrolizumab 2 mg/kg and 17.3 months with pembrolizumab 10 mg/kg versus 8.2 months after treatment with docetaxel. Based on these findings, the FDA approved pembrolizumab as single-agent, second-line therapy for patients with NSCLC whose tumors express PD-L1.[27]

Another phase III trial investigated pembrolizumab versus chemotherapy in previously untreated patients with advanced NSCLC, with PD-L1 expression on at least 50% of tumor cells and without epidermal growth factor receptor (EGFR) or anaplastic lymphoma kinase (ALK) mutations. The study demonstrated statistically significant improvements in OS for patients randomized to pembrolizumab compared with chemotherapy; the study also showed a significant improvement in median PFS of 10.3 months in the pembrolizumab group versus 6.0 months in the chemotherapy group.[28] This study led to the approval of pembrolizumab as first-line therapy for the treatment of patients with metastatic NSCLC whose tumors express PD-L1.

Several other PD-1/PD-L1 checkpoint inhibitors are in late-stage clinical testing for patients with lung cancer.

KEY CONCEPTS

FDA-Approved Checkpoint Inhibitors in Non–Small-Cell Lung Cancer

- Nivolumab: treatment of all subtypes of non–small-cell lung cancer (NSCLC)
- Pembrolizumab: single-agent, second-line therapy for NSCLC expressing PD-L1
- Pembrolizumab first-line therapy for metastatic NSCLC expressing PD-L1

Renal Cell Carcinoma

Antiangiogenic therapies have been the established treatment for patients with metastatic RCC. However, patients eventually develop innate or adaptive resistance to antiangiogenic therapy, which could be attributed partly to changes in the immune microenvironment,[29] proving a strong rationale to investigate the clinical effects of immune checkpoint therapy in RCC patients.

A phase III study in patients with advanced or metastatic RCC, who had received at least one prior treatment, demonstrated improved OS of 25 months after treatment with nivolumab compared with OS of 19.6 months in patients treated with everolimus.[30] Based on this trial, the FDA approved nivolumab as second-line therapy for patients with metastatic RCC. New immune-based treatment strategies for treating RCC patients are in development.

Bladder Cancer

Platinum-based chemotherapy has been the established first-line treatment for patients with metastatic bladder cancer; no new therapies have been established in the past 30 years, including no standardized second-line treatment because of the dismal objective response rates of 10% or less with tested chemotherapies. However, preclinical studies have demonstrated clinical benefits with immune checkpoint therapy in bladder tumors.[31] A single-arm phase II study investigated the role of an anti-PD-L1 antibody (atezolizumab), to block one of the ligands that interacts with PD-1, in patients with locally advanced or metastatic urothelial bladder cancer whose disease had progressed after previous treatment with platinum-based chemotherapy. Compared with the historical control ORR of 10%, treatment with atezolizumab resulted in a significantly improved ORR of 15% for all treated patients, regardless of PD-L1 expression.[32] This study led to FDA approval of atezolizumab in patients with locally advanced or metastatic urothelial bladder cancer whose cancers had progressed during or after treatment with platinum-based chemotherapy.

A phase I/II study investigated nivolumab in patients with locally advanced or metastatic urothelial cancer whose disease had progressed after previous treatment with platinum-based chemotherapy. This study demonstrated that nivolumab monotherapy led to an ORR of 24.4%.[33] Another, larger phase II study assessed the clinical activity and safety of nivolumab in 270 patients with metastatic bladder cancer who had progressive disease despite first-line platinum-based chemotherapy. This study reported an ORR of 19.6% with nivolumab therapy.[34]

Additionally, in another phase I/II study that treated patients with metastatic bladder cancer who progressed after platinum-based chemotherapy, patients who received combination treatment with ipilimumab 3 mg/kg plus nivolumab 1 mg/kg had an ORR of 38.5% versus an ORR of 26% for patients who received ipilimumab at 1 mg/kg plus nivolumab at 3 mg/kg; patients who received nivolumab alone had an ORR of 24.4%.[35] These early findings with higher-dose ipilimumab provide evidence for improved responses with combination immune checkpoint therapy.

Head and Neck Cancer

Head and neck cancer is one of the cancer types for which new immune-based cancer treatments are currently in development.

In 2016, pembrolizumab was FDA approved for the treatment of patients with recurrent or metastatic head and neck squamous-cell carcinoma (HNSCC) that continued to progress despite chemotherapy. The approval was based on data from a nonrandomized trial in patients with HNSCC who failed to respond to platinum-based chemotherapy.[36] Following treatment, pembrolizumab demonstrated an ORR of 16%. The median OS observed with pembrolizumab was 9.6 months, compared with the historical median OS of only 6 months. In this trial, responses were seen in patients with human papillomavirus (HPV)-positive tumors (24%) as well as in patients with HPV-negative tumors (16%), respectively.

Another phase III trial investigated nivolumab versus standard single-agent systemic therapy (methotrexate, docetaxel, or cetuximab) in patients with recurrent or metastatic HNSCC after platinum-based chemotherapy. This study demonstrated that treatment with nivolumab resulted in longer OS of 7.5 months compared with an OS of 5.1 months with standard single-agent therapy.[37] These findings led to the FDA approval of nivolumab in the treatment of platinum-refractory recurrent HNSCC.

Hodgkin Lymphoma

A phase I study of nivolumab in previously treated patients with relapsed or refractory Hodgkin lymphoma demonstrated an objective response of 87%, including 17% of patients with a complete response and 70% of patients with a partial response; the remaining 13% had stable disease. The Reed-Sternberg cells that define Hodgkin lymphoma have frequent amplification of chromosome 9p24.1, which leads to expression of PD-1 ligand in Reed-Sternberg cells and provides strong rationale for immune checkpoint therapy that blocks the PD-1/PD-L1 axis.[38] Findings from this study led to FDA approval of nivolumab for patients with previously treated relapsed or refractory Hodgkin lymphoma.

EARLY AND LATE-PHASE TRIALS WITH CHECKPOINT THERAPY IN OTHER TUMORS

Pancreatic Cancer

Pancreatic cancer has a poor prognosis with limited treatment options. A poor understanding of the immune microenvironment in pancreatic cancer has been a roadblock in developing effective treatment strategies. We have generated preliminary data to indicate that pancreatic tumors have immune infiltrates and differentially express inhibitory immune checkpoints compared with tumors that respond to checkpoint therapy, such as melanoma and bladder cancer.[39] Therefore treatment of pancreatic cancer with combination strategies that include immune checkpoint therapy is being explored.

Early-phase trials with ipilimumab have demonstrated delayed responses in patients with advanced pancreatic cancer.[40] Immune checkpoint therapies that are under development for treating pancreatic cancer include pembrolizumab, nivolumab with or without ipilimumab, and durvalumab with or without tremelimumab (ClinicalTrials.gov).

Prostate Cancer

Despite advances in the treatment of metastatic castration-resistant prostate cancer (mCRPC), novel immune-based strategies that can provide durable and long-term responses in patients are still needed. Sipuleucel-T (autologous cellular immunotherapy) is the only approved immunotherapy for patients with prostate cancer.[41] However, the exact mechanism of action of sipuleucel-T remains to be elucidated, and clinical metrics and/or biomarkers to evaluate responses with sipuleucel-T are currently lacking.

Several early-phase clinical trials have evaluated different doses, schedules, and combinations of immune checkpoint inhibitors in patients with mCRPC. A phase I/II trial (CA184–107) of ipilimumab as a single agent or in combination with radiotherapy in patients with previously treated mCRPC demonstrated a decline in prostate-specific antigen (PSA) in 15% of patients.[42] However, a phase III clinical trial (CA184–043) of bone-directed radiotherapy followed by either ipilimumab or placebo in patients with mCRPC failed to demonstrate an improvement in OS.[43]

Further studies are warranted to investigate the potential effects of combining immune checkpoint therapy with other treatment strategies, including other immunotherapies, chemotherapy, radiotherapy, and hormonal agents, in patients with prostate cancer in different treatment settings (metastatic hormone-naïve or metastatic castration-resistant disease).

IMMUNE COSTIMULATORY MOLECULES

Multiple ligand-receptor interactions play a pivotal role in shaping T-cell responses, including those that positively regulate T-cell function, thus making them potential targets for cancer immunotherapy. Many such costimulatory molecules are being investigated in preclinical or early-phase clinical studies. The two major families of costimulatory receptors expressed by T cells are the immunoglobulin-like (Ig) superfamily (including inducible T-cell costimulator [ICOS]) and the tumor necrosis factor receptor (TNFR) superfamily (including OX40 and 4–1BB).[44] These costimulatory molecules are suitable candidates for the development of targeted therapeutics. A schematic overview of the costimulatory and coinhibitory ligand-receptor interactions is illustrated in Fig. 77.4.

Inducible T-Cell Costimulator

ICOS is another member of the CD28/CTLA-4 immunoglobulin superfamily, whose expression is increased upon T-cell activation. ICOS can be expressed on activated effector T cells as well as on activated T regulatory cells (Tregs). ICOS interacts with its ligand, ICOSL, which is expressed on APCs, and provides a key costimulatory signal for T-cell proliferation and survival. We have shown that ICOS$^+$ effector T cells increase in patients after anti-CTLA-4 (ipilimumab) treatment.[45] An increased frequency of ICOS$^+$ CD4 T cells was also identified as a pharmacodynamic biomarker of anti-CTLA-4 therapy.[46] Another study with a different anti-CTLA-4 antibody (tremelimumab) also reported an increase in ICOS$^+$ CD4 T cells in blood samples obtained from patients with breast cancer.[47] Collectively, these findings indicated a biologically relevant role for ICOS in antitumor immune responses elicited by anti-CTLA-4 therapy.

Furthermore, preclinical studies demonstrated that combination therapy with anti-CTLA-4 plus an agonist to target the ICOS/ICOSL pathway led to a significant improvement in T-cell responses and tumor rejection.[48] These data led multiple pharmaceutical companies to pursue the development of an anti-ICOS antibody, including a recent trial with GSK3359609, an anti-ICOS antibody for treating patients with advanced solid tumors (NCT02723955; ClinicalTrials.gov).

ICOS previously was shown to be expressed on T-follicular helper cells (Tfh cells).[49] A clinical trial is ongoing with administration of an anti-ICOS antibody as treatment for patients with relapsed or refractory peripheral T-cell lymphoma follicular variant or angioimmunoblastic T-cell lymphoma (ClinicalTrials.gov).

4-1BB (CD137)

4-1BB (CD137) is an activation-induced costimulatory molecule that is expressed on activated T cells, NK cells, dendritic cells, eosinophils, mast cells, endothelial cells, and some tumor cells. Anti-4-1BB mAbs have demonstrated improved antitumor T-cell responses, with rejection of established syngeneic tumor cell lines in preclinical models. Several early-phase trials with anti-4-1BB antibodies have been initiated, and they have demonstrated antitumor responses, including partial remission and stabilization of disease in a subset of patients with advanced malignancies; however, immune-related adverse events (irAEs) have led to reevaluation of the dose and schedule of treatment with anti-4-1BB agonists.[50] Two separate phase I clinical studies

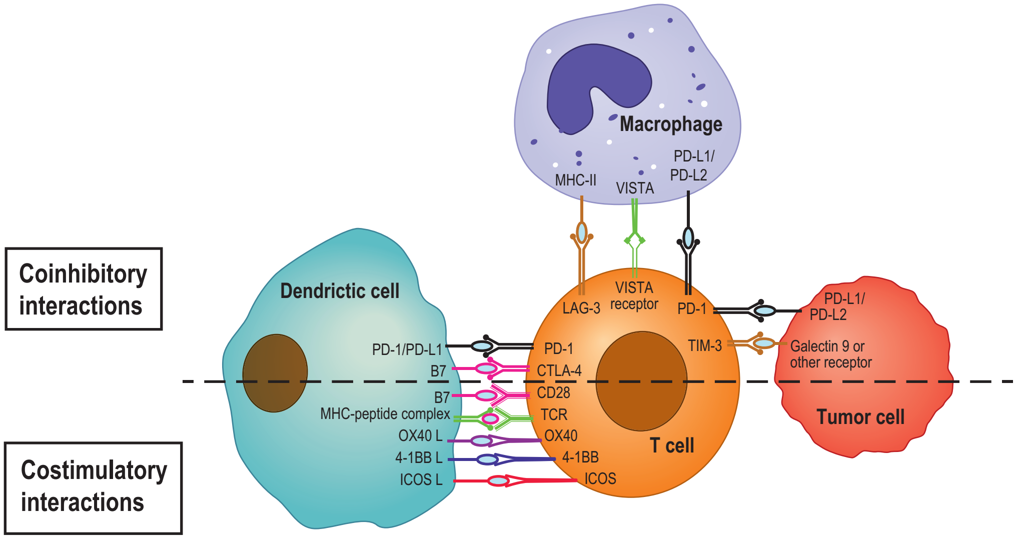


FIG 77.4 **Costimulatory and coinhibitory ligand–receptor interactions:** A schematic overview of the costimulatory and coinhibitory interactions between a T cell and dendritic cell; T cell and a macrophage and a T cell and a tumor cell in the tumor microenvironment.

to assess the safety and immunoregulatory activity of urelumab (BMS-663513) in patients with advanced and/or metastatic tumors and relapsed/refractory B-cell non-Hodgkin lymphoma are under investigation (ClinicalTrials.gov).

OX40 (CD134)

OX40 (CD134) is expressed on activated T cells, including activated effector and Tregs. Engagement of OX40 with its ligand enhances proliferation and survival of T cells. Preclinical studies with OX40 have shown increased antitumor responses and overall survival. A trial with the murine anti-OX40 agonist 9B12 mAb in patients with advanced cancers was well tolerated and demonstrated regression of at least one metastatic lesion in 12 out of 30 patients.[51] Humanized antibodies against OX40 are under clinical development. A phase I first-in-human study to evaluate the anti-OX40 antibody GSK3174998 as monotherapy and in combination with anti-PD-1 therapy (pembrolizumab) in patients with advanced solid tumors is ongoing (ClinicalTrials.gov). In addition, a phase Ib dose escalation study of the anti-OX40 antibody MOXR0916 plus the anti-PD-L1 antibody atezolizumab is ongoing in patients with locally advanced or metastatic solid tumors (ClinicalTrials.gov).

ADOPTIVE CELL TRANSFER

Several clinical studies have demonstrated that the reinfusion of autologous lymphocytes in cancer patients, isolated from their own tumor or peripheral blood, could inhibit tumor growth, and these initial studies formed the basis for adoptive T-cell therapy.[52]

For adoptive immunotherapy, the isolation of antigen-specific T cells from the peripheral blood or tumor is followed by clonal expansion and transfusion back into the tumor-bearing host. Adoptive T-cell transfer has achieved tumor regression and improved survival by increasing the number of reactive T cells, providing long-term immune protection and guaranteeing antigen specificity. The downside to this method is both the expense and the expertise needed to perform this labor-intensive therapy. Assuming that the T cells are already primed to the tumor and are antigen-specific, they can be polyclonally activated *ex vivo* before being reinfused back into the patient. Suppressing the immune system before adoptive transfer is important for improving the antitumor efficacy. Significant progress has been made to establish adoptive T-cell therapy as a standard-of-care method for the treatment of cancer, and multiple studies are ongoing.[53]

Adoptive Transfer of Tumor-Specific Cytotoxic T Cells

Studies in murine tumor models have demonstrated that the lymphocytes in transplantable tumors are capable of recognizing tumor cells *in vitro,* and the adoptive transfer of these syngeneic tumor-infiltrating lymphocytes (TIL) could mediate tumor regression in certain cancers. *In vitro* studies demonstrated that TIL obtained from melanoma patients could specifically recognize autologous tumors. Further studies based on this observation demonstrated that transfer of autologous TIL could mediate objective regression in patients with metastatic melanoma and across other cancers.[53] An alternative source of endogenous tumor-reactive T cells can be found in the peripheral blood of patients. Strategies to isolate and expand these rare circulating T cells are being developed for further investigation in clinical trials to target a broader population of patients.[54]

Adoptive Immunotherapy With Genetically Modified Lymphocytes

In early 2000, studies demonstrated that genetically engineered lymphocytes transduced with a retrovirus encoding a TCR to recognize melanoma-melanocyte antigen (MART-1) could mediate tumor regression.[53]

To improve the T-cell antigen specificity, a new class of "chimeric antigen receptor" (CAR) was developed. This model was capable of overcoming limitations associated with central and peripheral tolerance, generating efficient T cells. The advantage of CAR compared with conventional TCR is that the antigen need not be processed and presented by MHC. Importantly, this approach can be used in all patients who express the same tumor antigen, regardless of HLA type.[53,54]

Clinical Development of Chimeric Antigen Receptor T Cells

Chimeric antigen receptor T cells (CAR T cells) are defined as T cells that are genetically modified to express an artificial construct, consisting of a synthetic TCR targeted at a predetermined antigen expressed on tumor cells.

Coupling the CAR construct, which typically consists of a single-chain variable fragment (scFv; directed against a known tumor antigen), with the aid of a linker and a spacer to the transmembrane domain and a CD3ζ signaling domain, which provides "signal 1" necessary for T-cell activation, paved its way as first-generation CAR T cells. Preclinical studies using first-generation CAR T cells directed against CD19 expressing B-cell malignancies and HER2/Neu showed promising results.[55]

After this success, the second-generation CAR T cells were developed by coupling the intracellular portion of the construct with additional costimulatory signaling domains such as CD28 or 4–1BB, which provided "signal 2" for T-cell activation and further improved the efficacy of CAR T cells.

KEY CONCEPTS

Chimeric Antigen Receptor (CAR) T Cells Synthetic Receptor

- Single-chain variable fragment
- Linker and a spacer to the transmembrane domain
- CD3ζ signaling domain
- Additional costimulatory signaling domains

Further optimization led to the development of the third-generation CAR T cells, which incorporated two-tandem costimulatory domains by coupling ICOS, CD27, CD28, 4–1BB, or OX40 (*e.g.,* CD28/4–1BB/CD3ζ or CD28/OX40/CD3ζ), hence demonstrating varying degrees of *in vitro* and *in vivo* T-cell activation, proliferation, and cytokine (interleukin [IL]-2) production.[55]

The armored CAR T represent the next-generation CAR T cells, which incorporate not only the two-costimulatory domains but also an additional transgene for cytokines (*e.g.,* IL-12) or ligands (*e.g.,* CD40L or 4–1BBL); this helps these CAR T cells to survive and/or disrupt an immunosuppressive tumor microenvironment.[53,55] The key concepts of the four generations of CAR T cells are illustrated in Fig. 77.5.

CTL019 is the first investigational CAR T to receive breakthrough-therapy status by the FDA for the treatment of adult and pediatric relapsed/refractory acute lymphoblastic

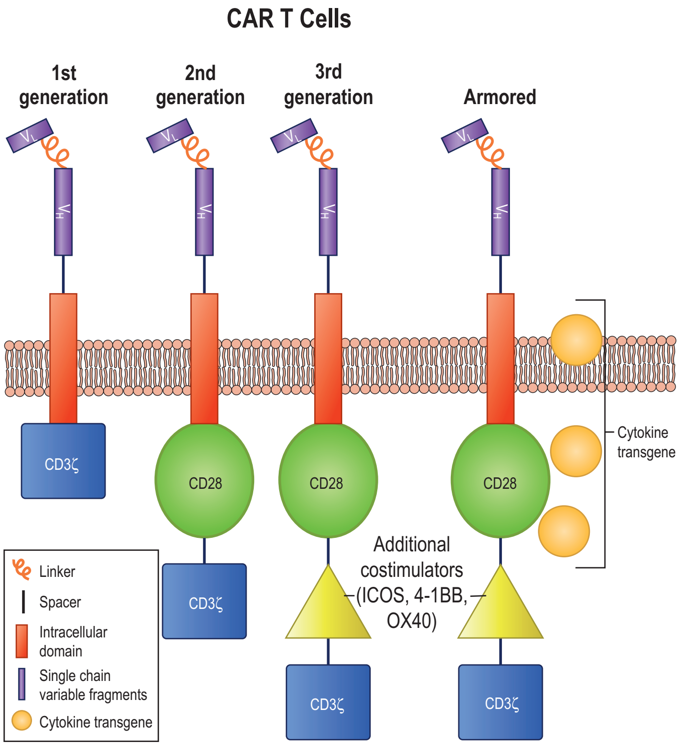


FIG 77.5 Different generations of CAR T cell constructs. The 1st generation CAR T comprising of a T cell signaling domain; the 2nd generation, with one costimulation domain and a T cell signaling domain; the 3rd generation, with two costimulation domains and a T cell signaling domain; and the armored CAR T cell with similar construct to third generation but with an additional pro-inflammatory cytokine transgene.

leukemia (NCT02228096; ClinicalTrials.gov). CTL019 is a second-generation CAR T cell that comprises an anti-CD19 extracellular scFv and CD3-ζ and 4–1BB intracellular signaling domains.

Adoptive transfer of CAR T cells has demonstrated significant promise in targeting hard-to-treat hematologic malignancies and solid tumors. Although the safety concerns related to cytokine release syndrome still need to be fully addressed, multiple CAR T cells are in clinical development.[56]

Adoptive Transfer of Viral-Specific T Cells

The adoptive transfer of T cells isolated or engineered to have specificity and reactivity for viral DNA can efficiently treat Epstein–Barr virus (EBV)-associated posttransplant lymphoproliferative disease.[57] It is known that CD8 and CD4 virus-specific T cells in chronic infections are characterized by the loss of T-cell function due to the upregulation of inhibitory molecules such as PD-1 and TIM-3; therefore the adoptive transfer of autologous T cells with high-affinity TCRs for viral proteins in combination with inhibitors of PD-1 or TIM-3 could be an ideal approach for the targeted therapy of virally induced malignancies.[58]

MONOCLONAL ANTIBODIES

The establishment of *hybridoma technology* in 1975 enabled the efficient development of mAbs. Immortalized B lymphocytes from immunized hosts were fused with nonimmunoglobulin murine myeloma cells, forming hybridomas, which produced large amounts of antibodies with high reactivity.[59] Later the development of *phage/yeasts* or *ribosome display technology* helped generate mAbs from any species and facilitated selection based on specificity, stability, and affinity.

The mounting success of a mAb as a therapeutic agent is based on three characteristics, including (i) an *Fc moiety* that mediates *antibody-dependent cellular cytotoxicity* (ADCC) and *complement-dependent cytotoxicity* (CDC); (ii) an *Fab moiety* that promotes high specificity and affinity for antigen binding; and (iii) a molecular mass of about 150 kDa, which extends the circulatory half-life of the mAb up to 21 days. Importantly, the mechanism of tumor cell killing by mAbs involves (i) binding to specific receptors on tumor cells via the Fab portion of the antibody and triggering cytotoxicity via the Fc portion of the antibody or (ii) binding to specific receptors on tumor cells via the Fab portion of the antibody and blocking important signaling pathways via the Fab portion of the antibody or a combination of both mechanisms.

Over the past 15 years, antibody-based immunotherapy has been established as a successful strategy for treating patients with hematological malignancies and some solid tumors. To enhance the effector functions, mAbs have been conjugated to radioisotopes, chemotherapeutic agents, bacterial toxins, cytokines, and enzymes[60]; therefore mAbs are classified as follows:

Naked mAbs

These mAbs are not attached to any drug or radioactive material and bind to specific molecules or antigens on tumor cells, which can block important signaling pathways, such as EGFR, or enable the immune system to destroy the tumor cells by ADCC or a combination of both.[60] ADCC is an important mechanism for tumor cell killing that is mediated by the interaction between the Fc region of an antibody and FcγRIIIa receptors on the surface of immune cells. mAbs may bind to a cell surface target via its Fab region and engage the effector cell expressing FcγRIIIa with the Fc region of the mAb, leading to subsequent killing of the tumor cell. ADCC is thought to be a key mechanism of activity of the FDA-approved mAbs rituximab and trastuzumab.[60]

Rituximab, a chimeric mAb targeted against the B-cell surface antigen CD20, was the first mAb to be FDA approved for therapeutic use against B-cell non-Hodgkin lymphoma (NHL), CD20-positive NHL, and chronic lymphocytic leukemia (CLL).

Trastuzumab (Herceptin) is another such mAb directed against human EGFR-2, a tyrosine kinase membrane receptor, which is overexpressed on approximately 30% of breast cancer cells in a subset of breast cancer patients. Trastuzumab monotherapy has demonstrated prolonged disease stability in a significant number of advanced-metastatic breast cancer patients.[61]

Other successful mAbs approved by the FDA for patients with solid tumors include bevacizumab targeting vascular endothelial growth factor (VEGF) (as first-line and second-line treatment of metastatic colon cancer in combination with 5-FU chemotherapy; in combination with carboplatin and paclitaxel chemotherapy as first-line treatment of advanced NSCLC patients who have not yet been treated with chemotherapy; as a single agent in patients with glioblastoma whose tumors have progressed after treatment; and for treating mRCC patients in combination with IFN-α); cetuximab targeting EGFR (as a single agent in HNSCC patients with failure to platinum-based therapy; in combination with radiotherapy for regionally advanced HNSCC; and as palliative treatment of pretreated metastatic EGFR-positive colorectal cancer); panitumumab targeting EGFR (as a single agent for treating pretreated EGFR-positive metastatic colorectal cancer); alemtuzumab targeting CD52 (as a single agent for treating CLL);

and ofatumumab targeting CD20 (for the treatment of CLL patients refractory to fludarabine and alemtuzumab).

Conjugated/Tagged/Labeled/Loaded mAbs

These mAbs target a molecule or antigen expressed on tumors and are coupled to a cytotoxic or radioactive agent to enable delivery of a toxic substance directly to the tumor site. Treatment with a radiolabeled mAb is termed *radioimmunotherapy* (RIT). Ibritumomab tiuxetan is a radiolabeled mAb against the CD20 antigen expressed on B lymphocytes and is FDA approved to treat different lymphomas. A randomized phase III trial comparing ibritumomab tiuxetan to rituximab in patients with relapsed or refractory NHL demonstrated a higher overall response rate (80% versus 56%) and complete response (CR) rate (30% versus 16%) for ibritumomab tiuxetan, but similar time to disease progression for both treatment groups.[62]

A different approach that uses mAbs coupled to chemotherapeutic agents is known as *antibody-drug conjugates* (ADC). Brentuximab vedotin (Zevalin) is a mAb conjugated to a chemotoxic drug called monomethyl auristatin E (MMAE), targeting the cell membrane antigen CD30 on lymphocytes. A phase II clinical study (NCT00848926; ClinicalTrials.gov) with brentuximab vedotin in patients with refractory Hodgkin lymphoma demonstrated complete remission in 34% of patients, partial remission in 40% of the patients, and tumor regression in 94% of patients.[63] Brentuximab vedotin is FDA approved to treat Hodgkin and anaplastic large-cell lymphoma. Trastuzumab emtansine (T-DM1) is another example of an antibody-drug conjugate, which is FDA approved for the treatment of HER2-positive breast cancer.[61]

Bispecific Monoclonal Antibodies

These mAbs are formed by coupling two different mAbs together and allowing the construct to bind to two separate proteins simultaneously, thus enabling the mAb to direct the immune system to act against the tumor.

Blinatumomab is one such example where one part of the mAb binds to the CD19 protein expressed on B-lineage acute lymphoblastic leukemia cells, while the second part of the mAb binds to the CD3 protein found on T cells, thereby linking these two cell types to activate the T cell to exert cytotoxicity on the $CD19^+$ tumor cells. Blinatumomab was first approved by the FDA in 2014, based on the observations of a phase II trial in patients with relapsed or refractory acute lymphoblastic leukemia in which 40% of patients achieved a CR or complete response with partial hematologic recovery (CRh).[64]

CYTOKINE THERAPY

Cytokines are low-molecular-weight proteins or glycoproteins that mediate cell-to-cell communication and play a pivotal role in regulating biological activities, including innate immunity, adaptive immunity, inflammation, and hematopoiesis.

IL-2 and IFN-α have been used previously to treat patients with metastatic melanoma and RCC. However, due to the nonspecific activation of the immune response and the toxicities associated with these cytokines, as well as the development of novel immunotherapy agents as discussed above, the use of high-dose IL-2 and IFN-α has diminished in clinical practice. Additional cytokines, including granulocyte macrophage–colony-stimulating factor (GM-CSF), IL-7, IL-12, IL-15, IL-18, and IL-21, have also been investigated in clinical studies for patients with advanced cancers.[65] Currently, the antitumor activity of PEGylated human IL-10 (AM0010) is being investigated in patients with advanced solid tumors.[66]

However, certain properties of cytokines pose significant clinical challenges in achieving antitumor responses. Some cytokines are pleiotropic and act on several immune cells and mediate opposite effects. High-dose IL-2 is known to promote cytotoxic activity of effectors (CD8 T cells and NK cells) and differentiation of CD4 T cells into T helpers. However, IL-2 can also preferentially expand Tregs due to a higher affinity to the IL-2 receptor (CD25) on these cells. Sometimes multiple cytokines have the same functional effect, leading to redundancy of cytokine signaling; therefore the therapeutic manipulation of cytokines can be challenging, as the modification of one cytokine can be compensated by another cytokine.[65]

CANCER VACCINES

Efforts to produce effective anticancer vaccines have been ongoing for decades. The two known categories are preventive cancer vaccines and therapeutic cancer vaccines.

Preventive vaccines consist of those that target infectious organisms that can cause cancer such as strains of HPV, which contribute to the development of some head/neck cancers, anal cancers, and cervical cancers. The three FDA-approved vaccines include Cervarix HPV (against HPV-16,-18, -31,-33,-45), Gardasil (against HPV-6,-11,-16,-18), and Gardasil 9 (against HPV-6,-11,-16,-18,-31,-33,-45,-52,-58).[67]

For the treatment of established cancers, sipuleucel-T was the first vaccine to be FDA approved. It is a cell-based vaccine consisting of autologous peripheral blood mononuclear cells (PBMC), which are collected from each patient and activated *ex vivo* with a recombinant fusion protein PA2024 (consisting of a prostate antigen and prostatic acid phosphatase fused to GM-CSF) before reinfusion into the patient. A phase III clinical trial demonstrated significant improvement in median OS after treatment with sipuleucel-T in asymptomatic men with castrate-resistant prostate cancer. Treatment with sipuleucel-T also resulted in an improvement in the rate of 3-year survival in 31.7% of patients compared with 23.0% for those receiving placebo.[67] The exact mechanism of action for sipuleucel-T remains to be defined.

ONCOLYTIC VIRUS IMMUNOTHERAPY

Oncolytic viruses represent a new class of therapeutic agents that promote antitumor immune responses, which depend on mechanisms that elicit selective tumor cell killing and induction of systemic antitumor immunity. A variety of native and genetically modified viruses are being developed and clinically investigated as oncolytic agents.[68]

IMLYGIC (talimogene laherparepvec, T-VEC) was the first oncolytic viral therapy to receive FDA approval for the treatment of unresectable cutaneous and subcutaneous and nodal lesions recurrent after surgery in melanoma patients. IMLYGIC demonstrated improved median OS in a phase III study in patients with advanced melanoma.[69] IMLYGIC is a genetically modified herpes simplex virus type designed to selectively replicate within the tumor cells and induce lysis.

CLINICAL CHALLENGES IN IMMUNOTHERAPY

Multiple challenges that remain for immunotherapy agents include resistance mechanisms and immune-related adverse events.

Resistance Mechanisms

Durable clinical responses with immune checkpoint therapy have been well documented; some of these responses last for years. However, many patients do not respond to initial therapy or relapse with disease after initially responding to treatment. Resistance mechanisms may involve the following: (i) activation of oncogenic pathways (WNT/B-catenin signaling, loss of PTEN expression), (ii) loss of interferon-gamma (IFNγ) signaling genes, (iii) recruitment of immunosuppressive cells such as Treg, myeloid-derived suppressor cells, M2 macrophages, and expression of other inhibitory immune checkpoints, and (iv) composition of gut microbiome.

KEY CONCEPTS

Mechanisms of Tumor Resistance to Immune-Mediated Therapy

- Activation of oncogenic pathways
- Loss of interferon-γ signaling genes in the tumor cells
- Recruitment of immunosuppressive cells (*e.g.,* Treg cells, myeloid-derived suppressor cells)
- Expression/secretion of inhibitory molecules
- Composition of the gut microbiome

Activation of Oncogenic Pathways

Recent research has identified an intricate interplay between oncogenic pathways and antitumor immunity.

In one such study, the whole exome sequencing and gene expression profiling of melanoma biopsies revealed that tumor-cell intrinsic activation of the WNT/β-catenin pathway correlates with absence of T cells in the tumor microenvironment.[70] Further investigation in a genetically engineered autochthonous mouse melanoma model confirmed that increased oncogenic β-catenin signaling results in a failure to recruit $CD8\alpha^+$ and $CD103^+$ DC populations in the tumor due to repressed expression of the chemokine CCL4. Consequently, mice with tumors expressing active β-catenin responded poorly to anti-CTLA-4 and anti-PD-L1 therapy, compared with mice with tumors lacking β-catenin expression.[70]

Apart from the activation of active β-catenin, the loss of PTEN expression is a frequent event in melanoma, especially in tumors with *BRAF* mutations. PTEN is a negative regulator of the PI3K-AKT pathway, and the complete loss of PTEN is associated with increased signaling through this pathway, correlating with shorter OS in patients with advanced melanoma.[71] A recent study in a melanoma mouse model demonstrated that the loss of PTEN expression can reduce the therapeutic activity of immune checkpoint therapy. The study additionally demonstrated that a selective PI3K-β-isoform inhibitor could induce synergy with immune checkpoint therapy in the preclinical model.[72] Similarly, findings in human melanoma specimens confirmed that the loss of PTEN correlates with exclusion of CD8 T cell in the tumor and increased PI3K activation. Importantly, patients with tumors lacking PTEN expression demonstrated poor clinical responses to anti-PD-1 therapy compared with patients with retained PTEN expression.[72]

Loss of Interferon γ Signaling

It is well established that IFNγ, a critical cytokine produced by T cells and other immune cells, plays a role in promoting innate and adaptive immune responses and in inhibiting tumor cell proliferation. IFNγ functions by binding to the IFNγ receptor, which initiates signaling through the Janus kinase/signal transducer and activator of transcription (JAK/STAT) pathway to regulate gene expression.[73] Data from two independent studies have established that the loss of genes involved in the IFNγ signaling pathway may lead to both primary[74] and adaptive resistance[75] to immune checkpoint therapy.

Previously published data from a study in patients with bladder cancer demonstrated increased frequency of a $CD4^+$ $ICOS^{hi}$ T-cell subset that produced IFNγ following anti-CTLA-4 therapy.[45] We further hypothesized that tumor cells harboring mutations in the IFNγ-signaling pathway are not susceptible to the IFNγ produced by these T cells, which may result in resistance to anti-CTLA-4 therapy.[74] Evaluation of genomic alterations of IFNγ-pathway genes in melanoma patients treated with ipilimumab demonstrated that the tumors of the nonresponders had loss of IFNγ-pathway genes at significantly higher frequencies compared with the responders.

Furthermore, mice bearing melanoma tumors with knockdown of IFNγ-receptor, resulted in impaired tumor rejection when treated with anti-CTLA-4 therapy.

Similarly, another study compared paired tumor biopsies that were collected from four melanoma patients before treatment with anti-PD-1 (pembrolizumab) and at the time of disease relapse.[75] In two out of the four patients whose disease relapsed after initial responses to pembrolizumab therapy, mutations were detected in the genes encoding interferon-receptor–associated Janus kinase 1 (JAK1) or Janus kinase 2 (JAK2), two critical components of the IFNγ signaling pathway.

Taken together, these findings indicate that the loss of IFNγ signaling limits the effectiveness of the immune system to eliminate tumor cells and induces resistance to immune checkpoint therapy.

Immunosuppressive Tumor Microenvironment

In addition to the above-mentioned genomic alterations in tumor cells, other resistance mechanisms within the tumor microenvironment include Treg cells, myeloid-derived suppressor cells (MDSCs), M2 macrophages, expression of other inhibitory immune checkpoints (TIM-3, LAG-3, VISTA), recruitment of inhibitory molecules/ligands (indoleamine dioxygenase [IDO], carcinoembryonic antigen-related cell adhesion molecule 1 [CEACAM1], PD-L1), and the secretion of inhibitory cytokines (IL-10, IL-35, transforming growth factor [TGF]-β); these may have a direct negative effect on effector T-cell function and hence contribute toward immunosuppression and resistance to immunotherapy.[76,77]

Gut Microbiome

Two independent preclinical studies reported that the composition of the gut microbiome can contribute to the difference in responses to immunotherapy in cancer patients.[78,79]

One study compared the tumor growth in mice with distinct gut microbiota and identified that mice that harbored the commensal *Bifidobacterium* species had enhanced spontaneous antitumor immunity compared with mice with a different gut microbiota composition. Importantly, the direct administration of the *Bifidobacterium* species into mice with established melanoma tumors improved tumor-specific immunity and response to anti-PD-L1 therapy.[78]

Similarly, another study in tumor-bearing mice and cancer patients demonstrated that the efficacy of anti-CTLA4 therapy is associated with T-cell responses specific for the microbiota species *Bacteroides fragilis.* The study reported that antibiotic-treated and

TABLE 77.2 Frequency of Immune-Related Adverse Events in Phase III Clinical Trials With Immune Checkpoint Therapies Across Diseases

Drug/ Disease/No. of pts	Any Toxicity All Grades %/ Grade ≥3 %	Diarrhea	Colitis	Pneumonitis	Hepatitis	Rash	Endocrine
Ipilimumab		33/5	8/5	NA	4/0	44/2	8/4
Melanoma 108							
{Hodi FS, 2010 #10}							
Ipilimumab Adjuvant							
Melanoma Resected/	99/55	49/10	16.3/8	NA	25/8	34/13	38/8
471{Eggermont AM, 201 #59}							
Ipilimumab							
MCRPC 393{Kwon ED, 2014 #56}	98/59	51/17	7/5	NA	12/4	17/1	5/2
Tremelimumab	96/52	51/18	NA included with diarrhea	NA	1/1	33/2	8/3
Melanoma 325{Ribas A, 2013 #58}							
Ipilimumab/ 311	99/56	33/6	12/9	2/0	7/2	35/0	11/2
Nivolumab/ 313	99/44	19/2	1/1	1/0	6/3	22/0	14/1
Combination/ 313	100/69	44/9	12/8			28/3	30/5
Melanoma							
{Larkin J, 2015 #1}							
	Related to Tx						
Pembro Q2wk / 278	80/13	17/3	2/1	0/0	1/1	15/0	17/1
Pembro Q3wk / 277	73/10	15/1	4/3	1/0	2/2	13/0	13/1
Ipilimumab / 256	73/20	23/3	8/7	0/0	1/0	15/1	7/2
Melanoma							
{Robert C, 2015 #9}							
Nivolumab							
NSCLC Squamous 272	58/7	8/0	1/1	5/1	2/0	4/0	2/ 0
{Brahmer J, 2015 #43}	69/10	8/1	1/1	4/1	1/1	13/1	8/0
NSCLC Non-Squamous 292							
{Borghaei H, 2015 #42}							
Pembro							
2 mg/kg / 339	63/15	7/1	1/1	5/2	1/<1	9/<1	14/1
10 mg/kg 343	66/16	6/1	1/<1	4/2	1/0	13/<1	16/1
NSCLC All{Herbst RS, 2015 #45}							
Nivolumab							
MRCC / 406{Motzer RJ, 2015 #31}	79/19	12/1	NA	4/1	NA	10/<1	NA

MCRPC, metastatic castrate-resistant prostate cancer; NSCLC, non–small-cell lung cancer; MRCC, metastatic clear-cell renal cell carcinoma; pembro, Pembrolizumab; NA, not available; pts, patients.

germ-free tumor bearing mice did not respond to anti-CTLA-4 therapy compared with mice harboring *Bacteroides fragilis*. Additionally, the study reported that ipilimumab-treated patients with metastatic melanoma tend to have *Bacteroides fragilis* as a significant component of their microbiome.[79]

Collectively, these studies demonstrate the importance of gut microbiota and its role in facilitating the anticancer effects of immune checkpoint therapy.

Immune-Related Adverse Events

The clinical use of immunotherapeutic strategies is expanding rapidly, and the side effect profile of immunotherapeutic agents is distinct from that of conventional anticancer agents.[80] Therefore it is important to recognize and manage the irAEs in the growing patient population treated with immunotherapy. Early recognition and treatment of symptoms, time to discontinuation of therapy in case of an adverse event, and adherence to established treatment algorithms are essential to prevent patient morbidity and mortality.

irAEs With Immune Checkpoint Immunotherapy

Immune checkpoint blockade with antibodies targeting CTLA-4 and PD-1/PD-L1 has been extensively evaluated in the clinical setting. In a subset of patients, nevertheless, treatment with these immune checkpoint antibodies can be associated with distinctive inflammatory tissue damage,[81] which has been termed *immune-related adverse events* (irAEs). Table 77.2 summarizes the adverse events associated with anti-CTLA-4 and anti-PD-1, respectively, in phase III clinical trials. The spectrum of irAEs related to anti-CTLA4 and anti-PD-1/PD-L1 immunotherapy includes common reactions involving the gastrointestinal tract (diarrhea, colitis), skin (rash, dermatitis), liver (transaminitis), or endocrine axis (hypophysitis, thyroiditis).

The combination of nivolumab (anti-PD-1) and ipilimumab (anti-CTLA-4) has a much higher incidence of irAEs compared with monotherapy with either antibody. For instance, a phase III clinical trial in patients with advanced melanoma receiving nivolumab or ipilimumab, or a combination of both antibodies, demonstrated grade 3/4 treatment-related adverse events in 16% of patients treated with nivolumab, in 27% of patients treated with ipilimumab, and in 55% of patients treated with the combination.[82]

The optimal management of irAEs is based on clinical experience, and many of these irAEs are transient and responsive to steroid therapy.[83] Various immunomodulatory or immunosuppressive agents such as azathioprine and mycophenolate mofetil (MMF) have also been effective in the management of steroid-refractory irAEs.[84]

Recent findings from a retrospective study involving two clinical trials with ipilimumab in patients with prostate cancer identified CD8 T-cell clonal expansion within the systemic circulation as a potential correlative biomarker of irAEs.[85] These findings highlight a potential biomarker to predict, and provide early intervention for, toxicities that occur in patients who receive ipilimumab therapy.

irAEs With Adoptive T-Cell Therapy

The adoptive transfer of T cells is generally well tolerated, but it may be associated with the occurrence of life-threatening irAEs. Serious irAEs after adoptive T-cell therapy may occur when T cells targeting differentiation antigens in the tumor also recognize these antigens on normal cells. For example, patients with melanoma who receive adoptive T-cell therapy with T cells that target differentiation antigens such as MART-1 and gp100 may develop vitiligo and uveitis as irAEs, since these antigens are also expressed on normal cells.[86,87] Additionally, patients with metastatic RCC who were treated with carbonic anhydrase-IX (CAIX)-specific CAR T cells developed liver toxicity.[88] In addition, fatal irAEs were reported in cancer patients who received anti-MAGE-A3 TCR-transduced T cells.[89]

Adoptive T-cell therapy with CAR T cells may also be associated with another potentially life-threatening toxicity termed as *cytokine release syndrome* (CRS). The hallmark of CRS is elevated circulating levels of cytokines—including IL-6, TNF-α, and interferon-γ—resulting in fever, rigors, hypotension, and hypoxia.[90,91]

Treatment of irAEs associated with adoptive T-cell therapy, such as uveitis and colitis, can often be managed with topical or systemic corticosteroids; as for CAR T-cell therapy–related CRS, it is generally managed with either high-dose corticosteroids or agents such as tocilizumab (an IL-6 receptor-blocking antibody).[92]

PERSPECTIVES ON FUTURE DEVELOPMENTS

Recent success with immunotherapy in the management of cancer has given credence to the long-held belief that the immune system can be used to treat cancer. However, the current immunotherapeutic strategies form only the tip of the iceberg in terms of potential targets and combinations that can serve to improve clinical responses. A large number of additional agonistic and antagonistic mAbs that target immune checkpoints, as well as costimulatory molecule to potentiate T cell–mediated immune responses, are in various stages of clinical development; these include antibodies against ICOS, OX40, 4–1BB, LAG-3, and VISTA (ClinicalTrials.gov). Furthermore, intense clinical investigation is currently ongoing for the development of effective anticancer vaccines and adoptive cell transfer strategies (ClinicalTrials.gov).

ON THE HORIZON BOX

- Develop optimal patient selection strategy for cancer immunotherapy
- Develop the most appropriate timing for cancer immunotherapy
- Maximize cancer therapy effectiveness through strategies for:
  - Developing combined immunotherapeutic approaches
  - Optimally integrating standard anticancer modalities with immunotherapy

Biomarkers

To identify potential biomarkers that are relevant to antitumor immune responses, clinical trials are being designed to obtain tumor tissues to gain a better understanding of the tumor microenvironment and of immunological changes that occur over time.[10] Several biomarkers have been associated with clinical outcomes to immune checkpoint therapy, including high levels of PD-L1 expression, infiltration of effector T cells, and mutational landscape.[70,93,94] However, it is unlikely that a single biomarker will predict clinical outcome to immunotherapy; therefore, it will be necessary to integrate multiple biomarkers in immune monitoring studies. Additionally, immune monitoring studies should include gene expression studies, epigenetics studies, flow cytometry and CYTOF studies, immunohistochemical studies, mutational load studies, neoantigen studies, and microbiome studies. Sufficient access to longitudinal patient samples will be required to conduct these types of studies.

Combination Therapy

Immune checkpoint therapy has resulted in durable clinical responses, but in only a fraction of patients. On the basis of the encouraging results with monotherapy, combination treatment strategies are being explored extensively.[95]

Because anti-CTLA-4 and anti-PD-1 antibodies have distinct mechanisms of action, preclinical studies evaluated combination treatment with anti-CTLA-4 and anti-PD-1/PD-L1, which demonstrated improved antitumor immune responses.[31,96]

A detailed study in preclinical tumor models demonstrated that treatment with a combination of anti-CTLA-4 and anti-PD-1 antibodies improves antitumor immunity and synergistically eradicates tumors, and this synergy relies on the interdependence between IL-7 and IFNγ signaling in T cells; the lack of either pathway abrogates the therapeutic effects of this combination therapy.[31]

In 2015 the FDA approved the first combination immunotherapy of nivolumab (anti-PD-1) plus ipilimumab (anti-CTLA-4) for patients with previously untreated and unresectable metastatic melanoma.[82,97] Patients receiving this combination had an objective response rate of 50%. Importantly, this combination was effective irrespective of the PD-L1 expression. Several other combination therapies are currently being tested in clinical trials (ClinicalTrials.gov).

Combination treatments with immune checkpoint therapy plus targeted therapy (angiogenesis inhibitors, tyrosine kinase inhibitors, BRAF inhibitors) also hold promise for improved efficacy. A phase I trial (NCT01472081; ClinicalTrials.gov) investigating the combination of nivolumab with sunitinib (a receptor tyrosine-kinase inhibitor) demonstrated tumor regression in patients with metastatic RCC.[98] The safety and efficacy of another tyrosine-kinase inhibitor (lenvatinib) in combination with pembrolizumab (anti-PD-1) are currently being explored in a phase I/II study (ClinicalTrials.gov).

Another interesting combination currently being evaluated is radiotherapy plus immunotherapy. Several studies have demonstrated that radiotherapy supports tumor-specific immunity by enhancing antigen presentation and increasing the vulnerability of tumor cells to T cell–mediated attack.[99,100] Therefore combination therapy with radiation plus immune checkpoint therapy or other immunotherapeutic strategies is warranted.[101] A phase I trial of hypofractionated radiotherapy in combination with MEDI4736 (an anti-PD-L1 antibody) and tremelimumab (anti-CTLA-4) is being tested in patients with metastatic melanoma and lung, breast, and pancreatic cancer (ClinicalTrials.gov).

CONCLUSIONS

Cancer immunotherapy has established itself as a critical pillar in cancer therapy. It has resulted in durable clinical responses in patients whose tumors were refractory to conventional anticancer therapies. Nonetheless, additional research is required to further optimize the existing immunotherapeutic approaches and also to assess new strategies in this emerging field with significant untapped potential. Current challenges include rational development and optimization of combination strategies, identification of biomarkers, and continued support for basic and translational research to provide a greater understanding of mechanisms that drive antitumor immune response. Cancer immunotherapy has made great strides in improving clinical outcomes for some patients, and there is great hope that continued research efforts will enable even more patients to benefit in the future.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Nauts HC. Bacterial pyrogens: beneficial effects on cancer patients. Prog Clin Biol Res 1982;107:687–96.
2. Bassi P. BCG (Bacillus of Calmette Guerin) therapy of high-risk superficial bladder cancer. Surg Oncol 2002;11(1-2):77–83.
3. Schreiber RD, Old LJ, Smyth MJ. Cancer immunoediting: integrating immunity's roles in cancer suppression and promotion. Science 2011;331(6024):1565–70.
4. Quail DF, Joyce JA. Microenvironmental regulation of tumor progression and metastasis. Nat Med 2013;19(11):1423–37.
5. Dunn GP, Old LJ, Schreiber RD. The three Es of cancer immunoediting. Annu Rev Immunol 2004;22:329–60.
6. Walunas TL, Lenschow DJ, Bakker CY, et al. CTLA-4 can function as a negative regulator of T cell activation. Immunity 1994;1(5):405–13.
7. Krummel MF, Allison JP. CD28 and CTLA-4 have opposing effects on the response of T cells to stimulation. J Exp Med 1995;182(2):459–65.
8. Leach DR, Krummel MF, Allison JP. Enhancement of antitumor immunity by CTLA-4 blockade. Science 1996;271(5256):1734–6.
9. Sharma P, Allison JP. The future of immune checkpoint therapy. Science 2015;348(6230):56–61.
10. Sharma P, Wagner K, Wolchok JD, et al. Novel cancer immunotherapy agents with survival benefit: recent successes and next steps. Nat Rev Cancer 2011;11(11):805–12.
11. Hamanishi J, Mandai M, Matsumura N, et al. PD-1/PD-L1 blockade in cancer treatment: perspectives and issues. Int J Clin Oncol 2016;21(3):462–73.
12. Sakuishi K, Apetoh L, Sullivan JM, et al. Targeting Tim-3 and PD-1 pathways to reverse T cell exhaustion and restore anti-tumor immunity. J Exp Med 2010;207(10):2187–94.
13. Huang CT, Workman CJ, Flies D, et al. Role of LAG-3 in regulatory T cells. Immunity 2004;21(4):503–13.
14. Wang L, Rubinstein R, Lines JL, et al. VISTA, a novel mouse Ig superfamily ligand that negatively regulates T cell responses. J Exp Med 2011;208(3):577–92.
15. Hodi FS, O'Day SJ, McDermott DF, et al. Improved survival with ipilimumab in patients with metastatic melanoma. NEJM 2010;363(8):711–23.
16. Robert C, Thomas L, Bondarenko I, et al. Ipilimumab plus dacarbazine for previously untreated metastatic melanoma. N Engl J Med 2011;364(26):2517–26.
17. Eggermont AM, Chiarion-Sileni V, Grob JJ, et al. Adjuvant ipilimumab versus placebo after complete resection of high-risk stage III melanoma (EORTC 18071): a randomised, double-blind, phase 3 trial. Lancet Oncol 2015;16(5):522–30.
18. Schadendorf D, Hodi FS, Robert C, et al. Pooled Analysis of Long-Term Survival Data From Phase II and Phase III Trials of Ipilimumab in Unresectable or Metastatic Melanoma. J Clin Oncol 2015;33(17):1889–94.
19. Weber JS, D'Angelo SP, Minor D, et al. Nivolumab versus chemotherapy in patients with advanced melanoma who progressed after anti-CTLA-4 treatment (CheckMate 037): a randomised, controlled, open-label, phase 3 trial. Lancet Oncol 2015;16(4):375–84.
20. Robert C, Long GV, Brady B, et al. Nivolumab in previously untreated melanoma without BRAF mutation. NEJM 2015;372(4):320–30.
21. Ribas A, Puzanov I, Dummer R, et al. Pembrolizumab versus investigator-choice chemotherapy for ipilimumab-refractory melanoma (KEYNOTE-002): a randomised, controlled, phase 2 trial. Lancet Oncol 2015;16(8):908–18.
22. Robert C, Schachter J, Long GV, et al. Pembrolizumab versus Ipilimumab in Advanced Melanoma. N Engl J Med 2015;372(26):2521–32.
23. Larkin J, Chiarion-Sileni V, Gonzalez R, et al. Combined Nivolumab and Ipilimumab or Monotherapy in Untreated Melanoma. N Engl J Med 2015;373(1):23–34.
24. Alexandrov LB, Nik-Zainal S, Wedge DC, et al. Signatures of mutational processes in human cancer. Nature 2013;500(7463):415–21.
25. Brahmer J, Reckamp KL, Baas P, et al. Nivolumab versus docetaxel in advanced squamous-cell non–small-cell lung cancer. NEJM 2015;373(2):123–35.
26. Borghaei H, Paz-Ares L, Horn L, et al. Nivolumab versus Docetaxel in Advanced Nonsquamous Non–Small-Cell Lung Cancer. NEJM 2015;373(17):1627–39.
27. Herbst RS, Baas P, Kim DW, et al. Pembrolizumab versus docetaxel for previously treated, PD-L1-positive, advanced non-small-cell lung cancer (KEYNOTE-010): a randomised controlled trial. Lancet 2016;387(10027):1540–50.
28. Reck M, Rodriguez-Abreu D, Robinson AG, et al. Pembrolizumab versus chemotherapy for PD-L1-Positive non-small-cell lung cancer. N Engl J Med 2016;375(19):1823–33.
29. Liu XD, Hoang A, Zhou L, et al. Resistance to antiangiogenic therapy is associated with an immunosuppressive tumor microenvironment in metastatic renal cell carcinoma. Cancer Immunol Res 2015;3(9):1017–29.
30. Motzer RJ, Escudier B, McDermott DF, et al. Nivolumab versus everolimus in advanced renal-cell carcinoma. NEJM 2015;373(19):1803–13.
31. Shi LZ, Fu T, Guan B, et al. Interdependent IL-7 and IFN-gamma signalling in T-cell controls tumour eradication by combined alpha-CTLA-4+alpha-PD-1 therapy. Nat Commun 2016;7:12335.
32. Rosenberg JE, Hoffman-Censits J, Powles T, et al. Atezolizumab in patients with locally advanced and metastatic urothelial carcinoma who have progressed following treatment with platinum-based chemotherapy: a single-arm, multicentre, phase 2 trial. Lancet 2016;387(10031):1909–20.
33. Sharma P, Callahan MK, Bono P, et al. Nivolumab monotherapy in recurrent metastatic urothelial carcinoma (CheckMate 032): a multicentre, open-label, two-stage, multi-arm, phase 1/2 trial. Lancet Oncol 2016;17(11):1590–8.
34. Sharma P, Retz M, Siefker-Radtke A, et al. Nivolumab in metastatic urothelial carcinoma after platinum therapy (CheckMate 275): a multicentre, single-arm, phase 2 trial. Lancet Oncol 2017;18(3):312–22.
35. Sharma P, Callahan MK, Calvo E, et al. Efficacy and safety of nivolumab plus ipilimumab in previously treated metastatic urothelial carcinoma:

First results from the phase I/II CheckMate 032 study. Presented at: 2016 SITC Annual Meeting; November 9-13, 2016; National Harbor, MD.

36. Muro K, Chung HC, Shankaran V, et al. Pembrolizumab for patients with PD-L1-positive advanced gastric cancer (KEYNOTE-012): a multicentre, open-label, phase 1b trial. Lancet Oncol 2016;17(6): 717–26.
37. Ferris RL, Blumenschein G Jr, Fayette J, et al. Nivolumab for recurrent squamous-cell carcinoma of the head and neck. N Engl J Med 2016;375(19):1856–67.
38. Ansell SM, Lesokhin AM, Borrello I, et al. PD-1 blockade with nivolumab in relapsed or refractory Hodgkin's lymphoma. N Engl J Med 2015;372:311–19.
39. Sharma A, Blando J, Allison JP, et al. Evaluation of immune infiltrates and inhibitory immune checkpoints in human pancreatic tumors compared to other tumors. Manuscript in review.
40. Royal RE, Levy C, Turner K, et al. Phase 2 trial of single agent Ipilimumab (anti-CTLA-4) for locally advanced or metastatic pancreatic adenocarcinoma. J Immunother 2010;33(8):828–33.
41. Kawalec P, Paszulewicz A, Holko P, et al. Sipuleucel-T immunotherapy for castration-resistant prostate cancer. A systematic review and meta-analysis. Arch Med Sci 2012;8(5):767–75.
42. Slovin SF, Higano CS, Hamid O, et al. Ipilimumab alone or in combination with radiotherapy in metastatic castration-resistant prostate cancer: results from an open-label, multicenter phase I/II study. Ann Oncol 2013;24(7):1813–21.
43. Kwon ED, Drake CG, Scher HI, et al. Ipilimumab versus placebo after radiotherapy in patients with metastatic castration-resistant prostate cancer that had progressed after docetaxel chemotherapy (CA184-043): a multicentre, randomised, double-blind, phase 3 trial. Lancet Oncol 2014;15(7):700v12.
44. Watts TH. TNF/TNFR family members in costimulation of T cell responses. Ann Rev Immunol 2005;23:23–68.
45. Liakou CI, Kamat A, Tang DN, et al. CTLA-4 blockade increases IFNgamma-producing CD4+ICOShi cells to shift the ratio of effector to regulatory T cells in cancer patients. Proc Natl Acad Sci USA 2008;105(39):14987–92.
46. Ng Tang D, Shen Y, Sun J, et al. Increased frequency of ICOS+ CD4 T cells as a pharmacodynamic biomarker for anti-CTLA-4 therapy. Cancer Immunol Res 2013;1(4):229–34.
47. Vonderheide RH, LoRusso PM, Khalil M, et al. Tremelimumab in combination with exemestane in patients with advanced breast cancer and treatment-associated modulation of inducible costimulator expression on patient T cells. Clin Cancer Res 2010;16(13): 3485–94.
48. Fan X, Quezada SA, Sepulveda MA, et al. Engagement of the ICOS pathway markedly enhances efficacy of CTLA-4 blockade in cancer immunotherapy. J Exp Med 2014;211(4):715–25.
49. Weber JP, Fuhrmann F, Feist RK, et al. ICOS maintains the T follicular helper cell phenotype by down-regulating Kruppel-like factor 2. J Exp Med 2015;212(2):217–33.
50. Bartkowiak T, Curran MA. 4-1BB Agonists: multi-potent potentiators of tumor immunity. Front Oncol 2015;5:117.
51. Curti BD, Kovacsovics-Bankowski M, Morris N, et al. OX40 is a potent immune-stimulating target in late-stage cancer patients. Cancer Res 2013;73(24):7189–98.
52. Yee C. The use of endogenous T cells for adoptive transfer. Immunol Rev 2014;257(1):250–63.
53. Rosenberg SA, Restifo NP. Adoptive cell transfer as personalized immunotherapy for human cancer. Science 2015;348(6230):62–8.
54. Hinrichs CS, Rosenberg SA. Exploiting the curative potential of adoptive T-cell therapy for cancer. Immunol Rev 2014;257(1):56–71.
55. Dai H, Wang Y, Lu X, et al. Chimeric antigen receptors modified T-Cells for cancer therapy. J Natl Cancer Inst 2016;108(7).
56. Teachey DT, Lacey SF, Shaw PA, et al. Identification of Predictive Biomarkers for Cytokine Release Syndrome after Chimeric Antigen Receptor T-cell Therapy for Acute Lymphoblastic Leukemia. Cancer Discov 2016;6(6):664–79.
57. Rooney C, Leen A. Moving successful virus-specific T-cell therapy for hematopoietic stem cell recipients to late phase clinical trials. Mol Ther Nucleic Acids 2012;1(11):e55.
58. Rooney CM, Leen AM, Vera JF, et al. T lymphocytes targeting native receptors. Immunol Rev 2014;257(1):39–55.
59. Kohler G, Milstein C. Continuous cultures of fused cells secreting antibody of predefined specificity. Nature 1975;256(5517):495–7.
60. Glassman PM, Balthasar JP. Mechanistic considerations for the use of monoclonal antibodies for cancer therapy. Cancer Biol Med 2014;11(1):20–33.
61. Ecker DM, Jones SD, Levine HL. The therapeutic monoclonal antibody market. MAbs 2015;7(1):9–14.
62. Witzig TE, Gordon LI, Cabanillas F, et al. Randomized controlled trial of yttrium-90-labeled ibritumomab tiuxetan radioimmunotherapy versus rituximab immunotherapy for patients with relapsed or refractory low-grade, follicular, or transformed B-cell non-Hodgkin's lymphoma. J Clin Oncol 2002;20(10):2453–63.
63. Report SGaM. Positive Data from Pivotal Trial of Brentuximab Vedotin (SGN-35) in Relapsed or Refractory Hodgkin Lymphoma at 2010 Annual Meeting of the American Society of Hematology (ASH) *(Corporate Press Release)*.
64. Topp MS, Gökbuget N, Stein AS, et al. Safety and activity of blinatumomab for adult patients with relapsed or refractory B-precursor acute lymphoblastic leukaemia: a multicentre, single-arm, phase 2 study. Lancet Oncol 2015;16(1):57–66.
65. Dranoff G. Cytokines in cancer pathogenesis and cancer therapy. Nat Rev Cancer 2004;4(1):11–22.
66. Todd Michael Bauer KPP, Karen A. Autio, Patrick Alexander Ott, Melinda Whiteside, Martin Oft, Jeffrey R. Infante, F. Stephen Hodi, Jedd D. Wolchok. A first-in-human study of pegylated recombinant human IL-10 (AM0010), daily administered for four months in selected advanced solid tumors. J Clin Oncol 2014;32(Suppl.; abstract TPS3126).
67. Yaddanapudi K, Mitchell RA, Eaton JW. Cancer vaccines: Looking to the future. Oncoimmunology 2013;2(3):e23403.
68. Kaufman HL, Kohlhapp FJ, Zloza A. Oncolytic viruses: a new class of immunotherapy drugs. Nat Rev Drug Discov 2015;14(9):642–62.
69. Andtbacka RH, Kaufman HL, Collichio F, et al. Talimogene Laherparepvec Improves Durable Response Rate in Patients With Advanced Melanoma. J Clin Oncol 2015;33(25):2780–8.
70. Spranger S, Bao R, Gajewski TF. Melanoma-intrinsic beta-catenin signalling prevents anti-tumour immunity. Nature 2015;523(7559):231–5.
71. Ascierto PA, Agarwala S, Botti G, et al. Future perspectives in melanoma research: meeting report from the "Melanoma Bridge." Napoli, December 1st-4th 2015. J Transl Med 2016;14(1):313.
72. Peng W, Chen JQ, Liu C, et al. Loss of PTEN promotes resistance to T Cell-mediated immunotherapy. Cancer Discov 2016;6(2):202–16.
73. Murray PJ. The JAK-STAT signaling pathway: input and output integration. J Immunol 2007;178(5):2623–9.
74. Gao J, Shi LZ, Zhao H, et al. Loss of IFN-gamma pathway genes in tumor cells as a mechanism of resistance to anti-CTLA-4 therapy. Cell 2016;167(2):397–404.e9.
75. Zaretsky JM, Garcia-Diaz A, Shin DS, et al. Mutations associated with acquired resistance to PD-1 blockade in melanoma. N Engl J Med 2016;375(9):819–29.
76. Gajewski TF, Schreiber H, Fu YX. Innate and adaptive immune cells in the tumor microenvironment. Nat Immunol 2013;14(10):1014–22.
77. Sharma P, Hu-Lieskovan S, Wargo JA, et al. Primary, adaptive, and acquired resistance to cancer Immunotherapy. Cell 2017;168(4): 707–23.
78. Sivan A, Corrales L, Hubert N, et al. Commensal Bifidobacterium promotes antitumor immunity and facilitates anti-PD-L1 efficacy. Science 2015;350(6264):1084–9.
79. Vetizou M, Pitt JM, Daillere R, et al. Anticancer immunotherapy by CTLA-4 blockade relies on the gut microbiota. Science 2015;350(6264):1079–84.

80. Friedman CF, Proverbs-Singh TA, Postow MA. Treatment of the Immune-Related Adverse Effects of Immune Checkpoint Inhibitors: A Review. JAMA Oncol 2016;2(10):1346–53.
81. Gao J, He Q, Subudhi S, et al. Review of immune-related adverse events in prostate cancer patients treated with ipilimumab: MD Anderson experience. Oncogene 2015;34(43):5411–17.
82. Larkin J, Hodi FS, Wolchok JD. Combined nivolumab and ipilimumab or monotherapy in untreated melanoma. N Engl J Med 2015;373(13):1270–1.
83. Weber JS, Dummer R, de Pril V, et al. Patterns of onset and resolution of immune-related adverse events of special interest with ipilimumab: detailed safety analysis from a phase 3 trial in patients with advanced melanoma. Cancer 2013;119(9):1675–82.
84. Michot JM, Bigenwald C, Champiat S, et al. Immune-related adverse events with immune checkpoint blockade: a comprehensive review. Eur J Cancer 2016;54:139–48.
85. Subudhi SK, Aparicio A, Gao J, et al. Clonal expansion of CD8 T cells in the systemic circulation precedes development of ipilimumab-induced toxicities. Proc Natl Acad Sci USA 2016;113(42):11919–24.
86. Yee C, Thompson JA, Roche P, et al. Melanocyte destruction after antigen-specific immunotherapy of melanoma: direct evidence of t cell-mediated vitiligo. J Exp Med 2000;192(11):1637–44.
87. Chandran SS, Paria BC, Srivastava AK, et al. Persistence of CTL clones targeting melanocyte differentiation antigens was insufficient to mediate significant melanoma regression in humans. Clin Cancer Res 2015;21(3):534–43.
88. Lamers CH, Sleijfer S, Vulto AG, et al. Treatment of metastatic renal cell carcinoma with autologous T-lymphocytes genetically retargeted against carbonic anhydrase IX: first clinical experience. J Clin Oncol 2006;24(13):e20–2.
89. Morgan RA, Chinnasamy N, Abate-Daga D, et al. Cancer regression and neurological toxicity following anti-MAGE-A3 TCR gene therapy. J Immunother 2013;36(2):133–51.
90. Lee DW, Gardner R, Porter DL, et al. Current concepts in the diagnosis and management of cytokine release syndrome. Blood 2014;124(2):188–95.
91. Tey SK. Adoptive T-cell therapy: adverse events and safety switches. Clin Transl Immunology 2014;3(6):e17.
92. Davila ML, Riviere I, Wang X, et al. Efficacy and toxicity management of 19-28z CAR T cell therapy in B cell acute lymphoblastic leukemia. Sci Transl Med 2014;6(224):224ra25.
93. Hamid O, Schmidt H, Nissan A, et al. A prospective phase II trial exploring the association between tumor microenvironment biomarkers and clinical activity of ipilimumab in advanced melanoma. J Transl Med 2011;9:204.
94. Patel SP, Kurzrock R. PD-L1 Expression as a Predictive Biomarker in Cancer Immunotherapy. Mol Cancer Ther 2015;14(4):847–56.
95. Callahan MK, Postow MA, Wolchok JD. CTLA-4 and PD-1 pathway blockade: combinations in the clinic. Front Oncol 2014;4:385.
96. Curran MA, Montalvo W, Yagita H, et al. PD-1 and CTLA-4 combination blockade expands infiltrating T cells and reduces regulatory T and myeloid cells within B16 melanoma tumors. Proc Natl Acad Sci USA 2010;107(9):4275–80.
97. Postow MA, Chesney J, Pavlick AC, et al. Nivolumab and ipilimumab versus ipilimumab in untreated melanoma. N Engl J Med 2015;372(21):2006–17.
98. Amin A, Plimack ER, Infante JR, et al. Nivolumab (anti-PD-1; BMS-936558, ONO-4538) in combination with sunitinib or pazopanib in patients (pts) with metastatic renal cell carcinoma (mRCC). J Clin Oncol 2014;32(Suppl.):5s, abstr 5010.
99. Sharma A, Bode B, Wenger RH, et al. gamma-Radiation promotes immunological recognition of cancer cells through increased expression of cancer-testis antigens in vitro and in vivo. PLoS ONE 2011;6(11):e28217.
100. Sharma A, Bode B, Studer G, et al. Radiotherapy of human sarcoma promotes an intratumoral immune effector signature. Clin Cancer Res 2013;19(17):4843–53.
101. Sharabi AB, Lim M, DeWeese TL, et al. Radiation and checkpoint blockade immunotherapy: radiosensitisation and potential mechanisms of synergy. Lancet Oncol 2015;16(13):e498–509.

MULTIPLE-CHOICE QUESTIONS

1. The immunotherapy regimen associated with the "maximum response" or the most patients with significant reduction in tumor burden based on phase III clinical trials completed to date is ———?
   A. Nivolumab
   B. Ipilimumab
   C. Nivolumab plus ipilimumab
   D. Pembrolizumab

2. A 64-year-old patient was diagnosed with a large 11.5-cm mass in the left kidney. The patient underwent nephrectomy, which revealed a pT3aN0cM0 clear-cell renal cell carcinoma. After 2 years of observation, he was found to have numerous pulmonary metastases. He was treated with sunitinib (Sutent) for 15 months and was later found to have disease progression. His physician is interested in treating him with nivolumab. Which pathway does nivolumab target?
   A. Programmed death 1 (PD-1)/PD-L1
   B. PD-L1/CD80
   C. Inducible T-cell costimulator (ICOS)/ICOS-L
   D. CD68/cytotoxic T lymphocyte antigen-4 (CTLA-4)

3. A 28-year-old female is diagnosed with cutaneous melanoma on her right shoulder. After surgical resection and imaging, she is diagnosed with stage III melanoma. She is interested in receiving adjuvant immune checkpoint therapy to reduce the risk of relapse at 3 years. Which immune checkpoint inhibitor has received US Food and Drug Administration (FDA) approval for this indication?
   A. Nivolumab
   B. Ipilimumab
   C. Atezolizumab
   D. Pembrolizumab

78

Lymphoid Leukemias

Moshe E. Gatt, Shai Izraeli

Leukemias are a group of hematological malignant clonal diseases arising in bone marrow. This chapter focuses on the two most common lymphoid leukemias: acute lymphoblastic leukemia (ALL) and chronic lymphocytic leukemia (CLL). The immunological aspects of both diseases are extensively discussed, with a special emphasis on the clinical and laboratory features of each disorder.

ACUTE LYMPHOBLASTIC LEUKEMIA

ALL manifests with the clonal proliferation and accumulation of malignant lymphoid progenitors. ALL is a developmental disease. Most childhood ALLs arise as a "developmental accident" during normal fetal lymphopoiesis. Studies of chromosomal translocations in ALL cells have identified key genes involved in normal lymphopoiesis and hematopoiesis. Conversely, basic studies of the development of the immune system and the immune receptors have provided important tools for the diagnosis and management of ALL. These achievements in basic and clinical research have led to the remarkable transformation of ALL from a uniformly fatal disorder to a disease that is curable in more than 80% of children.[1] Adults, however, have not fared as well.

Epidemiology and Etiology

ALL is the most common malignancy of childhood. One of every 2000 children will develop leukemia by 15 years of age. In contrast, ALL accounts for <20% of leukemias in adults. In developed countries, incidence peaks at 2–5 years of age. This low age at peak of incidence is characteristic of affluent societies.[2]

Most ALLs are sporadic. Less than 5% are associated with hereditary or constitutional syndromes. Children with Down syndrome have an approximately 20-fold increased risk of developing ALL. Other diseases associated with increased risk include inherited genomic instability syndromes, such as ataxia–telangiectasia, Bloom syndrome, and Li-Fraumeni syndrome. Conversely, ALL is more common in patients with congenital B-cell immune deficiencies (Chapter 34). These include X-linked agammaglobulinemia (XLA), selective immunoglobulin A (IgA) deficiency, and common variable immunodeficiency (CVID). Interestingly, some cases of CVID associated with leukemia have been recently shown to be caused by germline mutations in *IKZF1,* which encodes the lymphoid transcription factor Ikaros.[3] Germline mutations in additional hematopoietic and B-cell developmental genes, such as *ETV6,*[4], *RUNX1,*[5] and *PAX5,*[6] also predispose to ALL.

Although ALL is not hereditary, there is a markedly increased risk of leukemias in identical twins, and studies in these twins have shed light on the etiology of childhood ALL. This phenomenon has promoted the hypothesis that at least two genetic hits are required for the development of ALL[2] (Fig. 78.1). The first occurs during fetal lymphopoiesis, most likely because of a "developmental accident" of the V(D)J recombination machinery, and results in clonal proliferation of a preleukemic clone. Intrauterine metastasis of this clone from one twin to the other via their shared placental circulation is responsible for the concordant leukemia. Additional genetic hits in the preleukemic cells occur after birth and are required for the development of full-blown leukemia. The initial findings in identical twins with leukemia have been extended to sporadic ALL. In at least 70% of the patients, the preleukemic clone can be detected molecularly in the neonatal blood samples collected after birth (known as *Guthrie cards*). More recently, careful molecular analysis of the cord blood of normal infants has demonstrated that the occurrence of a preleukemic clone carrying a leukemia-defining chromosomal translocation might be common. However, only 1% of children born with such a preleukemic clone will develop leukemia, implying the impracticality of molecular screen for early diagnosis of childhood ALL.

KEY CONCEPTS

Environmental Factors in the Epidemiology of Childhood Acute Lymphoblastic Leukemia (ALL): Roles for Infection and Immunity?

"Common" B-cell precursor ALL at the preschool age is the most common type of ALL in the suburban regions of affluent countries. The causes for this phenomenon are unknown. A popular hypothesis suggests a modified immune response to delayed infections during infancy.

The causes of the relatively rare postnatal leukemogenic genetic hits are unknown. As the risk of B-cell precursor ALL during early childhood is markedly increased by higher socioeconomic status and a suburban style of living in which the exposure of children to infectious pathogens is typically delayed beyond the neonatal period, Greaves hypothesized that many childhood cases are the consequence of an abnormally late immunological response to common infections.[2] One proposed mechanism is that growth inhibitory factors, such as interferon (IFN) or transforming growth factor-β (TGF-β), secreted during this immune response provide a survival advantage to a preleukemic clone setting the stage for additional leukemogenic mutations. Recently a more direct suggestion for the involvement of infection (or the response to infection) in the pathogenesis of ALL has been provided. Pax5 heterozygous mice developed preB-cell ALL only after exposure to common mouse pathogens.[7]

Immunological and Molecular Classification of ALL

Immunological Classification

ALL subtypes are typically identified by their immunophenotype. They are positioned by the lymphoid developmental stage at which the leukemic cell appears to have been arrested (Table 78.1).

B-Cell Precursor Leukemias

B-cell precursor ALLs are the most common childhood leukemias. **ProB-cell ALL** is characterized by expression of CD19 and CD34 without CD10. This is the most common leukemia of infants that contains rearrangements of the *MLL*(*KMT2A*) gene on chromosome 11q23. It is associated with a poor outcome.

Leukemic blast cells of **early preB-cell ALL** resemble normal B-lymphoid cell precursors. They express CD19, CD22, and CD79a. CD10 and terminal deoxynucleotidyl transferase (TdT) are detectable in 90% of cases. CD34 is detected in >75% of cases. Early preB-cell ALL is the most prevalent type of ALL and is thus often called "common ALL" (cALL).

In **preB-cell ALL** lymphoblasts accumulate cytoplasmic immunoglobulin (Ig) heavy chains but have no detectable surface Igs. This subtype also expresses CD19, CD22, and CD79a. It often contains a translocation that fuses the *E2A*(*TCF3*) gene on chromosome 19 with the *PBX1* gene on chromosome 1.

Leukemic cells that express both cytoplasmic and surface Ig heavy chains have been designated **transitional preB-cell ALL**. The surface μ chains on these leukemic cells are noncovalently associated with surrogate light chain (VpreB λ5[λ14.1]), CD79a, and CD79b.

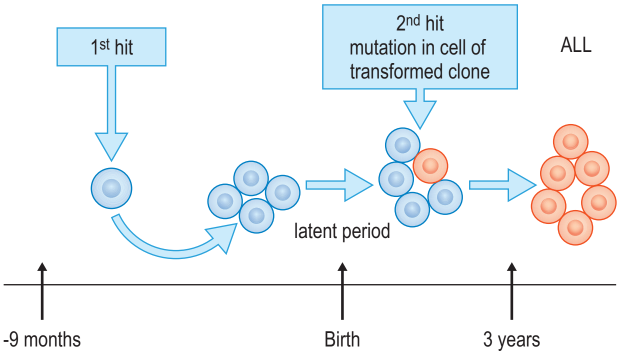


FIG 78.1 A Model for the Development of Childhood Acute Lymphoblastic Leukemia (ALL). The first acquired genomic hit (*e.g.,* chromosomal translocation or change in chromosomal copy number) occurs during fetal hematopoiesis and results in clonal proliferation of a preleukemic clone. This event is common, occurring in up to one in every 20 children. Additional genetic aberrations occurring after birth are required for the development of ALL. These events are rare and estimated to occur in about 1% of children born with a preleukemic clone.[5]

CLINICAL PEARLS

Mature Versus Precursor B-Cell Acute Lymphoblastic Leukemia (ALL)

Since the treatment of mature B-cell leukemia (the leukemic form of Burkitt lymphoma) is vastly different from the treatment of B-cell precursor ALL, it is critical to distinguish between the two. Burkitt leukemia is characterized by mature B-cell phenotype and by the presence of chromosomal translocations involving the *cMyc* gene (Chapter 79).

Mature B-cell ALL is the leukemic form of Burkitt lymphoma (Chapter 79). As treatment is dramatically different from that for B-cell precursor ALL, this subtype must be specifically ruled out as part of the immunophenotypic evaluation of ALL. Mature B-cell ALL cells express surface Ig μ heavy chains in association with either κ or λ light chains. In most cases, cells have L3 morphology and express CD19, CD22, and CD20. CD10 and CD23 are frequently expressed, whereas CD34 is absent.

T-Cell ALL

T-cell ALL is more prevalent in less affluent countries, which is likely a reflection of a lower incidence of the common B-lineage early childhood peak. In affluent countries, T-cell ALL occurs in 10–15% of children with ALL. It is also more common in adults. In many treatment protocols, T-cell ALL is considered to have a less favorable prognosis compared with B-cell precursor ALL.

In T-cell ALL, the cells express surface CD7 and cytoplasmic CD3 (cCD3) antigens. More than 90% of the T lymphoblasts express CD2, CD5, and TdT. Three stages of immunophenotypic differentiation can be determined: early ($CD7^+$, $cCD3^+$, surface $CD3^-$, $CD4^-$, and $CD8^-$), intermediate ($cCD3^+$, surface $CD3^-$, $CD4^+$, $CD8^+$, and $CD1^+$), and late (surface $CD3^+$, $CD1^-$, and either $CD4^+$ or $CD8^+$). However, immunophenotype does not conform to any of these maturation stages in as many as 25% of T-lineage ALL cases. T-cell receptor (TCR) proteins are heterogeneously expressed in T-lineage ALL. In approximately two-thirds of cases, membrane CD3 and TCR proteins are absent.

TABLE 78.1 Immunophenotypic Classification of Acute Lymphoblastic Leukemia (ALL)

	LEUKOCYTE ANTIGEN EXPRESSION (% OF CASES POSITIVE)									FREQUENCY (%)	
Subtype	**CD19**	**cCD22**	**CD79a**	**CD10**	**CD7**	**CD2**	**cCD3**	**cIg μ**	**sIg μ**	**sIg κ/λ**	**Children**	**Adults**
Pre-preB	100	>95[a]	>95	0	0	0	0	0	0	0	5	10
Early preB	100	>95[a]	>95	95	5	<5	0	0	0	0	60–65	50–55
PreB	100	100[a]	100	>95	0	0	0	100	0	0	20–25	10
Transitional preB	100	100[a]	100	50	0	0	0	100	100	0	1–3	?
B	100	100[a]	100	50	0	0	0	>95	>95	>95	2–3	4
PreT	<5	0	0–20	45	100	0	100	0	0	0	1	5
T	<5	0	0–20	45	100	95	100[a]	0	0	0	10–15	15–20

c, cytoplasmic; cIg μ, cytoplasmic immunoglobulin μ chain; sIg μ, surface immunoglobulin μ chain; sIg κ/λ, surface immunoglobulin κ or λ chains.

[a] Detectable on the cell-surface membrane in some cases.

TABLE 78.2 Frequencies of Major, Clinically Important Genetic Aberrations in Childhood and Adult Acute Lymphoblastic Leukemia (ALL)

Genetic Aberration	Children	Adults
B-Cell Lineage		
Hyperdiploidy (>50 Chromosomes)	30%	9%
Hypodiploid (<45 chromosomes)	1%	2%
Amplified 21q	2%	2%
TEL-AML1 (t12;21)	25%	3%
MLL rearrangements	9%	13%
BCR-ABL	4%	33%
E2A-PBX1	5%	4%
"Ph like" including CRLF2	8%	25%
MYC rearrangements	2%	5%
T-Cell Lineage		
Notch1 mutations	60%	70%
TAL1 (SCL) cluster	58%	33%
HOX11 (TLX1) cluster	3%	33%
HOX11L2 (TLX3) cluster	20%	5%
LYL1 cluster	12%	37%
MLL-ENL	2%	2%
NUP214-ABL	6% (?)	

In half these cases, however, TCR proteins (TCRβ, TCRαα, or both) are present in the cytoplasm. When membrane CD3 and TCR chains are expressed, the αβ form of the TCR predominates. Only a minority of cases express TCR γδ proteins.

The expression of myeloid-associated antigens, such as CD13, CD15, CD33, or CD65, in leukemic lymphoblasts confounded leukemia classification in early studies. However, the diagnosis of B-lineage ALL should be made when leukemic cells express cytoplasmic Ig, CD79a, or CD19 plus CD22, regardless of myeloid-associated antigen expression. Likewise, the diagnosis of T-lineage ALL should be made when leukemic cells express CD7 plus either surface or cytoplasmic CD3. An immature immunophenotype of T-cell ALL, early T phenotype (ETP) ALL, has been described and is similar to normal ETP precursors.[8] True biphenotypic leukemias usually coexpress B- or T-cell markers with cytoplasmic myeloperoxidase.[9]

Genetic and Molecular Classification

Virtually every leukemic cell contains acquired alterations in multiple genes. These alterations often manifest gross numerical or structural aberrations that frequently define a specific clinical subtype of ALL. Common and/or clinically significant genetic aberrations typically found in ALL are summarized in Table 78.2.

Numerical Chromosomal Aberrations

Deviation from the normal chromosomal modal number is called *aneuploidy* and is the most common chromosomal aberration in cancer. **High hyperdiploid ALL** (Fig. 78.2A), containing 50–60 chromosomes, is the most common type of B-lineage ALL in children and is associated with about 90% cure rate. There is typically an excess of specific chromosomes, most commonly chromosomes 6, 10, 14, 17, 18, 21, and X. **Hypodiploid ALL** (see Fig. 78.2B), containing <45 chromosomes, is much rarer and associated with a very poor prognosis.

Chromosomal Translocations

Chromosomal translocations can be divided into two general subtypes. The first involves translocation of a gene into the proximity of a strong regulatory region. This often results in the marked overexpression of the oncogene. Often these translocations are mediated by the V(D)J recombination machinery (Chapter 4) and can be therefore viewed as an unfortunate developmental "accident" caused by the physiological lymphocyte specific genomic instability. Examples include the activation of the *MYC* oncogene by t(8;14) in Burkitt lymphoma and of the *SCL* (*TAL1*) gene by t(1;14) or by rearrangement with the *STIL* gene on chromosome 1p32 in T-cell ALL. The second involves creation of novel fusion proteins consisting of parts of the genes that participate in the chromosomal translocation. For example t(12;21) fuses the *TEL* (*ETV6*) gene on chromosome 12 with the *AML1* (*RUNX1*) gene on chromosome 21.

Amplifications and Deletions

Amplifications and deletions of small chromosomal regions are another type of structural aberration that is often detected in leukemias. For example, deletions of the INK4A locus or of PAX5 are commonly detected in T- and B-cell precursor ALLs, respectively.

Oncogenic-Activating Mutations

Oncogenic-activating mutations in ALL are reported with an increasing frequency. The Notch pathway, for example, is normally involved in T-cell development. This pathway is activated by acquired mutations in >60% of T-cell ALL.

Genes Involved in Leukemogenesis Often Play Key Roles in Normal Development

Many of the genes modified by chromosomal translocations, amplification, deletions, or point mutations are active during normal lymphoid or hematopoietic development (Table 78.3). The acquired aberrations promote malignant transformation by either overexpression or dysregulated expression (in the wrong cell or in the wrong developmental stage) of the developmental gene.

A recent discovery involving the IL7Rα serves as an example. Loss of function mutations in this receptor causes $T^-B^+NK^+$ severe combined immunodeficiency (Chapter 35). Activating mutations are found in 10% of T-cell ALL.[10] An example involving both misexpression and activation of an immune receptor is the aberrant activation of the receptor to thymic stromal lymphopoietin (TSLP). This receptor, created by heterodimerization of IL7Rα and CRLF2, is normally expressed in dendritic cells (DCs) and some CD4 T cells. It is involved in allergy and inflammation. Activation of this receptor in B-lineage ALL is multistaged. First, chromosomal rearrangements cause the misexpression of CRLF2 in a B-cell precursor. This is followed by activating mutations in the TSLP receptor components—CRLF2, IL7Rα, or the downstream signaling enzymes JAK2 or JAK1.[10] The aberrant expression and activation of the TSLP receptor occurs in 10% of sporadic B-cell ALLs and in 50% of Down syndrome ALLs. It is associated with a poor prognosis. It may also be amenable to targeted therapy. Conversely, the acquired genetic aberration may block the normal developmental function of the involved gene(s). A good example is the genomic deletions or inactivating mutations in B-cell differentiation genes, such as *PAX5*, *EBF*, or *IKZF1*, detected in about 50% of B-lineage ALLs.

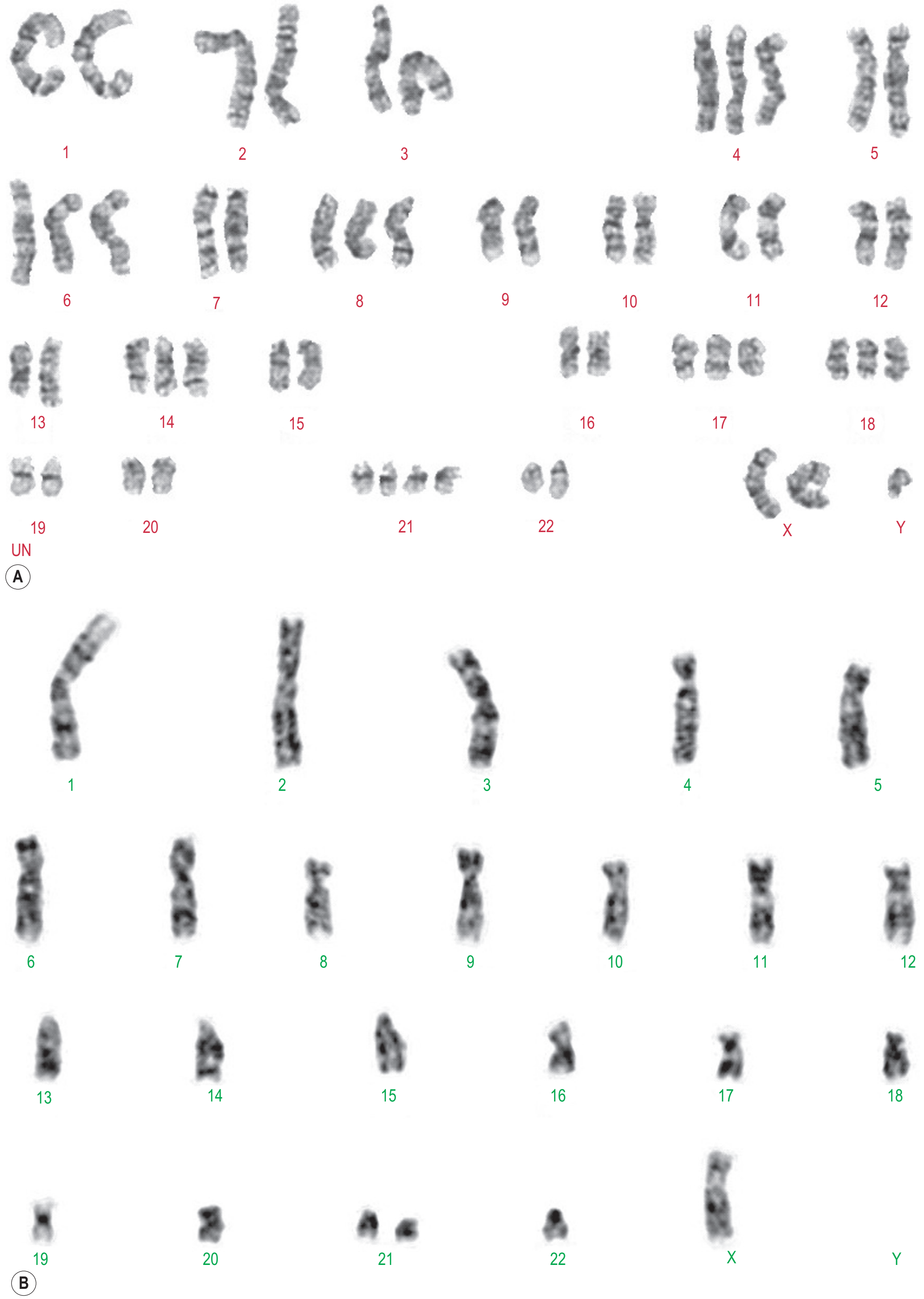

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
X
Y
UN
A
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
X
Y
B

FIG 78.2 Chromosomal Aneuploidy in Acute Lymphoblastic Leukemia (ALL). A typical karyotype of (A) hyperdiploid and (B) hypodiploid ALL. (Courtesy of B. Stark and D. Betts.)

TABLE 78.3 Example of Hematopoietic Genes Involved in the Pathogenesis of Leukemia

Gene(s) Names	Normal Hematopoietic Development	Leukemic Involvement
SCL (TAL1)	Hemangioblast specification Erythropoiesis and megakaryopoiesis	T-ALL
LMO1/2	Similar to SCL	T-cell ALL
NOTCH1	T lymphocytes	T-cell ALL
HOX11	Spleen	T-cell ALL
E2A	T and B lymphocytes	BCP-ALL
PAX5	B lymphocytes	BCP-ALL, B-NHL
SLP-65	B lymphocytes	BCP-ALL
TEL	Bone marrow hematopoietic stem cells	BCP-ALL, T-cell ALL rarely myeloid malignancies
RUNX1 (AML1, CBFA2)	Definite hematopoiesis Megakaryopoiesis and T lymphocytes	BCP-ALL, AML (M0-M1) Hereditary FPD/AML
CBFB	Same as RUNX1	AML (M4e)
C/EBP 1-3	Myeloid cells	AML (M1, M2)
PU.1	Myeloid and lymphoid stem cells	AML
GATA1	Erythropoiesis, megakaryopoiesis, and mast cells	AML (M7) associated with trisomy 21
FLT3	Hematopoiesis and lymphopoiesis	AML and ALL
MLL	Hematopoiesis stem cells	AML and ALL
IL7R	T lymphocytes	ALL

Major, Clinically Relevant, Molecular Subtypes of ALL

(See Table 78.2.)

B-Lineage ALL

Hyperdiploid ALL and ALL cells *TEL/AML1* gene translocation comprise the majority of "common ALL" leukemias that occur typically in young children and rarely in adults. Both are associated with an extremely good prognosis. In contrast, the much less common hypodiploidy (<45 chromosomes) and the internal amplification of chromosome 21q (iAMP 21q) are associated with a poor prognosis.

Another aberration with a poor prognosis is the BCRF-ABL fusion protein created by the t(9;22) "Philadelphia chromosome." Its frequency is low in children (3–5%) and much higher in adults (at least 30%). The prognosis dramatically improved by treatment with BCR–ABL1 inhibitors.

The *MLL(KMT2A)* gene located on chromosome 11q23 is involved in fusion translocations with >80 different partner genes. The most common translocation in ALL fuses the *MLL* with the *AF4* gene on chromosome 4. It is characteristic of infant leukemia and is associated with a poor prognosis.

The clinical significance of the E2A (TCF3)-PBX1 fusion protein caused by t(1;19) translocation, occurring in <5% of childhood ALL, has greatly improved by more intensive chemotherapy. The change in the prognostic significance of this particular translocation with improved therapy exemplifies the general principle that the prognostic impact of a clinical or biological parameter is highly dependent on the specific treatment protocol. E2A-PBX1 ALLs also represents a subtype of ALL with activation of signaling downstream to the preB-cell receptor (pre-BCR). Recent therapies with BCR signaling inhibitors effective in CLL or lymphomas may prove useful for pre-BCR ALLs.[11] The *E2A* gene is rarely fused with the *HLF* gene in the t(17;19) translocation. This is an extremely rare translocation that is associated with diffuse intravascular coagulation (DIC), hypercalcemia, and an extremely poor prognosis.[12]

Several genetic lesions have prognostic significance and may be included in risk stratification in future treatment protocols. Examples are deletions of Ikaros (IKZF1) and CRLF2 aberrations. A novel subgroup of ALLs has been termed "Philadelphia-like ALL" because of gene expression that is similar to BCR-ABL. It is generally associated with markedly poor prognosis.[13,14] These leukemias are characterized by activation of either the ABL or the JAK kinase pathway and are caused by many types of fusion genes or aberrant expressed receptors. Their proper diagnosis is important because of the potential therapeutic activity of either ABL or JAK inhibitors.

T-Lineage ALL

Although multiple genetic and molecular subtypes of T-cell ALL have been recently described, their clinical significance is presently unclear, except for the bad prognosis of ETP ALL.[8] Most of the genetic aberrations in T-cell ALL results in the abnormal expression of transcription factors. Examples include SCL (TAL1), which forms a complex with LMO1 or LMO2, HOX11L2 (TLX3) or HOX11 (TLX1), LYL1, and MYB. Other abnormalities include MLL–ENL fusion and amplification of the ABL oncogene. Cooperating mutations are activating mutations in NOTCH1 or interleukin-7R (IL-7R) and inactivating mutations in the E3 ligase FBW7 or the phosphatase PTEN.

Clinical Features

The clinical signs and symptoms relate to the replacement of bone marrow cells by leukemic blasts and to the infiltration of extramedullary sites. Pallor, fatigue, petechiae, bleeding, or fever may be caused by the pancytopenia. Bone pain and arthralgias, the onset of a limp and refusal to walk, and even frank arthritis are not uncommon. The musculoskeletal symptoms may be confused with osteomyelitis or juvenile rheumatoid arthritis, which may delay the diagnosis. Uncommonly, central nervous system (CNS) involvement may present as signs and symptoms of increased intracranial pressure, such as headaches and papilledema; or as cranial nerve palsies, nuchal rigidity, and rarely hyperphagia and obesity caused by infiltration of the hypothalamus. Overt testicular leukemia manifests as painless testicular enlargement. Mediastinal involvement, common in T-cell ALL, may cause dyspnea and superior vena cava syndrome.

Clinical laboratory findings often include anemia and thrombocytopenia. Approximately 20% of children present with leukocyte counts >50 000/µL. Importantly, approximately 40% of children have leukocyte counts of <10 000/µL, and leukemic blasts may or may not be seen on peripheral smears. Therefore the diagnosis of leukemia may occasionally be missed in routine automated blood count. Elevated serum lactate dehydrogenase activity and uric acid and phosphorous concentrations are common in patients with a large leukemic cell burden.

The diagnosis of ALL is established by bone marrow examination. The normal bone marrow contains <5% blasts. A minimum of 25% lymphoblasts on differential examination of the bone marrow aspirate is necessary for the diagnosis of ALL. Most children with ALL have a hypercellular marrow with blasts constituting 60–100% of the nucleated cells.

Traditionally, CNS leukemia has been defined as the presence of at least five leukocytes per microliter of cerebrospinal fluid (CSF) and the detection of leukemic blast cells, by the presence of cranial nerve palsy, or by retinal involvement, as detected by ophthalmoscopy. Although overt CNS leukemia is relatively rare, submicroscopic CNS involvement is present at diagnosis in at least half the patients in the absence of any neurological symptoms. Thus CNS-directed therapy is routinely included in ALL therapy.

The differential diagnosis of ALL includes neoplastic and nonneoplastic diseases. Because children with ALL present with a variety of nonspecific symptoms, several pediatric nonmalignant conditions may be confused with leukemia. Since treatment with steroids may mask the presence of ALL, serious consideration of the diagnosis of ALL must be given before starting treatment with steroids to any pediatric nonmalignant disorder. Bone marrow examination is recommended in case of uncertainty.

Idiopathic thrombocytopenic purpura (ITP) is a common cause of bruising and petechiae in children. ITP is characterized by the absence of any other hematological abnormalities. Bone marrow should be examined if anemia or hepatosplenomegaly are present.

Infectious mononucleosis may present with fever, malaise, adenopathy, splenomegaly, rash, and lymphocytosis. The atypical lymphocytes may morphologically resemble the leukemic lymphoblasts. Rarely, flow cytometry may be necessary to distinguish between the activated atypical lymphocytes and the immature leukemic lymphoblasts.

Leukemoid reactions, observed in sepsis, acute hemolysis, and other disorders, are usually easy to distinguish from ALL by morphological examination of peripheral blood smear. Since occasionally ALL presents with pancytopenia, aplastic anemia is also on the differential diagnosis list.

CLINICAL PEARLS

Acute Lymphoblastic Leukemia (ALL) and Rheumatoid Disorders

- ALL can mimic juvenile idiopathic arthritis (JIA; Chapter 53) and other musculoskeletal disorders.
- Because leukemic blasts may be absent from the peripheral blood, bone marrow examination should be considered in any child with JIA, especially prior to commencing steroid therapy.

As many as 10% of children with ALL are first evaluated at pediatric rheumatology clinics. Fever, arthralgias, arthritis, or a limp accompanied by anemia, mild splenomegaly, and lymphadenopathy frequently can be confused with juvenile idiopathic arthritis (Chapter 53) or osteomyelitis. These patients may be treated with antibiotics and antiinflammatory agents for several weeks to months before the diagnosis of ALL is finally made. Bone marrow examination should be seriously considered in such patients.

As leukemic lymphoblasts are seen as small round blue cells when stained with hematoxylin and eosin, they may be rarely be confused with metastatic small round cell pediatric tumors, including neuroblastoma, rhabdomyosarcoma, and retinoblastoma.

Special Diagnostic Tests

The classification and the risk stratification for treatment protocols of ALL is based on detailed immunophenotyping and genotyping analysis.[15] In most centers, karyotype by classic cytogenetics (see Fig. 78.2) is performed. However, because of the lack of metaphases, the yield of karyotypic analysis of ALL in multicenter protocols is often <70%. Furthermore, normal karyotypes are sometimes derived from the normal cells in bone marrow rather than the leukemia blasts. This drawback can be overcome by interphase fluorescent in-situ hybridization (FISH), a technique that does not require metaphases. All clinically relevant structural and numerical chromosomal aberrations can be detected with the use of commercially available FISH probes (Fig. 78.3). Fusion translocations, such as *BCR–ABL*, *TEL–AML1*, and *MLL–AF4*, can all be detected by using reverse transcription–polymerase chain reaction (RT-PCR).

The elucidation of the human genome and the invention of the genomic technologies are in the process of revolutionizing the diagnostics of leukemias. New methodologies of next-generation genome sequencing (NGS) are likely to transform both our understanding of leukemia biology and the diagnostic approach. Routine use of NGS mutation panels and copy number genomic arrays are likely to replace routine cytogenetic analysis in the near future.

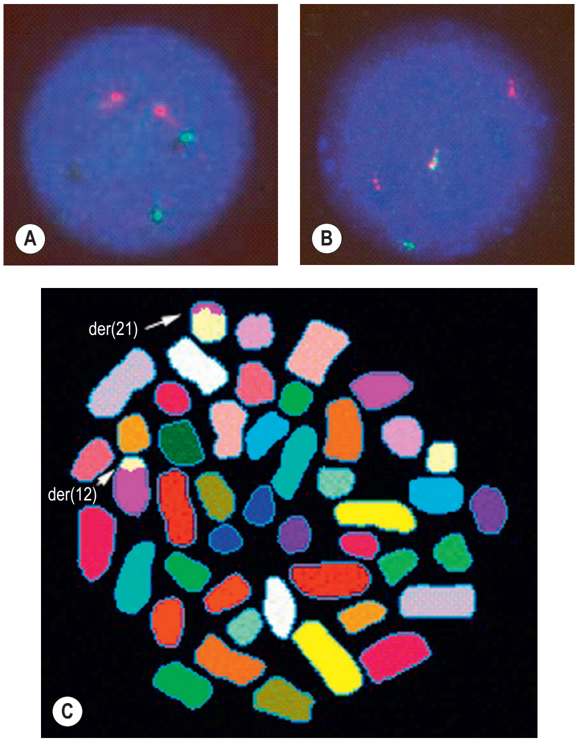


FIG 78.3 Molecular Cytogenetic Techniques for the Diagnosis of Chromosomal Translocations in Acute Lymphoblastic Leukemia (ALL). (A) and (B) display interphase fluorescent in-situ hybridization (FISH) with probes to the *AML1* (*RUNX1*) gene on chromosome 21 (red) and to the *TEL* (*ETV6*) gene on chromosome 12 *(green)*. (A) displays a normal cell, (B) displays a leukemic cell that has undergone a fusion TEL-AML1 translocation *(arrow)*, and (C) displays the same translocation as depicted in B, but on metaphase chromosomes. It uses a molecular cytogenetic technique called spectral karyotyping *(arrows)*. Classic cytogenetic analysis often misses this translocation. (Courtesy of Dr. L. Trakhtenbrot.)

Principles of Therapy[1]

Supportive therapy is given before the initiation of leukemia specific therapy. This includes hydration, treatment, and preventive therapy of hyperuricemia; blood and platelet transfusion; and treatment of emergencies, such as respiratory insufficiency associated with mediastinal leukemia. It is highly recommended that children and adults with ALL be treated in specialized centers as a part of clinical prospective studies. Such clinical trials have led to the dramatic improvement in the outcome of patients with ALL achieved over the last several decades.

Typical remission induction regimens include a glucocorticoid (prednisone, prednisolone, or dexamethasone), vincristine, l-asparaginase, and, in many protocols, anthracycline. The rate of complete remission now ranges from 97% to 99% in children and from 75% to 90% in adults. Intensification therapy uses either high doses of multiple agents not used during the induction phase or repeats the induction regimen. Regimens used for children include high-dose methotrexate with or without 6-mercaptopurine; high-dose l-asparaginase given for an extended period; an epipodophyllotoxin plus cytarabine; or a combination of dexamethasone, vincristine, l-asparaginase, doxorubicin, and thioguanine, with or without cyclophosphamide. Another integral component of many protocols is reinduction therapy. This treatment employs drugs similar to those used during the initial phase of induction therapy and has improved the outcomes of both children and adults with ALL. Maintenance therapy consist a combination of methotrexate administered weekly and mercaptopurine administered daily. Some protocols add intermittent pulses of vincristine and dexamethasone.

CNS prophylactic therapy consisting of cranial irradiation plus intrathecal chemotherapy, introduced after the induction of complete remission, became one of the cornerstones of ALL therapy in the 1970s. More recently, because of concerns about neurotoxicity and the occurrence of brain tumors, intensive intrathecal and systemic chemotherapy are used instead for most patients.

Allogeneic stem cell transplantation (SCT) is reserved for relapsed or refractory leukemia and for patients with a very-high-risk leukemia that is likely to demonstrate a slow response to therapy. With modern therapy, including targeted therapy with imatinib for BCR–ABL, there is almost no indication for SCT in first remission. With improvements in the prevention of transplant-related toxicities, suitable marrow donors are not only human leukocyte antigen (HLA)–identical siblings or single-antigen mismatched family members but also matched unrelated donors and, in some situations, two- and three-antigen mismatched family members.

Prognostic Factors

Modern therapy for ALL is based on adjustment of the intensity of therapy to the risk assessment of the relapse hazard (Table 78.4). The two major clinical parameters of prognostic significance are age at diagnosis and leukocyte count. A presenting age between 1 and 9 years and a leukocyte count $<50 \times 10^9$/L are favorable prognostic factors. In adults, the outcome of therapy worsens with increasing age and leukocyte count. Patients >60 years of age and/or a leukocyte counts $>100 \times 10^9$/L have a particularly poor response to treatment. Females fare somewhat better compared with males. The prognostic significance of the major genetic aberrations has been described above.

TABLE 78.4 Major Prognostic Factors in Acute Lymphoblastic Leukemia (ALL)[a]

Prognostic Factor	Good Prognosis	Worse Prognosis
Age at diagnosis	Age 1 year to <10 years (children)	<1 year; >10 year (children) >60 years (adult)
Peripheral blood WBC	<50 000 cells/μL	>100 000 cells/μL
Response to therapy	Early response to therapy Negative MRD at the end of induction	Slow response to therapy. High MRD
Genetic abnormalities	Hyperdiploidy (>50 chr.); *TEL/AML1 (ETV6/RUNX1)*	*BCR/ABL MLL/AF4* Hypodiploidy <45 chr.

[a] The most important prognostic factor is the *treatment protocol*. Thus the prognostic significance of various clinical and laboratory variables may differ between protocols. Here, the significant parameters common to most studies are listed.
WBC, white blood cells; MRD, minimal residual disease; chr., chromosomes.

The most significant prognostic factor is the initial response to therapy. A rapid clearance of leukemic cells from blood or bone marrow confers a favorable prognosis. More recently, the level of minimal residual disease after the induction of clinical remission has emerged as a powerful tool for gauging treatment response and predicting outcome.

Where Immunology Meets Oncology—Minimal Residual Disease

Modern treatment protocols have led to morphological complete remission in the majority of patients, which is defined as <5% blasts in bone marrow examination. If treatment is discontinued at that stage, in most patients the disease will eventually relapse. Indeed, all prospective clinical studies have shown that ALL should be treated for at least 2 years. These facts indicate that at the completion of remission induction not all clonogenic malignant lymphoblasts have been destroyed, even though most of the patients are in clinical and morphological remission. Indeed, by this criterion, patients may have as many as 10^{10} undetectable neoplastic cells when in remission. Since, by definition, leukemic cells must constitute at least 1–5% of the nucleated cells in bone marrow to be detected by microscopic examination, morphological examination is clearly inadequate for evaluation of the quality of remission in patients with ALL. Therefore more sensitive techniques for the detection of rare leukemic cells are required. This is the rationale behind the recent incorporation of modern techniques of detection of minimal residual disease (MRD) into treatment protocols of childhood ALL.

During the last 2 decades, two general methodologies have been developed for the sensitive detection of submicroscopic residual leukemic cells. These methodologies could not have been developed without elucidation of the scientific basis of the developmental phenotype of immune cells (Chapters 7 and 8) and of the elaborate process of Ig gene rearrangements (Chapter 4).

The most widely studied DNA-based MRD methodology is based on the identification of clonospecific rearrangements of Ig genes or TCRs (Ig/TCR-PCR).[16] This approach exploits the physiological process of somatic rearrangement of Ig and TCR gene loci that occur during the early differentiation of B cells and T cells. Thus any single T or B lymphocyte carries a unique rearrangement that is not shared by any other lymphoid cell.

Since leukemia is clonal (*i.e.*, it originates from one lymphoid cell), all of the leukemic cells of a particular person carry the same Ig and/or TCR rearrangements. Because leukemic cells are genetically unstable, they often (>90% of the cases) carry multiple rearrangements, a fact that facilitates the usefulness of using these rearrangements as a clonal marker for MRD detection. The major advantages of this technique are the exquisite sensitivity (at least 10^{-5}), reliability, reproducibility, and its applicability to >90% of children with ALL. The application of NGS techniques is likely to lower the costs and complexity of this approach.

Current strategies for flow cytometric detection of MRD rely on combinations of leukocyte markers that do not normally occur in cells of the peripheral blood and bone marrow. Such leukemia-associated phenotypes can be identified by multiple color staining techniques (Fig. 78.4). Flow cytometric analysis of these immunophenotypes allows the detection of one leukemic cell among 10^{-4} or more normal cells. The advantages of flow cytometry for MRD detection are adequate sensitivity, and the presence of immunophenotyping facilities in most major centers that facilitate timely performance of the analysis on fresh cells at a reasonable cost. There are, however, a few disadvantages. It requires a flow cytometry operator with a high level of expertise, and it is more difficult to standardize compared with PCR. It may be difficult to distinguish between regenerating B-cell progenitors and leukemic blasts. Thus flow MRD is most reliable and sensitive at very early stages (up to 4 weeks) of therapy.

Many clinical studies of patients enrolled on ALL treatment protocols have revealed strikingly similar results. Fast clearance of leukemic blasts to below 10^{-4} cells within the first 2–4 weeks of therapy is detected in ≈40% of children with ALL and is associated with an extremely good prognosis. Conversely, the presence of >0.1% blasts after 2 or 3 months of therapy defines a very-high-risk group. These initial findings have led the BFM-AEIOP study group to initiate a prospective study on children with ALL that utilize the MRD level determined by Ig/TCR-PCR for risk classification. This study, which involved 3184 patients, confirmed the strength of PCR MRD over genetic classification as a prognostic marker.[17] At present, MRD studies have been incorporated into most treatment protocols. Patients with high MRD are stratified into a high-risk arm and receive more intensive chemotherapy.

Course and Prognosis

With current protocols, the cure rate of children with ALL is ≈80%. More than 90% of the children with low-risk leukemia and ≈70% of those with high-risk leukemia are cured. The cure rate for adults is significantly lower.

Most relapses occur during treatment or within the first 2 years after completion of therapy. However, late relapses (even 10 years after diagnosis) are increasingly seen. The clinical relapse can occur in an extramedullary site, most often the CNS and the testes. Leukemic relapse occasionally occurs at other sites, including the eye, ear, ovary, uterus, bone, muscle, tonsil, kidney, mediastinum, pleura, and paranasal sinuses. The prognosis of relapse depends on the time from diagnosis (earlier is worse), leukocyte count, immunophenotype, and genotype (similar to

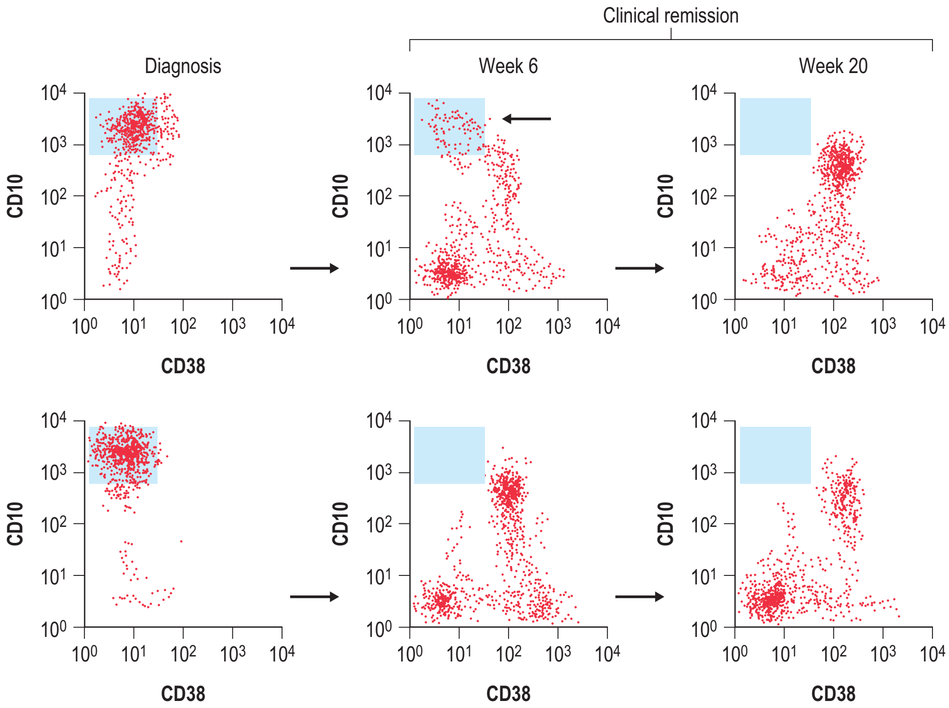


FIG 78.4 Different Kinetics of Leukemia Cytoreduction Revealed by Minimal Residual Disease (MRD) Studies With Flow Cytometry. The left panels illustrate the leukemia-specific immunophenotype ($CD10^+$, $CD38^-$) determined at diagnosis in two children with ALL. This phenotype is not found in normal bone marrow. Bone marrow samples were collected during clinical remission from both patients. In one patient *(top panels),* 0.04% of mononuclear cells expressed the leukemia-specific phenotype at week 6. MRD was undetectable by week 20. In the other patient *(bottom panels),* a profound remission (MRD <0.01%) was achieved by week 6 and maintained at week 20.

the prognostic factors at diagnosis). Isolated extramedullary relapse carries a better prognosis. Although some relapses can be treated with chemotherapy only, most are being treated with SCT. The most important prognostic factor that has emerged is the MRD level before transplantation. More than half the late relapses in patients with "good prognosis" ALL are curable. The prognosis of early relapses and relapsed T-cell ALL is grim.

Treatment Sequelae

In the 1990s, improvements in supportive care reduced the rate of early death to <2%. However, the death rate among older patients during remission induction therapy remains as high as 30% because of hematological, hepatic, and cardiac toxicities.

Long-term toxicity is of major concern for children cured from ALL. Aseptic necrosis of various bones has emerged as a common late toxicity of glucocorticoids, especially in adolescent girls treated with dexamethasone. The rate of long-term neurotoxicity has been reduced as a result of replacement of cranial irradiation with high-dose and intrathecal methotrexate. However intensive methotrexate therapy can also have late neurodegeneration effects. The dose of anthracyclines used in most ALL protocols is unlikely to produce severe cardiomyopathy.

Secondary cancers are a major concern. The development of therapy-related acute myeloid leukemia (AML) has been linked to the use of topoisomerase II inhibitors (teniposide and etoposide), and the risk is apparently dependent on the treatment schedule and the concomitant use of other agents. Children who received cranial irradiation at 6 years of age or younger are most susceptible to the development of brain tumors.

Long-term follow-up studies reported from Scandinavia of children cured of ALL confirm excellent long-term outcomes, as judged by their socioeducational achievements and by reproductive success resulting in healthy offspring.[18]

Current Controversies and Future Perspectives

The high cure rates in patients with ALL (nearly 90% of children and 50% of adults) attest to the steady progress that has been made in treating this disease. A further increase in cure rates will require efforts to maximize the efficacy and minimize the toxicity of current therapy. Currently MRD measurement can reliably identify the extremely good-risk and extremely poor-risk patients. However, about half the patients belong to an "intermediate risk" group in which most of the relapses occur. Optimization of diagnosis and treatment of this group of patients is one of the current major challenges.

Advanced genomic technologies carry the promise of discovering the full spectrum of leukemogenic pathways and the identification of targets for new therapies. Genomic analysis has also discovered marked subclonal heterogeneity.[19] Most of the relapses apparently arise from a minor subclone present at diagnosis. The identification and specific targeting of these resistant cells is a major future challenge.

An exciting development is the advent of immunotherapy. Blinatumomab, a novel bispecific antibody engaging autologous T cells against CD19 positive leukemic cells, has been recently approved for therapy.[20] New cell therapy based on autologous T cells engineered to express anti-CD19 receptor coupled to intracytoplasmic stimulatory molecules (CAR-T cells) have also shown a tremendous efficacy against highly resistant B-cell precursor ALL.[21] This is therefore an exciting period in novel treatments of ALL beyond chemotherapy.

ON THE HORIZON

Translational Research of Acute Lymphoblastic Leukemia (ALL)

- Whole genomic analyses will discover all leukemia genetic abnormalities, enabling further personalization of treatment.
- Novel therapies targeting specific leukemia-associated abnormalities will be implemented.
- Immunotherapy utilizing antibodies directed against lymphoid antigens and conjugated either to toxins or to T-cell engaging molecules or modified T cells will be routinely implemented in treatment of high-risk ALL.

CHRONIC LYMPHOCYTIC LEUKEMIA

Chronic lymphocytic leukemia (CLL) is an often indolent lymphoproliferative neoplasm of mature peripheral circulating B cells. It is the most common leukemia in adults living in countries in the Western hemisphere. CLL originates from a clonal lymphoid evolved mature stem cell that can be identified by its distinct B-cell Ig gene rearrangement. Both clinically and at the molecular level, it is a heterogeneous disease. Some patients have an indolent course, whereas other have a more rapid and aggressive disease. During its progression, CLL may be associated with significant immune deficiencies and autoimmune phenomena that can complicate its course and treatment. These abnormalities may be profound and thus alter the nature of the disease. They are attributed to the clonal nature of its B-cell origin. Less than 5% of patients have T-cell CLL. Understanding the molecular pathways involving BCR moieties enables focusing of novel and targeted chemoimmunotherapy.

Epidemiology

The extensive use of automated peripheral blood lymphocyte counts has increased the rate of diagnosis of asymptomatic patients with CLL. The incidence rate of CLL increases logarithmically from age 35 years, with a median age at the time of diagnosis of 65 years. There is a male predilection, and the disease appears to have geographical and ethnic variations in incidence. In the United States, CLL is uncommon in people of Asian descent. Because many of these patients may never require tissue diagnosis or inpatient treatment, cases among them are not likely to be recorded in a tumor registry, thus making the true annual incidence of the disease higher than previously thought (6.8 per 100 000 population).[22] With sensitive techniques, a monoclonal population of B lymphocytes that is indistinguishable in immunophenotype from CLL cells may be found in the blood of 3.5% of persons >40 years of age.[23]

The presence of B lymphocytes <5000/µL showing clonality is defined as monoclonal B-lymphocytosis (MBL). MBL is an indolent disorder that may progress to frank CLL at a rate of 1–2% per year, and almost all patients with CLL begin with MBL. This state is divided into low or high MBL (below or above 500/µL clonal lymphocytes in the peripheral blood). Low MBL is more abundant and rarely progresses to frank CLL, whereas high MBL may progress at the above-mentioned rate.[24] Some patients with MBL will experience autoimmune phenomenon, as later described for CLL. However, the importance of the MBL clone is not fully understood and does not necessarily imply a future progression to CLL.[24]

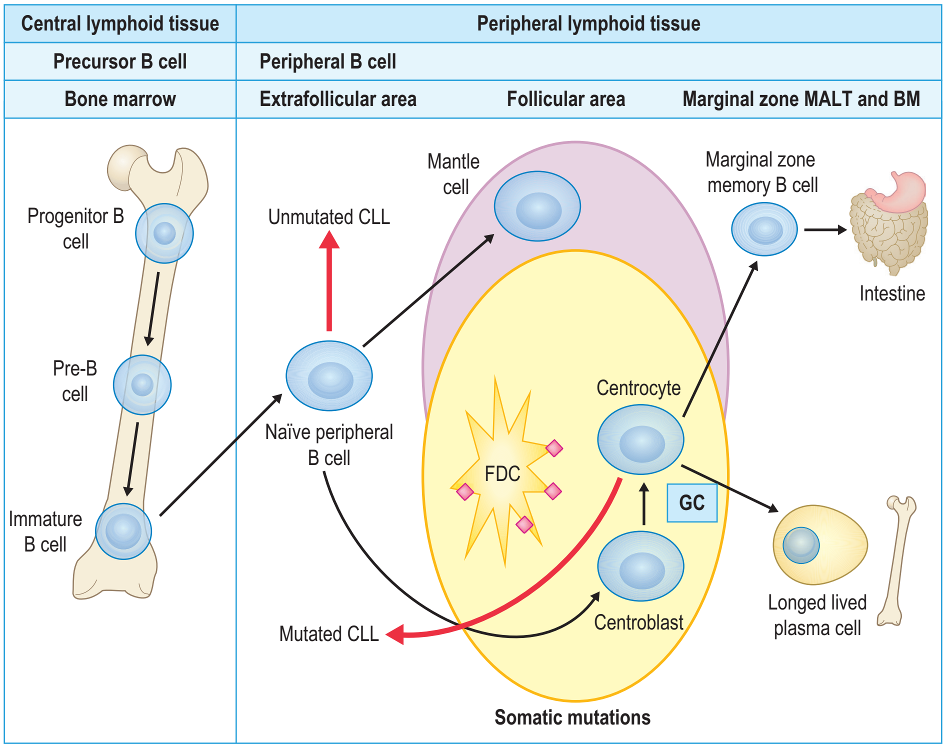


FIG 78.5 **B-Cell Development and Neoplasia.** B-cell lymphoproliferative disorders are related to different stage of normal B-cell development. Maturation starts at bone marrow. The more mature cells exit to the peripheral blood but do not experience the somatic hypermutation in the germinal center (GC), hence giving rise to the unmutated chronic lymphocytic leukemia (CLL) form. When the B cells mature and differentiate within the GC into the hypermutated centrocytes, they will pathologically become the mutated form of CLL. Other forms of B-cell neoplasms may arise at other stages. ProB and preB cells give rise to acute lymphoblastic leukemia; intermediate stages give rise to mantle zone lymphoma; and late and more mature stages give rise to marginal zone lymphomas and plasma cell–derived multiple myelomas.

The etiology of CLL is still unknown. However, as with other forms of malignancy, there is increasing evidence for the role of inherited factors in its development. Family surveys show a genetic predisposition in first-degree relatives, who also demonstrate increased susceptibility to other lymphoproliferative disorders, including other lymphomas.[25] Anticipation, the phenomenon of earlier onset, and a more severe phenotype in successive generations have been reported in families of patients with CLL.

Pathogenesis and the Biology of Leukemic Lymphocytes

CLL is now viewed as two related entities, both of which originate from B lymphocytes but differ in their activation state and in the maturation state of the cellular subgroup.[23.]

Normal B lymphocytes mature in bone marrow (Chapter 7). In the process, they undergo rearrangement of Ig V(D)J gene segments to create the code for an Ig molecule that serves as the BCR for antigen (Chapter 4). When an antigen of adequate affinity engages the receptor, the cell enters a germinal center located in a lymphoid follicle. There, as a centroblast, it rapidly divides and the V domains of its Ig undergo somatic hypermutation. Cells with receptors that have enhanced antigen-binding affinity proliferate in the presence of the antigen, whereas centrocytes with receptors that no longer bind the antigen (or bind autoantigens) are normally eliminated. Once the centrocytes are selected, they become plasma or memory B cells (Fig. 78.5).

The sequence and structure of Igs expressed by CLL cells from various patients have been found to be similar, suggesting a common pathogenetic antigen, such as an autoantigen-provoking clonal expansion. These findings suggest that antigenic stimulus plays a role in the promotion of CLL proliferation.

CLL may originate from a clone with few or no V domain mutations or from a more mature clone where these domains are hypermutated. These differences in the extent of V domain mutation suggest differing entities with two different developmental histories. Both originate from antigen-stimulated mature B lymphocytes. However, in the CLL with mutated Ig V domains, the proliferating B cell may have traversed the germinal centers, whereas in CLL with unmutated Ig V domains, the malignant B cell may derive from a naïve, pre–germinal center B cell.

The mutational status of the V domains strongly correlates with prognosis in that patients with an unmutated clone have a much worse prognosis compared with patients with mutated clones. These patients may differ also in their association with specific genetic aberrations. 11q22-23 (the ataxia–telangiectasia mutated gene) or 17p13 (the *p53* gene) deletions are associated with poor outcome and with an unmutated V domain profile. These genes regulate apoptosis and resistance to chemotherapy.

13q14 deletion or a normal karyotype is associated with a mutated profile and better prognosis. These chromosomal aberrations have independent prognostic significance (unrelated to the mutational status).

Surface membrane antigens typically found on B-cell CLL cells include CD19, CD21, and CD23. Expression of membrane IgM, IgD, and CD79b is reduced and is thus consistent with a phenotype matching that of mature, activated B lymphocytes. The pathological features of biopsy specimens of lymph nodes are those of a small lymphocytic lymphoma.[23] The coexpression of CD5, a T cell–associated antigen, is a phenotypic characteristic and part of the disease defining criteria. CD5$^+$ B cells can be found in the peripheral blood of normal adults, suggesting that specific subsets of CD5$^+$ B cells from the mantle zone may be the normal counterparts of B-cell CLL. The low expression of the BCR is the hallmark of CLL cells and contributes to impairments in the activation of the cell following BCR stimulation.

BCR signaling in CLL has been extensively studied and has shed considerable light on the pathogenesis of the disease, thus aiding in the design of targeted therapies. The BCR is a multimeric complex containing the antigen-specific surface Ig and the two membrane-bound signal transduction elements CD79A and CD79B. Binding of antigen to the BCR induces activation of upstream kinases, which, in turn, activate other kinases by means of their cytoplasmic moieties, including SYK and the Src kinase. These kinases further activate the Bruton tyrosine kinase (BTK), PI3K, and other downstream pathways, including phospholipase C gamma 2 (PLC-γ2), calcium signaling, protein kinase C (PKC), nuclear factor (NF)-κB signaling, mitogen-activated protein kinases (MAPKs), and nuclear transcription.[26] The upstream kinases BTK and PI3K are currently targeted by specific agents, and other kinases and pathways are being explored for novel therapies.

The CD38 surface molecule supports B-cell interactions and differentiation. Under certain circumstances, CD38 also augments signaling of BCR, delivering signals that regulate the apoptosis of B cells. Expression of CD38 correlates with expression of unmutated V domains and suggests a bad prognosis. Another molecule that influences the BCR is the ζ-associated protein 70 (ZAP70). High levels of this receptor associated protein tyrosine kinase (usually found in T and natural killer [NK] cells, but not in normal B cells) are detected in the majority of unmutated CLLs and correlate with a poor prognosis.

The microenvironment may play a role in the pathogenesis of CLL. Interactions with stromal cells rescue CLL cells from apoptosis *in vitro.* Activated T cells support the growth of CLL cells; cytokines, such as IL-4 and vascular endothelial growth factor (VEGF), and chemokines, such as CXCL12, support the expansion of CLL clones. When compared with blood-derived cells, expression profiling of CLL cells in the patient's lymph nodes showed BCR and NF-κB activation promoting cell proliferation, thus indicating the strong effects of the tumor microenvironment.[27] Mesenchymal stromal cells are also commonly found in secondary lymphatic tissues of patients with CLL. There they provide survival and migration signals to CLL cells, as well as protection from apoptosis and modulation of antigen presentation.[26] Endothelial cell and follicular dendritic cells (FDCs) may also play a role.

Interaction with and evasion from the normal immune system has been shown to be of significance.[26] The number of circulating T cells, oligoclonal in both the CD4 and the CD8 compartments, is increased. The function of these T cells is impaired, and they express exhaustion markers, including programmed death protein 1 (PD-1), at higher levels. Accordingly, CLL cells express high levels of PD-1 ligand (PD-L1). In addition, NK cells have reduced effector activities[26] that are associated with low expression levels of the activating receptors NK-cell p30-related protein (NKp30) and NK group 2 member D (NKG2D). Together, these findings provide an explanation for ability of CLL cells to evade immune-mediated destruction.

CLL has been classically characterized by the accumulation of mature B cells that evade apoptosis, with high levels of the BCL2 antiapoptotic protein. Contradicting this dogma is the measurement of CLL kinetics that has shown CLL cells to proliferate at a high dynamic rate of up to 1% of the clone per day.[28] In addition, proliferation-related genes (*e.g., c-MYC* and *E2F1*) have been found to be upregulated, especially in the unmutated CLLs.[27] This finding suggests that CLL is not solely an accumulative disease but also has a proliferative element.

Clinical Features of CLL

CLINICAL PEARLS

Chronic Lymphocytic Leukemia (CLL): Clinical Manifestations

- Absolute blood lymphocytosis >5000 /mm^3 sustained over a period of 4 weeks (to exclude transient lymphocytosis related conditions as viral infections)
- At least 30% lymphocytes in a normo- or hypercellular marrow
- Phenotypically monoclonal lymphocytes
- They express mature B-cell markers (*e.g.,* CD19) and CD5
- Morphologically mature-appearing lymphocytes
- Most patients having some degree of lymphadenopathy on physical examination
- Progression and the need for treatment depend on lymphocyte doubling time, bulky symptomatic disease, anemia, thrombocytopenia, autoimmune phenomenon

The clinical diagnosis of CLL requires an absolute lymphocytosis with a threshold of $>5000 \times 10^9$/L mature-appearing lymphocytes in the blood smear, persistence of the lymphocytosis for >4 weeks, and a distinct immunophenotype (as described above). Approximately a quarter of patients with CLL are asymptomatic at diagnosis. The clinical characteristics at presentation include lymphadenopathy (87%), splenomegaly (54%), hepatomegaly (14%), high white blood cell (WBC) count, anemia, and thrombocytopenia (20%). Very high WBC counts are rare but may develop along the course of the disease. Hyperleukocytosis, which causes leukostasis and necessitates emergency treatment, is extremely rare. Other organs that are involved include other lymphoid tissues and rarely solid organs or skin.

A prognostic evaluation of the patient with CLL begins with study of the patient's blood and bone marrow. Lymph node biopsy is not necessary but, if performed, may reveal the diagnosis of small lymphocytic lymphoma (SLL), which is considered a different manifestation of the same disease. The differential diagnosis includes other low-grade lymphoproliferative disorders, such as a leukemic phase of lymphoma and mantle cell lymphoma (usually negative for CD23); hairy cell leukemia (CD5 and CD21 negative and CD103 and CD25 positive); T-cell leukemia (other T-cell markers, such as CD3, CD4, and CD7); and prolymphocytic leukemia (PLL), which is distinguished by morphologically immature-appearing cells and the presence of FMC7 and CD79b on the cell surface. T-cell CLL is rare (<5%).

TABLE 78.5 **Major Factors Associated With Prognosis of Chronic Lymphocytic Leukemia (CLL)**[30,4,5,6,7,9,31,33,34]

	Definition	Median Survival (Years)	Used in Clinical Practice
Rai stage 0	Leukocytosis	12.5	Yes
1	Leukocytosis and lymphadenopathy	8.4	
2	Lymphocytosis plus hepatosplenomegaly	6.9	
3	Lymphocytosis plus anemia (<11 gr/dL)	1.5	
4	Lymphocytosis plus thrombocytopenia (<100 000 × 10^9/L)	1.5	
Binet	Leukocytosis and lymphadenopathy	Age-matched	Yes
Stage A	Lymphadenopathy of more than two involved areas	7	
Stage B	Anemia or thrombocytopenia	2	
Stage C			
β_2 microglobulin	Normal	9.7	Yes
	Elevated	4.5	
CD 38	>30%	2.9 to <10 years	Yes
Early-stage CLL	<30%	9 to >26 years	
Chromosomal aberrations	17p-	2.7	Sometimes
(FISH analysis)	11q-	6.5	
	Trisomy 12	9.5	
	Normal karyotype	9.3	
	13q-	11.1	
Zeta-associated protein 70 (ZAP-70)	>20%	7.5	If available
Early-stage CLL	<20%	Not reached	
Mutational status	Unmutated	5.7 to <9.9 years	If available
	Mutated	10.2 to >24 years	

The oldest staging systems of CLL risk stratification rely on measurement of disease bulk as reported by Rai[29] and Binet[30] (Table 78.5). These staging systems are very useful for identifying patients who will need treatment at the time of diagnosis, but not for predicting who will subsequently need treatment. About 33% of patients never require treatment and have a long survival. In another third, an initial indolent phase is followed by disease progression. The remaining third of patients exhibit an aggressive disease at the onset and need immediate treatment.

The most powerful predictors for rapid and aggressive progression are the mutational status of the V domains and chromosomal abnormalities, as described above (Fig. 78.6; see Table 78.5). In contrast to genomic aberrations and serum markers, such as CD38, the mutational profile has the advantage of remaining constant during disease evolution. NGS has identified new genetic markers that have altered prognosis for patients with CLL.[31] *TP53, ATM, NOTCH1,* and *SF3B1* mutations indicate a worse prognosis.

Treatment

As CLL remains an incurable tumor, treatment may be delayed and the patient monitored until becoming symptomatic (see Table 78.4). Indications for therapeutic intervention include the development of symptoms, a worsening anemia and/or thrombocytopenia, autoimmune cytopenias, progressive splenomegaly, progressive lymphadenopathy, or a lymphocyte doubling time of ≤6 months. No prospective data yet exist to support the early treatment of asymptomatic patients with adverse prognostic features. However, this group warrants close monitoring.

Chlorambucil, alone or combined with corticosteroids, has been the most commonly used drug. It is advantageous in relieving symptoms, even in advanced disease. However, several randomized controlled trials have failed to demonstrate improved survival, and hardly any patient achieves complete remission (CR). Purine analogues (most commonly fludarabine) with or without cyclophosphamide have been shown to induce higher response rates, with some patients achieving complete remission. Bendamustine is a novel agent that contains both alkylating properties and a purine-like benzimidazole ring. It has been shown to be effective in CLL and less toxic than the fludarabine and cyclophosphamide combination.[32]

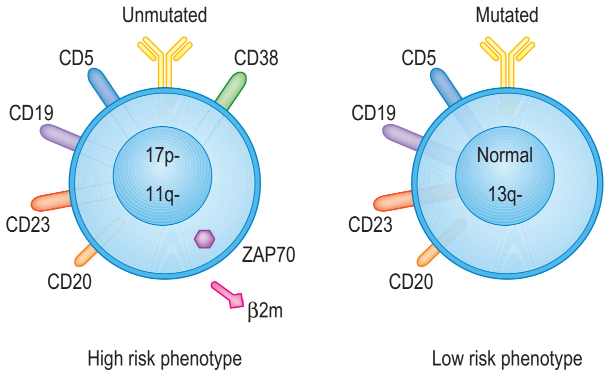


FIG 78.6 Risk Markers and Stratification in Chronic Lymphocytic Leukemia (CLL). CLL is one disease with two different entities. It commonly expresses normal levels of the B-cell surface antigens CD19 and CD23, low levels of CD20, and coexpression of the pan T-cell antigen CD5. The unmutated genotype correlates with a worse prognosis, featuring positive CD38 surface antigen, intracellular ZAP70, chromosomal aberrations as 17p- and 11q-, and high levels of soluble β_2 microglobulin. Conversely, the mutated form of CLL has a better prognosis, lacks these phenotypic features, and has fewer chromosomal aberrations, including 13q-.

THERAPEUTIC PRINCIPLES

Chronic Lymphocytic Leukemia (CLL) Is Incurable, but It Is Possible to Ameliorate Symptoms

- High rates of complete remissions and encouraging survival curves for high-risk patients
- Watch and wait approach in the case of asymptomatic patients with early-stage CLL
- Supportive care (immunoglobulins, blood supplements and erythropoietin, treatment of infections)
- Steroids with or without chemotherapy, especially for autoimmune phenomenon
- Conventional chemotherapy (alkylating agents, such as chlorambucil and cyclophosphamide; purine analogues, such as fludarabine and cladribine; other lymphoma regimens; bendamustine)
- Targeted therapy with or without chemotherapy (*i.e.*, rituximab, ofatumumab, obinutuzumab (all anti-CD20 antibodies) and alemtuzumab (anti-CD52 antibody)
- B-cell receptor pathway inhibitors—BTK (ibrutinib), PI3K-d (idelalisib), and BCL2 antagonists (venetoclax)
- Allogeneic (mostly with reduced intensity) stem cell transplantation

The CD52 antigen is present on lymphocytes (B, T, and NK cells), monocytes, and some granulocytes. The humanized anti-CD52 monoclonal antibody (mAb; CAMPATH-1, alemtuzumab) has demonstrated high activity in previously treated patients and those with refractory CLL.[33] Nevertheless, because of significant B- and T-cell depletion, patients are exposed to severe infectious complications, particularly reactivation CMV infection.

The anti-CD20 mAb rituximab (Mabthera) as a single agent has shown limited efficacy in CLL, possibly because of the weak receptor expression on CLL cells. However, in combination with chemotherapy, particularly with fludarabine and cyclophosphamide or bendamustine,[32] it appears to act synergistically and to achieve high rates of response, including molecular complete remission and prolongation of disease-free and overall survival. A newer fully human mAb targeting CD20 (ofatumumab) has a very potent effect as a single agent in both rituximab-naïve and rituximab-treated patients.[34] It targets a different epitope on the CD20 antigen. A second-generation anti-CD20 antibody, obinutuzumab, was added to chlorambucil in the treatment of the older adult population. When compared with chlorambucil, alone or in combination with rituximab and chlorambucil, obinutuzumab was shown to be superior and relatively safer.[32] Other investigational agents showing efficacy include lenalidomide (an immune modulator related to thalidomide), albeit with an interestingly initial flair-up phenomenon of disease-enlarged lymph nodes.

A major breakthrough in treatment was achieved through drugs targeting the BCR pathway. A BTK inhibitor, ibrutinib, has shown an impressive progression-free survival in relapsing patients and as first-line treatment[35,36] for patients with resistant or high-risk 17p-deleted CLL. In combination with rituximab, another pathway inhibitor, the PI3Kδ inhibitor idelalisib had an efficacy superior to rituximab alone in frail, older patients who were relapsing.[37] Furthermore, clinical trials with small molecules targeting BCL2, venetoclax, have shown promising results, even in patients with relapsing and highly resistant CLL.[38] Under investigation are various novel agents targeting other surface molecules (CD23, CD79b, and other CD20 inhibitors), pathway inhibitors (SYK, PI3K, second-generation BTK, and others), and immune checkpoint inhibitors. Some immunological novel approaches include expanded autoreactive activated T cells (chimeric antigen receptor–modified T cells [CAR-T]), which is showing promise.[39]

Allogeneic hematopoietic stem cell transplantation (allo-HSCT) is the only curative treatment for CLL. Allogeneic HSCT relies on myeloablative doses of chemoradiotherapy, which makes the treatment unacceptably risky for the majority of patients with CLL. In nonmyeloablative, or reduced-intensity, approaches, rates of engraftment are similar to fully ablative conditioning regimens, but with lower rates of early toxicity. Early evidence suggests that the graft-versus-leukemia (GVL) effect is present. Patients with deletion 17p (p53 involvement) with an extremely poor prognosis are physically fit candidates for allo-HSCT,[32] even as newer agents (*i.e.*, ibrutinib, idelalisib, and venetoclax) are showing high efficacy in this high-risk population. Studies involving autologous transplantation and high-dose chemotherapy for CLL have a limited survival advantage.

With treatment advances, the overall survival of patients with CLL has improved across the globe as a result of targeted therapy, even in high-risk and in patients with relapsed and refractory CLL. Thus classic prognostication (as presented in Table 78.5) is becoming less accurate as treatment improves.

Immunological Aspects of CLL

CLINICAL PEARLS

Chronic Lymphocytic Leukemia (CLL): Immunologic Manifestations

- Panhypogammaglobulinemia
- A monoclonal immunoglobulin peak, usually of the immunoglobulin M type
- Downregulation of T-cell function and aberrant cytokine production
- Defects in the complement system
- High risk of recurrent infections—encapsulated bacteria and opportunistic infections
- Autoimmune-associated phenomena:
  - Autoimmune hemolytic anemia
  - Autoimmune thrombocytopenia
  - Pure red cell aplasia and autoimmune neutropenia
  - Other autoimmune disorders (myositis, vasculitis, pemphigus vulgaris, acquired angioedema, glomerulonephritis)

CLL is characterized by multiple immune deficiencies and autoimmune phenomena. It is reasonable to hypothesize that immune incompetence and autoimmunity are two sides of the same coin.

The Pathophysiological Rationale

CLL cells secrete TGF-β, which is a potent inhibitor of B-cell proliferation. They also release high levels of circulating IL-2 receptor, which downregulates the T-helper (Th) cell function. Unlike normal cells, activated B cells, both B-cell CLL cells and anergic normal B cells, fail to present soluble antigen and alloantigens. Moreover, the T cells in patients with CLL often demonstrate profound abnormalities of their antigen receptor (TCR) repertoire and appear dysfunctional in terms of cytokine secretion.[40] This cytokine imbalance may be the cause of upregulation of the BCL-2 antiapoptotic activity. T-cell dysfunction could also explain the higher incidence of autoimmune complications, such as autoimmune hemolytic anemia, among patients receiving purine analogues therapy, which induce T-cell depletion. Certain

T-cell subsets appear to prevent the development of autoreactive B cells. When these are absent (*e.g.,* after treatment with purine analogues), autoreactive B-cell clones may easily emerge and expand.

Immunological Deficiencies

Patients with CLL are extremely sensitive to a number of infectious agents. A monoclonal Ig peak, usually of the IgM type, is found in 5% of patients with CLL, and a small amount of a monoclonal component can be identified in the serum or urine of 60% of patients. Hypogammaglobulinemia occurs in at least 60% of B-cell CLL cases and may include all three classes (IgG, IgA, and IgM). The pathogenesis of hypogammaglobulinemia in B-cell CLL is poorly understood, as this phenomenon is rare in other B-cell malignancies except multiple myeloma. Low Ig levels correlate with recurrent infections of encapsulated organisms. In patients who receive intravenous immune globulin (IVIG), there is a decrease in the incidence of major bacterial infections.

Infections are a major cause of morbidity and mortality in patients with CLL. Impaired humoral and cellular immunities, defects in the complement systems, and variable neutropenia, depending on marrow infiltrates, all contribute to the high rate of infections. Opportunistic infections are initially uncommon as the result of the relative preservation of cellular immunity early in the disease. Infection risk increases following purine analogue therapy because of the side effects of myelosuppression and marked lymphopenia with T-cell depletion. The addition of rituximab, the anti–B cell marker CD20 antibody, to nucleoside analogue-based therapy does not appear to increase the risk of early or late infections but may increase the rate of neutropenia. Active immunization with vaccines is hampered by the patient's inability to generate or retain a long and significant immune response.

Autoimmune Phenomena

Autoimmune-associated features are common in CLL. These manifestations primarily affect hematopoietic cells. For example, the most common known cause of autoimmune hemolytic anemia (AIHA) is CLL.[41] Positive result of direct antiglobulin test (direct Coomb test) has been reported to be as high as in 7–35% of patients with CLL, and AIHA itself occurs in 10–25% of patients during the course of their disease, twice as often in patients with unmutated genes as in those with mutated ones. Autoantibodies against red blood cells (RBCs) are warm-reactive polyclonal IgG. They are not secreted by the malignant clone, but rather by normal B cells.[41] Cold agglutinins are rare. AIHA is thought to arise from the imbalance among lymphocyte subsets, contributed to by therapy, resulting in the emergence of the autoimmune clone. It is usually observed in advanced stages of the disease, correlates with a poor prognosis, and has a close relationship with the activity of the CLL. After therapy, the autoimmune antibodies may remit in 70% of the treated patients.

Idiopathic thrombocytopenic purpura (ITP) is observed in about 2–3% of cases and presents as increased megakaryocytes in bone marrow. It should be distinguished from immune thrombocytopenia induced by marrow infiltration, which is very common in up to 50% of patients at presentation.[41] Two-thirds of patients with CLL-associated ITP also have AIHA (Evan syndrome). Pure red cell aplasia (PRCA) and autoantibodies against neutrophils are only rarely observed but are part of the CLL-related autoimmunity repertoire. Interestingly, the addition of cyclophosphamide represses the emergence of fludarabine-induced AIHA, but the latter may also be seen with other chemotherapies (*i.e.,* bendamustine).[42] Autoimmune phenomena in patients treated with purine analogues (mostly fludarabine-related) are of a more severe nature.

Other rare entities are reported as paraneoplastic autoimmune disorders with connective tissue disease manifestations, such as polymyositis, dermatopolymyositis, and focal myositis or as vasculitis, pemphigus vulgaris, and acquired angioedema. These autoimmune disorders are related to T-cell dysfunction and may be associated with purine analogue treatment. Paraneoplastic pemphigus also occurs in patients with CLL and may be triggered by chemotherapy and radiotherapy. Glomerulonephritis and nephrotic syndrome are seldom reported but, when present, are related to different mechanisms, such as cryoglobulins and antineutrophil cytoplasmic antibodies (ANCAs).

Therapy of autoimmune phenomena includes high-dose steroids and disease control.[41] In patients refractory to or relapsing after steroid therapy, more aggressive treatment is warranted. High-dose Igs offers transient amelioration in some patients. Splenectomy or splenic irradiation, cytotoxic agents, or cyclosporine may represent valid rescue approaches. In cases where AIHA has been triggered by fludarabine, further exposure is hazardous. Rituximab may be an alternative agent for the treatment CLL-associated autoimmune diseases, including rare autoimmune phenomena, such as pemphigus and PRCA.

Other Malignancies

Second malignancies (hematological and solid tumors) are not uncommon in CLL. The most common hematological malignancy is the Richter transformation to diffuse large B-cell lymphoma, which occurs in ≈5% of patients, as well as other high-grade lymphoproliferative diseases. Dermatological tumors, such as basal cell carcinoma, are the most frequent of the solid tumors encountered in patients with CLL, and these malignancies are more likely to be locally aggressive and metastatic.[32] The pathogenesis of these second cancers is not fully understood, and although disease-related genetic factors (*i.e.,* 17p deletion, notch mutation) are a major determinant, it is probably multifactorial and includes Epstein-Barr virus (EBV) infection and BCR configuration to respond to multiple autoantigens and immune/inflammatory stimuli present in the microenvironment.[32,43]

CONCLUSIONS

CLL is a common indolent lymphoid neoplasm with a wide clinical heterogeneity. It is suspected and diagnosed more commonly because of routine blood tests. Diagnosis is made with simple immunophenotyping. Cytogenetics and molecular diagnostic techniques are needed to determine the prognosis. The complications of CLL appear to be unique to this neoplasm and are part of a failing immune system with T-cell and B-cell dysregulation causing both deficiencies predisposing patients to recurrent infections and autoimmune diseases. New molecular and protein markers are key to finding novel effective targeted therapies.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Pui CH, Yang JJ, Hunger SP, et al. Childhood Acute Lymphoblastic Leukemia: Progress Through Collaboration. J Clin Oncol 2015;33:2938–48.
2. Greaves M. Infection, immune responses and the aetiology of childhood leukaemia. Nat Rev Cancer 2006;6:193–203.
3. Kuehn HS, Boisson B, Cunningham-Rundles C, et al. Loss of B Cells in Patients with Heterozygous Mutations in IKAROS. N Engl J Med 2016;374:1032–43.
4. Noetzli L, Lo RW, Lee-Sherick AB, et al. Germline mutations in ETV6 are associated with thrombocytopenia, red cell macrocytosis and predisposition to lymphoblastic leukemia. Nat Genet 2015;47:535–8.
5. Linden T, Schnittger S, Groll AH, et al. Childhood B-cell precursor acute lymphoblastic leukaemia in a patient with familial thrombocytopenia and RUNX1 mutation. Br J Haematol 2010;151:528–30.
6. Hyde RK, Liu PP. Germline PAX5 mutations and B cell leukemia. Nat Genet 2013;45:1104–5.
7. Martin-Lorenzo A, Hauer J, Vicente-Duenas C, et al. Infection Exposure is a Causal Factor in B-cell Precursor Acute Lymphoblastic Leukemia as a Result of Pax5-Inherited Susceptibility. Cancer Discov 2015;5:1328–43.
8. Coustan-Smith E, Mullighan CG, Onciu M, et al. Early T-cell precursor leukaemia: a subtype of very high-risk acute lymphoblastic leukaemia. Lancet Oncol 2009;10:147–56.
9. Arber DA, Orazi A, Hasserjian R, et al. The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. Blood 2016;127:2391–405.
10. Tal N, Shochat C, Geron I, et al. Interleukin 7 and thymic stromal lymphopoietin: from immunity to leukemia. Cell Mol Life Sci 2014;71:365–78.
11. Buchner M, Swaminathan S, Chen Z, et al. Mechanisms of pre-B-cell receptor checkpoint control and its oncogenic subversion in acute lymphoblastic leukemia. Immunol Rev 2015;263:192–209.
12. Fischer U, Forster M, Rinaldi A, et al. Genomics and drug profiling of fatal TCF3-HLF-positive acute lymphoblastic leukemia identifies recurrent mutation patterns and therapeutic options. Nat Genet 2015;47:1020–9.
13. Roberts KG, Li Y, Payne-Turner D, et al. Targetable kinase-activating lesions in Ph-like acute lymphoblastic leukemia. N Engl J Med 2014;371:1005–15.
14. Izraeli S. Beyond Philadelphia: 'Ph-like' B cell precursor acute lymphoblastic leukemias - diagnostic challenges and therapeutic promises. Curr Opin Hematol 2014;21:289–96.
15. Harrison CJ, Haas O, Harbott J, et al. Detection of prognostically relevant genetic abnormalities in childhood B-cell precursor acute lymphoblastic leukaemia: recommendations from the Biology and Diagnosis Committee of the International Berlin-Frankfurt-Munster study group. Br J Haematol 2010;151:132–42.
16. Szczepanski T, Orfao A, van der Velden VH, et al. Minimal residual disease in leukaemia patients. Lancet Oncol 2001;2:409–17.
17. Conter V, Bartram CR, Valsecchi MG, et al. Molecular response to treatment redefines all prognostic factors in children and adolescents with B-cell precursor acute lymphoblastic leukemia: results in 3184 patients of the AIEOP-BFM ALL 2000 study. Blood 2010;115:3206–14.
18. Koch SV, Kejs AM, Engholm G, et al. Leaving home after cancer in childhood: a measure of social independence in early adulthood. Pediatr Blood Cancer 2006;47:61–70.
19. Anderson K, Lutz C, van Delft FW, et al. Genetic variegation of clonal architecture and propagating cells in leukaemia. Nature 2011;469:356–61.
20. Klinger M, Benjamin J, Kischel R, et al. Harnessing T cells to fight cancer with BiTE(R) antibody constructs–past developments and future directions. Immunol Rev 2016;270:193–208.
21. Lee DW, Kochenderfer JN, Stetler-Stevenson M, et al. T cells expressing CD19 chimeric antigen receptors for acute lymphoblastic leukaemia in children and young adults: a phase 1 dose-escalation trial. Lancet 2015;385:517–28.
22. Zent CS, Kyasa MJ, Evans R, et al. Chronic lymphocytic leukemia incidence is substantially higher than estimated from tumor registry data. Cancer 2001;92:1325–30.
23. Chiorazzi N, Rai KR, Ferrarini M. Chronic lymphocytic leukemia. N Engl J Med 2005;352:804–15.
24. Strati P, Shanafelt TD. Monoclonal B-cell lymphocytosis and early-stage chronic lymphocytic leukemia: diagnosis, natural history, and risk stratification. Blood 2015;126:454–62.
25. Goldin LR, Bjorkholm M, Kristinsson SY, et al. Elevated risk of chronic lymphocytic leukemia and other indolent non-Hodgkin's lymphomas among relatives of patients with chronic lymphocytic leukemia. Haematologica 2009;94:647–53.
26. Ten Hacken E, Burger JA. Microenvironment interactions and B-cell receptor signaling in Chronic Lymphocytic Leukemia: Implications for disease pathogenesis and treatment. Biochim Biophys Acta 2015;1863:401–13.
27. Herishanu Y, Perez-Galan P, Liu D, et al. The lymph node microenvironment promotes B-cell receptor signaling, NF-kappaB activation, and tumor proliferation in chronic lymphocytic leukemia. Blood 2011;117:563–74.
28. Messmer BT, Messmer D, Allen SL, et al. In vivo measurements document the dynamic cellular kinetics of chronic lymphocytic leukemia B cells. J Clin Invest 2005;115:755–64.
29. Rai KR, Sawitsky A, Cronkite EP, et al. Clinical staging of chronic lymphocytic leukemia. Blood 1975;46:219–34.
30. Binet JL, Lepoprier M, Dighiero G, et al. A clinical staging system for chronic lymphocytic leukemia: prognostic significance. Cancer 1977;40:855–64.
31. Nabhan C, Raca G, Wang YL. Predicting Prognosis in Chronic Lymphocytic Leukemia in the Contemporary Era. JAMA Oncol 2015;1:965–74.
32. Stilgenbauer S, Furman RR, Zent CS. Management of chronic lymphocytic leukemia. Am Soc Clin Oncol Educ Book 2015;164–75.
33. Dighiero G. Unsolved issues in CLL biology and management. Leukemia 2003;17:2385–91.
34. Abou-Nassar K, Brown JR. Novel agents for the treatment of chronic lymphocytic leukemia. Clin Adv Hematol Oncol 2010;8:886–95.
35. Byrd JC, Brown JR, O'Brien S, et al. Ibrutinib versus ofatumumab in previously treated chronic lymphoid leukemia. N Engl J Med 2014;371:213–23.
36. Burger JA, Tedeschi A, Barr PM, et al. Ibrutinib as Initial Therapy for Patients with Chronic Lymphocytic Leukemia. N Engl J Med 2015;373:2425–37.
37. Furman RR, Sharman JP, Coutre SE, et al. Idelalisib and rituximab in relapsed chronic lymphocytic leukemia. N Engl J Med 2014;370:997–1007.
38. Roberts AW, Davids MS, Pagel JM, et al. Targeting BCL2 with Venetoclax in Relapsed Chronic Lymphocytic Leukemia. N Engl J Med 2016;374:311–22.
39. Porter DL, Levine BL, Kalos M, et al. Chimeric antigen receptor-modified T cells in chronic lymphoid leukemia. N Engl J Med 2011;365:725–33.
40. Scrivener S, Goddard RV, Kaminski ER, et al. Abnormal T-cell function in B-cell chronic lymphocytic leukaemia. Leuk Lymphoma 2003; 44:383–9.
41. Hamblin TJ. Autoimmune complications of chronic lymphocytic leukemia. Semin Oncol 2006;33:230–9.
42. Goldschmidt N, Gural A, Ben-Yehuda D, et al. Short communication: bendamustine-related hemolytic anemia in chronic lymphocytic leukemia. Cancer Chemother Pharmacol 2013;72:709–13.
43. Rossi D, Gaidano G. Richter syndrome: pathogenesis and management. Semin Oncol 2016;43:311–19.

MULTIPLE-CHOICE QUESTIONS

1. Which of the following cell surface antigens is not consistent with the diagnosis of acute lymphoblastic leukemia (ALL)?
   A. CD3
   B. CD19
   C. Surface Immunoglobulin M (IgM)
   D. CD22
   E. Intracytoplasmic IgM

2. The detection of minimal residual disease of ALL is based on:
   A. Bone marrow magnetic resonance imaging
   B. Positron emission tomography with computed tomography
   C. Determination of secreted immunoglobulins
   D. Determination of clonal V(D)J rearrangements
   E. Serial bone marrow biopsies

3. CRLF2:
   A. Is a signaling cytoplasmic molecule
   B. Is part of the receptor to thymic stromal lymphopoietin
   C. Is a major driver of T-acute lymphoblastic leukemia
   D. Modulate the antileukemic effects of mast cells
   E. Inhibits the growth of leukemic cells

79

Lymphomas

Stefania Pittaluga, Tapan Bhavsar, Elaine S. Jaffe

The classification of malignant lymphomas continues to undergo revision based on insights gained through the application of immunological and molecular techniques and the application of these discoveries to individualized therapeutic approaches. Early classifications were based on morphological characteristics of the neoplastic elements; however, with increasing knowledge of the complexity of the immune system, a more functional approach was sought. Differentiation schemes provided a useful starting point for understanding lymphomas (Fig. 79.1). High-throughput genomic studies of both tumor RNA and tumor DNA have been applied to lymphomas to define their molecular signatures with the aim of improving the understanding of oncogenic pathways and their clinical implications. These studies have led to new prognostic and diagnostic tools, and as a result, more targeted therapies have emerged.[1]

KEY CONCEPTS

Lymphoma

- Classification consists of a list of individual disease entities defined by morphological, immunophenotypic, genetic, and clinical features.
- Neoplastic cells are related to the postulated normal counterpart, when possible.
- Histological grade should be applied within individual diseases.
- Clinical factors for individual patients, as measured by the International Prognostic Index (IPI) and gene expression profiling, are useful in predicting clinical outcome.

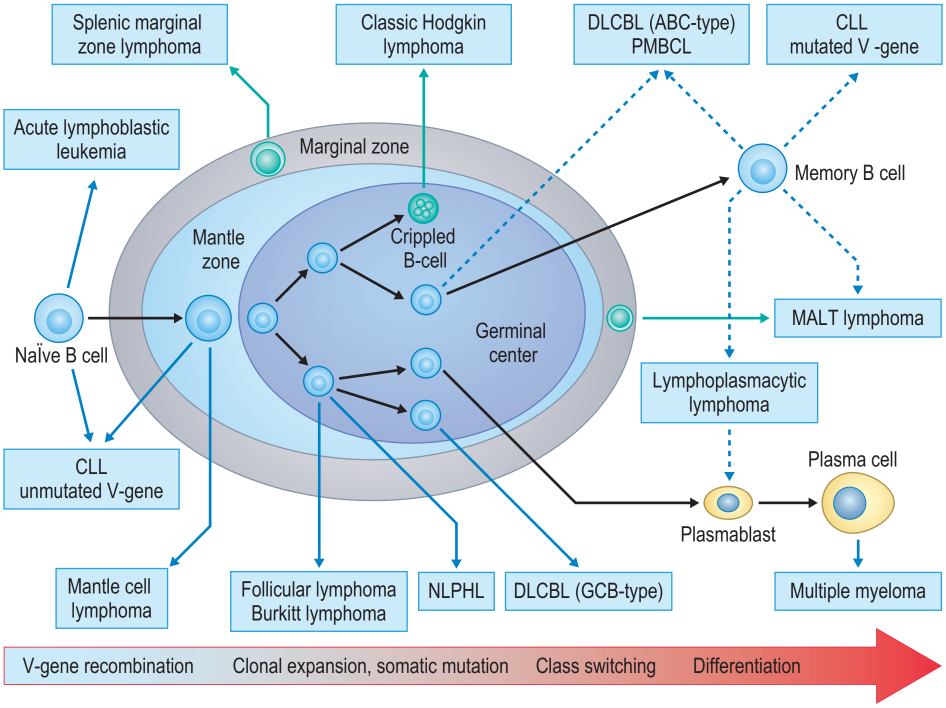

Splenic marginal zone lymphoma
Classic Hodgkin lymphoma
DLCBL (ABC-type) PMBCL
CLL mutated V -gene
Acute lymphoblastic leukemia
Marginal zone
Memory B cell
Mantle zone
Crippled B-cell
MALT lymphoma
Naïve B cell
Germinal center
Lymphoplasmacytic lymphoma
CLL unmutated V-gene
Plasma cell
Plasmablast
Mantle cell lymphoma
Follicular lymphoma Burkitt lymphoma
NLPHL
DLCBL (GCB-type)
Multiple myeloma
V-gene recombination
Clonal expansion, somatic mutation
Class switching
Differentiation

FIG 79.1 **Normal B-Cell Differentiation in Relation to a Secondary B Follicle, Mutational Stages of the Immunoglobulin Genes, and Cellular Counterparts for B-Cell Lymphomas.** Simplified version of B-cell development indicates points at which V-gene recombination, clonal expansion, and somatic mutations occur in relation to a secondary B follicle. B-cell lymphomas are related to different stages of B-cell differentiation and function. ABC-type, activated B-cell-like type; CLL, chronic lymphocytic leukemia/lymphoma; DLBCL, diffuse large B-cell lymphoma; GCB-type, germinal-center B-cell-like type; MALT lymphoma, marginal zone B-cell lymphoma of mucosa-associated lymphoid tissue (MALT) type; NLPHL, nodular lymphocyte-predominant Hodgkin lymphoma; PMBCL, primary mediastinal B-cell lymphoma.

The guiding principles of the World Health Organization (WHO) classification of neoplasms of the hematopoietic and lymphoid tissues, published in 2001 and updated in 2008 and 2016,[2] were based on the Revised European American Classification of Lymphoid Neoplasms (REAL), which was published by the International Lymphoma Study group in 1994. The focus was on the identification of individual diseases based on an integration of morphological, immunophenotypic, genetic, and clinical features. The recent application of gene expression profiling (GEP) in lymphomas has generated distinct molecular "signatures" for a variety of disease entities, in some cases corresponding more closely to different stages of lymphoid differentiation and in other instances offering insights into mechanisms of neoplastic transformation with delineation of alterations in specific pathways. The advent of whole-genome sequencing contributes additional new insights into pathogenetic mechanisms. Recently, a revision of the nearly 8-year-old WHO classification of lymphoid neoplasms (Table 79.1) was published clarifying

TABLE 79.1 WHO Classification of Tumors of Hematopoietic and Lymphoid Tissues (2016)

B-Cell Neoplasm

- Precursor B-cell lymphoblastic leukemia/lymphoma
    - B-lymphoblastic leukemia/lymphoma, not otherwise specified
    - B-lymphoblastic leukemia/lymphoma with recurrent genetic abnormalities

Mature B-Cell Neoplasm

- Chronic lymphocytic leukemia/small lymphocytic lymphoma
- Monoclonal B-cell lymphocytosis*
- B-cell prolymphocytic leukemia
- Splenic B-cell marginal zone lymphoma
- Hairy cell leukemia
- _Splenic B-cell lymphoma/leukemia, unclassifiable_
    - _Splenic diffuse red pulp small B-cell lymphoma_
    - _Hairy cell leukemia-variant_
- Lymphoplasmacytic lymphoma
- Waldenstrom macroglobulinemia
- Monoclonal gammopathy of undetermined significance (MGUS), IgM*
- μ heavy-chain disease
- γ heavy-chain disease
- α heavy-chain disease
- Monoclonal gammopathy of undetermined significance (MGUS), IgG/A*
- Plasma cell myeloma
- Solitary plasmacytoma of bone
- Extraosseous plasmacytoma
- Monoclonal immunoglobulin deposition diseases*
- Extranodal marginal zone lymphoma of mucosa-associated lymphoreticular tissue (MALT) lymphoma
- Nodal marginal zone lymphoma
    - Pediatric nodal marginal zone lymphoma
- Follicular lymphoma
    - _In situ_ follicular neoplasia*
    - Duodenal-type follicular lymphoma*
    - Pediatric-type follicular lymphoma*
    - _Large B-cell lymphoma with IRF4 rearrangements_*
    - Primary cutaneous follicle-center lymphoma
- Mantle cell lymphoma
- _In situ mantle cell neoplasia_ *
- Diffuse large B-cell lymphoma (DLBCL), NOS
    - Germinal-center B-cell type*
    - Activated B-cell type*
- T-cell/histiocyte rich large B-cell lymphoma
- Primary DLBCL of the central nervous system
- Primary cutaneous DLBCL, leg type
- EBV$^+$ DLBCL, NOS*
- _EBV$^+$ mucocutaneous ulcer_*
- DLBCL associated with chronic inflammation
- Lymphomatoid granulomatosis
- Primary mediastinal (thymic) large B-cell lymphoma
- Intravascular large B-cell lymphoma
- ALK positive large B-cell lymphoma
- Plasmablastic lymphoma
- Large B-cell lymphoma arising in HHV8-associated multicentric Castleman disease
- Primary effusion lymphoma
- _HHV8 positive DLBCL, NOS_*
- Burkitt lymphoma
- _Burkitt-like lymphoma with 11q aberration_*
- High grade B-cell lymphoma, with MYC and BCL2 and/or BCL6 rearrangements*
- High grade B-cell lymphoma, NOS*
- B-cell lymphoma unclassifiable, with features intermediate between diffuse large B-cell lymphoma and classic Hodgkin lymphoma

T-Cell Neoplasm

- Precursor T-cell lymphoblastic leukemia/lymphoma

Mature T-Cell and NK-Cell Neoplasms

- T-cell prolymphocytic leukemia
- T-cell large granular lymphocytic leukemia
- _Chronic lymphoproliferative disorder of NK-cells_
- Aggressive NK leukemia
- Systemic EBV$^+$ T-cell lymphoma of childhood*
- Hydroa vacciniforme-like lymphoproliferative disorder*
- Adult T-cell leukemia/lymphoma
- Extranodal NK-/T-cell lymphoma, nasal type
- Enteropathy-associated T-cell lymphoma
- Monomorphic epitheliotropic intestinal T-cell lymphoma*
- _Indolent T-cell lymphoproliferative disorder of the GI tract_*
- Hepatosplenic T-cell lymphoma
- Subcutaneous panniculitis-like T-cell lymphoma
- Mycosis fungoides
- Sézary syndrome
- Primary cutaneous CD30$^+$ positive T-cell lymphoproliferative disorders
    - Primary cutaneous anaplastic large-cell lymphoma
    - Lymphomatoid papulosis
- Primary cutaneous γ/δ T-cell lymphoma
- _Primary cutaneous CD8 positive aggressive epidermotropic cytotoxic T-cell lymphoma_
- _Primary cutaneous acral CD8$^+$ T-cell lymphoma_*
- _Primary cutaneous CD4 positive small/medium T-cell lymphoproliferative disorder_*
- Peripheral T-cell lymphoma, NOS
- Angioimmunoblastic T-cell lymphoma
- _Follicular T-cell lymphoma_*
- _Nodal peripheral T-cell lymphoma with TFH phenotype_*
- Anaplastic large-cell lymphoma, ALK positive
- Anaplastic large-cell lymphoma, ALK negative*
- _Breast implant-associated anaplastic large-cell lymphoma_*

Hodgkin Lymphoma

- Nodular lymphocyte predominant Hodgkin lymphoma
- Classic Hodgkin lymphoma
- Nodular sclerosis Hodgkin lymphoma
- Lymphocyte-rich classic Hodgkin lymphoma
- Mixed cellularity classic Hodgkin lymphoma
- Lymphocyte-depleted classic Hodgkin lymphoma

*Refers to changes from the 2008 classification.

Note: More common entities are underlined. Provisional entities in italics.

From Swerdlow SH, Campo E, Pileri SA, et al. The 2016 revision of the World Health Organization (WHO) classification of lymphoid neoplasm. Blood 2016;127:2375–90.

the diagnosis and management of lesions at the very early stages of lymphomagenesis, refining diagnostic criteria for some entities, and detailing the expanding genetic/molecular landscape of numerous lymphoid neoplasms and their clinical correlates.[3]

This chapter focuses on the classification of neoplasms derived from mature B cells, T cells, and natural killer (NK) cells, with emphasis on malignant lymphomas. Other chapters in this volume cover lymphoid leukemias (Chapter 78) and immunosecretory disorders, including plasma cell neoplasms (Chapter 80). Special attention is devoted to the impact of clinical features (*e.g.*, age and anatomical site) on disease definition and a greater appreciation of early events in neoplastic transformation. These early lesions can sometimes be detected in otherwise healthy individuals, and these lesions may or may not progress to overt lymphoma or leukemia. These entities appear to carry fewer genetic aberrations compared with the conventional forms of the disease, which perhaps explains their indolent clinical behavior.[4]

CLINICAL PEARLS

Indolent Lymphomas

- Natural history: survival measured in years
- Least sensitive to therapy
- Good response to low-dose oral alkylating agents, radiotherapy, and steroids, but no curability
- Higher response rate and complete remission with combination of standard chemotherapy and anti-CD20 monoclonal antibody
- Gene expression profiling can help identify patients who might benefit from high-dose chemotherapy and autologous stem-cell transplantation, which is a potentially curative modality

MATURE B-CELL NEOPLASMS

KEY CONCEPTS

Somatic Mutation in Relation to Normal B-Cell Development

- *Premutational stage:* circulating naïve B cells (immunoglobulin (Ig)M$^+$/D$^+$) before antigen exposure
- *Stage of somatic mutation, clonal expansion, and isotype switch:* at the germinal center
- *Postmutational stage:* selected B cells move to the periphery (post–germinal center) or to the recirculating pool (memory B cells), or undergo terminal differentiation (plasma cells)

Lymphoplasmacytic Lymphoma

Lymphoplasmacytic lymphoma (LPL) is a disease of adult life (median age in the 60s), usually presenting with generalized lymphadenopathy, constitutional symptoms, and splenomegaly.

Histologically, there is a diffuse interfollicular proliferation of small lymphocytes (many with plasmacytoid features) and plasma cells, with or without immunoglobulin (Ig)–filled intranuclear inclusions (Dutcher bodies) and sparing of the sinuses. An increased number of mast cells and iron-laden macrophages can be seen. Although many B-cell neoplasms occasionally show maturation to plasmacytoid or plasma cells with cytoplasmic Ig, the term LPL should be restricted to tumors lacking features of other well-defined entities, such as chronic lymphocytic leukemia (CLL) or mantle-cell lymphoma (MCL), which occasionally can manifest plasmacytoid differentiation. Many but not all patients with LPL may have clinical evidence of Waldenström macroglobulinemia (WM), which is based on the detection of an IgM monoclonal gammopathy of any concentration and is associated with bone marrow involvement by LPL (Chapter 80).

In LPL, cells have surface and cytoplasmic Ig, usually IgM (and usually lacking IgD), and express B cell–associated antigens (CD19, CD20, CD22, CD79a). They are generally negative for CD5 and cyclinD1, distinguishing LPL from CLL and MCL, respectively. CD25, CD10, or CD11c may be weakly expressed in some cases. The postulated normal counterpart is thought to be a postfollicular medullary cord B cell, based, in part, on the presence of somatic mutations in the Ig heavy-chain and light-chain variable region genes.

The recent identification of *MYD88* L265P mutation (found in approximately 90% of WM cases) has become a reliable marker supporting a diagnosis of LPL.[5] This mutation is also found in a significant proportion of IgM, but not IgG or IgA, monoclonal gammopathy of undermined significance (MGUS) cases.

Mantle-Cell Lymphoma

MCL usually presents in adults (median age 62; variable male predominance) with advanced-stage disease—involving lymph nodes, Waldeyer ring lymphoid tissue, spleen, bone marrow, and peripheral blood. Gastrointestinal (GI) tract involvement is common and is associated with a lymphomatous polyposis picture. Retrospective studies have shown a poor prognosis (median survival 3–5 years), with a high relapse rate following initial remission. MCL is composed of small lymphoid cells with slightly irregular nuclear contours, finely clumped chromatin, and scant cytoplasm. Specifically, blastoid and pleomorphic variants have been associated with a more aggressive clinical course. The postulated normal counterpart is a CD5$^+$ "naïve" sIgM$^+$ and sIgD$^+$ B cell, found in peripheral blood and in the mantles of reactive follicles.

MCL is characterized by a molecular hallmark, t(11;14)(q13;q32)—involving cyclin D1 (*CCND1*) and the *IGH* genes; an overexpression of *CCND1* is believed to be essential in the pathogenesis. Rare variants negative for *CCND1* but with similar immunomorphology and GEP have also been identified.[6] Half of the *CCND1* expression/rearrangement-negative forms have *CCND2* translocations, often with IGK or *IGL* as a partner locus, which could be of diagnostic utility.[7] SOX11 is overexpressed in most *CCND1*-positive and *CCND1*-negative cases.[8] Besides the *CCND1/IGH* translocation, additional alterations involving other cell cycle regulatory proteins (RB, p53, CDK inhibitors) have been described in the more aggressive forms of MCL.

In-situ mantle cell neoplasia (ISMCN) (according to new 2016 classification; previously "*in-situ*" MCL) describes clonal proliferation of cyclin D1-positive cells restricted to the inner mantle cuffs in an otherwise reactive lymph node/lymphoid tissue and usually represents an incidental finding. Some cases will eventually progress to overt MCL[2]; however, the risk of progression is difficult to ascertain, as the number of reported cases is few. Another newly identified nonnodal variant is characterized by a leukemic phase without nodal disease but often long-standing splenomegaly. These cases develop from IgHV-mutated SOX11-negative B cells, carry t(11;14) with few additional chromosomal abnormalities but lack expression of SOX11.[9]

The proliferation rate based on Ki67-positivity has been considered of prognostic relevance. More recently, GEP using genes involved in cell cycle progression and DNA synthesis has identified a proliferation signature—defining different prognostic groups and showing correlation with cytological subtype to some extent. For example, a high proliferation rate has been shown in the blastoid variant.

The treatment approach for newly diagnosed patients with MCL depends on their eligibility for stem cell transplantation (SCT). However, purging of the neoplastic cells is difficult to achieve because of frequent bone marrow involvement, and autologous SCT has not proven be curative. The incorporation of rituximab into chemotherapeutic regimens has become an evidence-based standard of care. Of the current drugs used for MCL, the application of the BTK inhibitor ibrutinib might change treatment paradigms by obviating the need for transplantation in younger patients and chemotherapy in older patients.[10] The patients with the nonnodal variant as described above do not appear to require aggressive chemotherapy.[9]

Follicular Lymphoma

Follicular lymphoma (FL) is the most common subtype of non-Hodgkin lymphoma (non-HL) in the United States and accounts for approximately 45% of all newly diagnosed cases. It has a peak incidence in the fifth and sixth decades and is rare under the age of 20 years. Both sexes are equally affected. Most patients have stage 3 or 4 disease at diagnosis, with generalized lymphadenopathy and bone marrow involvement. Approximately 10% of patients have circulating malignant cells; appropriate immunophenotypic or molecular analyses may reveal involvement of peripheral blood in a higher proportion of patients.

FL is composed of varying proportions of follicle center–type cells, centrocytes, and centroblasts, with centroblasts representing the proliferative component. According to the WHO classification, all low-grade FL are combined into a single category, grade 1/2—with an overall predominance of centrocytes and fewer than 15 centroblasts per high-power field (hpf). Grade 3 (with >15 centroblasts/hpf) is further subdivided into 3A and 3B, based on the presence or absence of background centrocytes. FL represents the neoplastic counterpart of the reactive germinal-center cells; intraclonal heterogeneity with high numbers of somatic and ongoing Ig mutations can be detected in the neoplastic cells, as in the normal counterparts. Biologically, grade 3B is more closely related to diffuse large B-cell lymphoma (DLBCL) compared with FL.

The vast majority of FL (approximately 90%) is associated with a t(14; 18) involving rearrangement of the *BCL2* gene. This translocation appears to result in constitutive expression of BCL2 protein, which is capable of inhibiting apoptosis in lymphoid cells. The cells of FL accumulate and are at risk for secondary mutations, which may be associated with histological progression. It is thought that the *BCL2* translocation occurs at a very early stage of B-cell development, during immunoglobulin gene rearrangement. Mutations in *BCL2* can lead to loss of BCL2 protein in the presence of the translocation; fluorescence *in situ* hybridization (FISH) studies for t(14; 18) in these cases are informative. Of note, t(14; 18) is not specific to FL and can be seen in CLL and other low-grade B-cell lymphomas as well. Grade 3B FL is less commonly associated with *BCL2* translocation and carries genetic aberrations more commonly seen in DLBCL. The neoplastic cells in FL have a mature B-cell phenotype with expression of the B-cell antigens (CD19, CD20, and CD22). Surface Ig is positive, most commonly with IgM expression, but IgG or IgA can also be seen in many cases. CD10 and BCL6 are positive, but CD5 is usually negative.

The 2008 WHO classification had recognized variants of FL, namely, *pediatric FL, GI tract FL,* and other *extranodal FL.* Pediatric-type FL (now a definite entity in the 2016 classification) is more common in males and presents as localized nodal disease. It typically has an expansile serpiginous pattern, with relatively monotonous, medium-sized, blastoid cellular composition. *BCL2, BCL6,* and *MYC* rearrangements are lacking. Complete remissions may be obtained with either surgical excision or local radiation therapy. Some studies have raised the possibility that pediatric-type FL might be a "benign clonal proliferation with low malignant potential."[11,12]

The WHO classification recognizes "FL *in situ*" (now termed *in-situ follicular neoplasia* [ISFN]) as a distinctive lesion. It should be distinguished from partial involvement of FL. ISFN shows involvement of germinal centers by CD10 and BCL2-positive cells carrying t(14;18) in an otherwise reactive lymph node, which is typically an incidental finding. Patients have a very low risk of progression, but ISFN may be detected in association with other forms of B-cell lymphoma, necessitating additional clinical assessment.[13] Fewer chromosomal abnormalities are noted in IFSN in comparison with partial or overt FL.[14] Patients lacking evidence of FL at staging have a low risk of developing the disease; this phenomenon appears to represent the tissue counterpart of circulating clonal B-cells carrying t(14;18) as detected in healthy individuals. A higher level of circulating t(14;18) positive lymphocytes ($>10^{-4}$ of total cells) indicates a higher risk for FL.

The new 2016 classification addresses GI tract FL as duodenal-type FL, as these can also be seen elsewhere in the GI tract and have features overlapping with ISFN and extranodal mucosa-associated lymphoreticular tissue (MALT) lymphoma.[14] These present as small mucosal polyps or nodules, and most cases are discovered incidentally on endoscopy. The lesions are usually low-grade and have an indolent clinical course, and most patients have been managed without therapy. Local recurrences in the intestine may occur, but spread beyond the small intestine is rare.

Primary cutaneous follicle center lymphoma frequently lacks the *BCL2* translocation and BCL2 expression. However, when BCL2 expression is detected, the possibility that this may represent a secondary site of involvement should be considered. These lesions are generally managed locally with conservative approaches and have a low risk of spread beyond the skin. The scalp is a common site of presentation.

Two recent variants or subtypes of FL have been described; one subtype described by Karube et al. occurring in older individuals showed a $CD10^-/MUM1^+$ immunophenotype that tends to show grade 3 morphology, usually lacks *BCL2* translocations, and often shows *BCL6* abnormalities.[15] This is in contrast to another recently described distinct new provisional entity of a low-stage large B-cell lymphoma with *IRF4* rearrangement common in children or young adults, typically involving the Waldeyer ring and/or the cervical lymph nodes. These cases show strong IRF4/MUM1 expression, usually with BCL6 and a high proliferative fraction, and are considered more aggressive than other pediatric-type FL; however, when treated, they show good response.[16] Another distinctive subtype described by Katzenberger et al.[17] is characterized by t(14;18)–negative nodal grade 1/2 FL with a high proliferation rate that predominantly shows diffuse growth pattern and deletions in the chromosomal region 1p36.

This entity usually presents with low clinical stage and large or localized inguinal tumors.[17]

FL is indolent but is still incurable with available therapeutic modalities. The occurrence of *BCL2* translocation at a very early stage of B-cell development might be contributory to the difficulty of eradicating the neoplastic clone with chemotherapy. Clinical parameters have been used to develop prognostic indexes (*e.g.,* the Follicular Lymphoma International Prognostic Index, adjusted from the International Prognostic Index for aggressive B-cell lymphoma). Recently, much has also been learned about the mutational landscape of FL. Mutations in *CREBBP, MLL2,* and *EZH2* have been shown as extremely common early events and may be potential therapeutic targets.[18] The natural history of the disease is associated with histological progression at both cellular composition and overall pattern levels (Fig. 79.2).

Mucosa-Associated Lymphoid Tissue Lymphomas

Most lymphomas of marginal zone derivation present in extranodal sites and have histopathological and clinical features that are part of the spectrum of MALT lymphomas. MALT lymphomas occur most frequently in the stomach, lung, thyroid gland, salivary gland, and lacrimal gland. Other less common sites of involvement include the orbit, breast, conjunctiva, bladder, kidney, and thymus. MALT lymphomas are characterized by a heterogeneous cellular composition, including centrocyte-like cells, monocytoid B cells, small lymphocytes, and plasma cells. In most cases, large transformed cells are uncommon, but reactive germinal centers are nearly always present. Historically, the distinction from reactive lesions has been problematic. Clonality can be established on the basis of light-chain restriction or molecular studies. Follicular colonization by the neoplastic cells can simulate FL. The clinical course is usually quite indolent. MALT lymphomas tend to relapse in other MALT-associated sites. MALT lymphomas of the salivary gland, thyroid gland, and mediastinum (thymus) are usually associated with a history of autoimmune diseases, predominantly Sjögren syndrome.

MALT lymphomas have several recurring cytogenetic abnormalities, including t(11;18)(q21;q21), t(1;14)(p22;q32), t(14;18)(q32;q21), t(3;14)(q27;q32), and t(3;14)(p14.1;q32); these abnormalities are observed with variable frequency, often depending on the anatomical site.[19] Although several genes are involved in these translocations, at least three of them—t(11;18), t(1;14), and t(14;18)—share a common pathway, leading to the activation of NFκB and its downstream targets.[20] By genome-wide DNA profiling integrated with GEP, differences were detected among the three main types of marginal zone lymphomas, lending support to the current WHO classification that separates these three entities.[18] The translocation t(11;18)(q21;q21) is associated exclusively with low-grade extranodal MALT and is not detected in cases with concurrent low-grade and high-grade tumors, or in primary extranodal large-cell lymphomas—raising doubts whether these primary extranodal lymphomas are, in fact, related to low-grade MALT. The term *extranodal marginal zone lymphoma* (MZL) of MALT type applies only to low-grade MALT; the term *high-grade MALT* should not be used for extranodal large B-cell lymphomas in a MALT site.

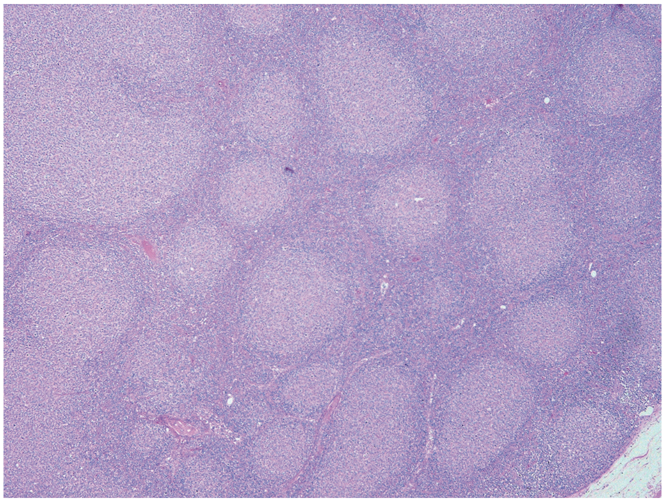


FIG 79.2 Follicular Lymphoma. The neoplastic follicles are similar in size and are partially surrounded by lymphoid cuffs. In contrast to reactive germinal centers, they lack polarization and tingible body macrophages ("starry-sky pattern").

There is a strong association between chronic infection with *Helicobacter pylori* and gastric MALT lymphoma. Other infectious agents have been described in MALT lymphomas involving skin (*Borrelia burgdorferi*), ocular adnexae (*Chlamydia psittaci*), and the small intestine (*Campylobacter jejuni*); however, a causal relationship has not yet been demonstrated. Chronic antigen stimulation is critical to both the development of a MALT lymphoma and the maintenance of the neoplastic state. Indeed, in some cases lacking aforementioned genetic aberrations, eradication of *H. pylori* by antibiotic therapy has led to the spontaneous remission of gastric MALT lymphoma. MALT lymphomas are positive for B cell–associated antigens but are usually negative for CD5 and CD10. BCL6 and CD10 are helpful markers to identify residual reactive germinal-center cells, especially in cases of follicular colonization. The putative cell of origin of MZL is a post–germinal-center memory B cell.

Nodal Marginal Zone Lymphoma

Nodal marginal zone lymphoma (NMZL), a primary nodal disease, resembles other marginal zone lymphomas, extranodal or splenic types. Adult patients often present with bone marrow involvement and tend to have a more aggressive clinical course compared with those with extranodal MALT. In contrast, when NMZL occurs in children, it shows a striking male predominance, presence of disrupted follicles with progressive transformation of germinal centers, presents with localized (head and neck) disease, and can be managed with local therapies.[19]

The neoplastic proliferation is mostly composed of small- to medium-sized B cells, often with pale cytoplasm. The immunophenotype is similar to that of other marginal zone B-cell lymphomas, $CD20^+$, $CD5^-$, $CD10^-$, with variable IgD expression. Some NMZLs have a morphology and immunophenotype similar to those of splenic marginal zone lymphomas (SMZLs) (see below), and in these cases, the tumor cells are IgD^+.

Splenic Marginal Zone Lymphomas

SMZLs present in adults and have a slight female gender predilection, usually with splenomegaly but without peripheral lymphadenopathy. Most patients have bone marrow involvement with modest lymphocytosis. Some evidence of plasmacytoid differentiation and the presence of a low level M-protein may also be seen. The course is reportedly indolent, and splenectomy may be followed by a prolonged remission.

Histologically, the spleen shows expansion of the white pulp with a characteristic biphasic pattern—central zone of small lymphocytes surrounded by a peripheral zone of larger cells resembling marginal zone cells. The residual polytypic mantle is not present in this instance. The abundant pale cytoplasm evident in tissue sections may also be seen in peripheral blood smears, and the cytological features may be mistaken for those of hairy-cell leukemia. The phenotype resembles other marginal zone B-cell lymphomas with a more frequent sIgD expression.

Deletion of 7q31 has been described in SMZL; although it can also be seen in isolated cases of splenic B-cell lymphoma or leukemia unclassifiable, it has not been detected in other splenic lymphomas.[21] Studies of the Ig variable genes and IgD expression have revealed two groups of naïve and memory B-cell origin.

CLINICAL PEARLS

Aggressive Lymphomas

- Natural history: survival measured in months.
- Successful therapy can be achieved with combination chemotherapy.
- Relapses from chemotherapy-induced remission may be cured with high-dose chemotherapy with hematopoietic support.
- In addition to the International Prognostic Index (IPI), gene expression profiling can be useful in predicting prognosis and survival of individual patients.

Diffuse Large B-Cell Lymphoma, Not Otherwise Specified

Diffuse large B-cell lymphoma, not otherwise specified (DLBCL, NOS) is one of the more common subtypes of non-HL, representing up to 40% of cases. This diagnosis is used for both primary DLBCL as well as for cases that transformed from a low-grade lymphoma. It may be nodal or involve extranodal sites, including bone, skin, the thyroid gland, the GI tract, and lungs.

DLBCL, NOS is composed of large transformed lymphoid cells with nuclei at least twice the size of a small lymphocyte (Fig. 79.3). The nuclei generally have vesicular chromatin, prominent nucleoli, and basophilic cytoplasm—resembling either centroblasts or immunoblasts, albeit with an overall greater cellular pleomorphism.

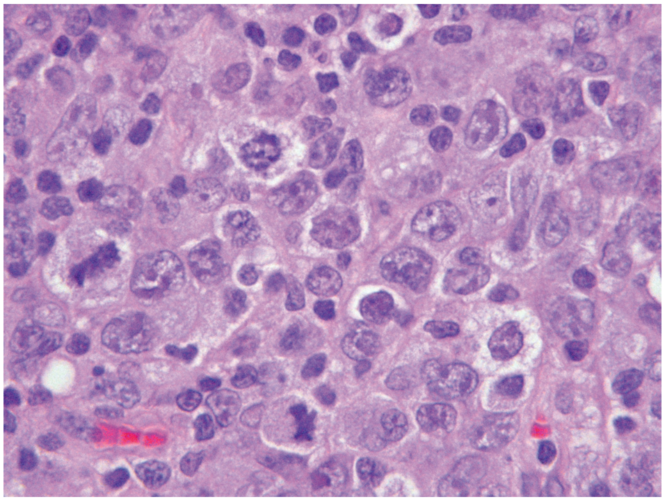


FIG 79.3 **Diffuse Large-Cell Lymphoma.** The neoplastic cells have large round to oval nuclei with vesicular chromatin and multiple eosinophilic nucleoli. Numerous mitoses are also present.

In terms of morphology and phenotype, DLBCL is one of the most heterogeneous categories in the WHO classification; currently, several morphological variants as well as specific subtypes are recognized.[2] There has been a great interest in identifying DLBCL features that might be prognostically relevant. To address issues traditionally not resolved by morphological or immunophenotypic features, DLBCL was among the first lymphomas to be analyzed by complementary DNA (cDNA) microarray. Based on the differential expression of a large set of genes by GEP, *germinal-center B cell–like* (GCB) group and *activated B cell–like* (ABC) group have been identified,[22] in addition to the previously recognized *primary mediastinal (thymic) large B-cell lymphoma* (PMBCL). GCB DLBCLs express a set of genes that are associated with normal germinal-center B cells, whereas ABC DLBCLs show downregulation of these genes and share similarities with post–germinal-center B cells.

The t(14;18) translocation involving *BCL2* and immunoglobulin heavy-chain gene has been detected in the GCB subtype, but not in the other subtypes. Previous studies have shown reduced disease-free survival in DLBCL cases with BCL2 overexpression, irrespective of the translocation. Although no absolute correlation between morphology and GEP has been established, interestingly, the majority of DLBCLs with centroblastic morphology falls into the GCB subtype, whereas those with immunoblastic morphology usually correlate with the ABC subtype.

DLBCL has an aggressive natural history but generally responds well to chemotherapy. The complete remission rate with modern regimens is 75–80%. Currently, with R-CHOP (rituximab–cyclophosphamide, doxorubicin, vincristine, and prednisone) regimen, the 10-year progression-free and overall survival rates for older patients with advanced-stage DLBCL are 36.5% and 43.5%, respectively.[23]

Immunophenotype-based algorithms using CD10/BCL6 positivity for GCB and MUM1/IRF4 expression for the ABC subtype have been proposed as surrogates for cDNA microarray; additionally, BCL2 and IPI are also informative to stratify the DLBCL cases. Because of the therapeutic impact, in the recent publication of the proposed changes in the WHO 2016 classification, it has been recommended that these two subtypes be identified, on the basis of either commonly used immunophenotype-based algorithms or GEP.

T cell/histiocyte–rich large B-cell lymphoma (THRLBCL) is a distinct clinicopathological entity, rather than a morphological variant. THRLBCL tends to occur in younger patients compared with other DLBCL, NOS, and often presents with advanced stage and bone marrow involvement with an aggressive clinical behavior.

The WHO classification recognizes that lymphomas arising from certain anatomical sites may have distinctive clinical and biological features. Among these are *primary DLBCL of the central nervous system* (CNS) and *primary cutaneous DLBCL, leg type.* Primary DLBCL of the CNS with some distinctive GEP features shares some similarities with DLBCL arising in other immune-privileged sites, such as the testis. Primary cutaneous DLBCL, leg type, has a GEP resembling the ABC subtype of DLBCL, presents most often in older females, and generally has an aggressive clinical course.

Several Epstein-Barr virus (EBV)–positive B-cell lymphoproliferations are often grouped with DLBCL. In the revised WHO 2016 classification, EBV-positive DLBCL, NOS[3] has now replaced a previous provisional entity "*EBV-positive DLBCL of the elderly*" in the 2008 WHO classification, since it may occur in younger patients with a broader morphological spectrum and better survival[24,25] as a consequence of decreased immune surveillance. These should be distinguished from EBV-associated atypical hyperplasia and lesions associated with iatrogenic or age-related immunosuppression, as well as those typified by an isolated and circumscribed cutaneous or mucosal presentation, with an indolent behavior and a self-limiting clinical course as EBV-positive mucocutaneous ulcer.[26] *Lymphomatoid granulomatosis* is an EBV-positive B-cell lymphoproliferative disorder (LPD) associated with an inflammatory background rich in T cells. The lung is nearly always involved; additionally, skin, kidneys, the liver, and the brain are frequently affected as well. *DLBCL associated with chronic inflammation* are EBV-driven large B-cell proliferations encountered in diverse clinical settings, usually associated with a confined anatomical space and a background of chronic inflammation. These cases appear to have a good prognosis if successfully resected.

Human herpes virus 6/Kaposi sarcoma herpes virus (HHV-8/KSHV)–associated LPDs are also documented. These include *primary effusion lymphoma* (PEL), *multicentric Castleman disease* (MCD), and lymphomas arising in the context of MCD. The cells of PEL are usually coinfected with EBV, and the disease is most often diagnosed in the setting of HIV infection and immunosuppression. Although pleural or peritoneal effusions are most common, extracavitary PEL can present as a tumor mass, usually in extranodal sites. PEL has a phenotype resembling terminally differentiated B cells (*i.e.,* plasmablastic).

Two other lymphomas with a plasmablastic phenotype include *plasmablastic lymphoma* (PBL) and *ALK-positive large B-cell lymphoma.* PBL is usually positive for EBV, most often extranodal, and is associated with immunosuppression from either human immunodeficiency virus (HIV) infection or advanced age. Recent studies have identified a high incidence of *MYC* translocation in PBL.[27] ALK-positive large B-cell lymphomas show overexpression of ALK, usually as a consequence of translocation. They mainly affect older individuals but can occur at any age.

Primary Mediastinal Large B-Cell Lymphoma

PMBCL has a distinct constellation of clinical and morphological features. PMBCL shows marked female gender predominance in adolescents and young adults. The clinical presentation is that of a rapidly growing anterior mediastinal mass with frequent superior vena cava syndrome and/or airway obstruction. Nodal involvement is uncommon at presentation and also at relapse. Frequent extranodal sites of involvement, particularly at relapse, include the liver, kidneys, adrenal glands, ovaries, the GI tract, and the CNS. The treatment approach includes aggressive systemic chemotherapy plus rituximab, along with radiation therapy used in some centers. A relatively abundant pale cytoplasm with distinct cytoplasmic membrane and fine compartmentalizing sclerosis are characteristic. The tumor appears to be derived from medullary B cells within the thymus gland. These cells express CD20 and CD79a, but not surface Ig. CD23 is frequently positive, and MUM-1/IRF4 coexpression is also common. A unique signature was identified by GEP in PMBCL that shared similarities with HL cell lines, including constitutive activation of the nuclear factor-κB and recurrent gains and amplification of c-Rel.[28] The expression of the *MAL* gene has been detected in PMBCL but not in other DLBCLs.

The 2008 WHO classification included borderline categories, one of which manifests features intermediate between DLBCL, especially PMBCL and classic HL (CHL) "gray-zone lymphomas."[29] A close relationship between PMBCL and CHL has been supported by GEP.[28] Tumor necrosis factor receptor–associated factor 1 expression and c-REL amplification also are seen in both types of neoplasms and could also be detected with appropriate immunohistochemical studies.

Gray-zone lymphomas are more common in males than females, present with bulky mediastinal masses, and appear to have a more aggressive clinical course compared with either PMBCL or CHL. Previously, methylation profiling showed a signature distinct from both CHL and PMBCL.[30] However, as shown by FISH studies, gray-zone lymphomas, PBMCL, and CHL all share a number of common cytogenetic aberrations, including gains at 2p16.1 (*REL/BCL11A* locus), 9p24.1 (*JAK2/PDL2*), and rearrangements of 16p13.13 (*CIITA*). DA-EPOCH R has been effective in both PMBL and DLBCL; however, it is not clear how these patients should be managed clinically, but they appear to benefit from combined modality therapy.[31,32]

CLINICAL PEARLS

Highly Aggressive Lymphomas

- More common in children
- Natural history similar to acute leukemia
- Successful treatment includes high-dose chemotherapy (induction, consolidation, and maintenance phases) with central nervous system prophylaxis

Burkitt Lymphoma

Burkitt lymphoma (BL) occurs most commonly in children and accounts for up to one-third of all pediatric lymphomas in the United States. Three clinical variants of BL are recognized—endemic, sporadic, and immunodeficiency-associated. The endemic cases are seen in equatorial Africa and mostly involve the jaw and other facial bones. African BL occurs in malaria-endemic regions. In non–malaria-endemic regions, such as the United States, extranodal sites are frequent, including the ileocecal region, ovaries, kidneys, or breasts. Bone marrow involvement is a sign of poor prognosis.

EBV has been shown to be a cofactor for the development of BL and shows varying degrees of positivity in the variant subtypes. BL is one of the more common tumors associated with HIV. It can present at any time during the clinical course and also can be the initial acquired immunodeficiency syndrome (AIDS)–defining illness. Recent GEP data have supported a common pathogenetic mechanism in cases of HIV infection and endemic malaria–related immunosuppression.

Cytologically, BL shows monomorphic medium-sized cells with round nuclei, multiple (2–5) basophilic nucleoli, and moderately abundant deeply basophilic cytoplasm. Cytoplasmic lipid vacuoles reflecting a high proliferation rate and apoptosis are frequent. It is the most rapidly growing of all lymphomas, with 100% of the cells in cell cycle at any time. The characteristic "starry sky" pattern of BL is a manifestation of the numerous benign macrophages that have ingested karyorrhectic or apoptotic tumor cells. BL has a mature B-cell phenotype, expressing CD19, CD20, CD22, CD79a, and monoclonal surface Ig, nearly always IgM. CD10 is positive in almost all cases, whereas CD5, CD23, and BCL2 are consistently absent.

The pathogenesis of BL is related to the *MYC* oncogene translocations seen in virtually 100% of cases. The *MYC* translocation is considered a primary event and often is the sole karyotypic abnormality detected. This is in contrast to other aggressive lymphomas, in which the *MYC* translocation occurs as a secondary event in a more complex karyotype.[33] Most of the translocations involve the *IGH* gene on chromosome 14 and, less commonly, the light-chain genes on chromosomes 2 and 22.

Recently, a subset with chromosome 11q alterations that morphologically, largely phenotypically, and by GEP resemble BL but lack *MYC* translocations has been described[34] as a new provisional entity "Burkitt-like lymphoma with 11q aberration" in the revised classification. Mutations in the transcription factor *TCF3* or its negative regulator *ID3* are present in approximately 70% of sporadic and immunodeficiency-related BL and 40% of endemic cases.

Few cases show GEP of BL but carry additional *BCL2* or *BCL6* cytogenetic abnormalities. These "double-hit" or "triple-hit" lymphomas have an aggressive clinical course and poor prognosis and are now reclassified from the previous borderline category of the 2008 WHO classification now as HGBL with and without *MYC* and *BCL2* or *BCL6* translocations (except for those fulfilling the criteria for follicular or lymphoblastic lymphoma).[3] Of note, the "double-hit" lymphoma are usually of GCB subtype, whereas "double expressors" are of the ABC subtype. However, this category should not be utilized for otherwise typical DLBCLs with a *MYC* translocation.

T-CELL AND NK-CELL NEOPLASMS

Overview of the Classification of T-Cell Neoplasms

Peripheral T-cell lymphomas (PTCLs) are uncommon, representing fewer than 10% of all non-HLs. The classification of PTCL has always been controversial. The genetic landscape of many entities has only recently been defined, and immunophenotypic markers are less specific for apparently distinct disease entities. For these reasons, the WHO classification relied, to a considerable extent, on clinical presentation to subdivide these tumors.[2]

Extranodal NK/T-Cell Lymphoma, Nasal Type

Extranodal NK/T-cell lymphoma, nasal type, is a distinct clinicopathological entity highly associated with EBV. It affects adults (median age 50 years), and the most common clinical presentation is a destructive nasal or midline facial lesion. Palatal destruction, orbital swelling, and edema can be prominent. NK/T-cell lymphomas have been reported in other extranodal sites, including skin, soft tissue, testes, the upper respiratory tract, and the GI tract. It is much more common in Asians and indigenous populations of the Americas than in those of European background, and this indicates that genetic factors play a role in the pathogenesis of these lymphomas.

Extranodal NK/T-cell lymphoma, nasal type, is characterized by a broad cytological spectrum. Although the cells express some T cell–associated antigens, most commonly CD2, other T-cell markers, such as surface CD3, are usually absent. Cytoplasmic CD3 is positive, but most cases lack clonal T-cell gene rearrangement. In favor of an NK-cell origin, the cells are nearly always CD56$^+$, although CD16 and CD57 are usually negative. EBV is invariably positive as shown by *in situ* hybridization.

The clinical features and treatment response of nonnasal extranodal NK/T-cell lymphoma are different from of those of nasal presentation of this lymphoma in that the addition of radiotherapy in early-stage nasal cases may offer a survival benefit.[35] A hemophagocytic syndrome is a common clinical complication and adversely affects survival. Emerging oncogenic pathways have been identified by GEP.

Aggressive NK-cell leukemia is a closely related entity with a similar phenotype, EBV association, and epidemiology. It presents at a younger age than extranodal NK/T-cell lymphoma, is associated with systemic disease, and has a fulminant clinical course.

There are other EBV$^+$ T-cell and NK-cell proliferations seen mainly in Asian children and in indigenous populations from Central and South Americas and Mexico. These show a broad range of clinical manifestations from indolent, localized forms involving skin, such as *hydroa vacciniforme-like lymphoproliferative disorder* (name changed from lymphoma) and *mosquito bite allergy,* the latter usually being derived from NK cells, to *systemic EBV$^+$ T-cell lymphoma* (name changed from lymphoproliferative disease),[3] a more systemic form characterized by fever, hepatosplenomegaly and lymphadenopathy with or without cutaneous manifestations[36] and a fulminant clinical course, and hemophagocytic lymphohistiocytosis.

Nodal T-Cell Lymphoma With T_{FH} Phenotype: Angioimmunoblastic T-Cell Lymphoma

Recently angioimmunoblastic T-cell lymphoma (AITL) has been grouped under an umbrella category "nodal T-cell lymphoma with T follicular helper cell (T_{FH}) phenotype" to highlight a spectrum of nodal lymphomas with T_{FH} phenotype, including follicular T-cell lymphoma and other nodal PTCLs with a T_{FH} phenotype, in addition to AITL.[3]

AITL presents in adults with generalized lymphadenopathy and prominent systemic symptoms, including fever, weight loss, and skin rash. Polyclonal hypergammaglobulinemia is usually seen.

Histologically, the nodal architecture is generally effaced, but peripheral sinuses are often open or dilated. Proliferation of high endothelial venules (HEVs) is often prominent. Follicles are regressed, but a proliferation of dendritic cells (DCs) often entrapping HEVs is typically seen. The atypical lymphoid cells have clear cytoplasm and are admixed with small lymphocytes, immunoblasts, plasma cells, and histiocytes, with or without eosinophils. A relationship to T_{FH} has been confirmed by GEP. The atypical T cells are usually positive for CD4, CD10, and PD-1, a phenotype characteristic of T_{FH}. Chemokine ligand-13 (CXCL-13), a chemokine involved in B-cell trafficking into the germinal centers, is also expressed. In keeping with an established derivation from T_{FH}, B-cell proliferation, including marked polyclonal plasmacytosis, is often seen. In some cases, the plasma cells may be monoclonal. Background EBV$^+$ B cells are almost constant, and progression to EBV$^+$ B-cell lymphoma may occur. The exact role of EBV in AILT is uncertain; however, plausible theories include expansion resulting from decreased immune surveillance. The majority of cases show clonal rearrangements of T-cell receptor (TCR) genes, but proliferation of B cells also contributes to the histological and clinical picture. Patients may initially respond to steroids or mild cytotoxic chemotherapy, but progression usually occurs. More aggressive combination chemotherapeutic regimens have led to a higher remission rate, but patients are prone to secondary infectious complications. The median survival is usually less than 5 years.

New genetic insights have shown recurrent mutations, including *TET2, IDH2, DNMT3A, RHOA,* and *CD28,* as well as gene fusions, such as *ITK-SYK* or *CTLA4-CD28,* which affect a significant proportion of PTCL, NOS cases with a T_{FH} phenotype, and AITL cases.

Peripheral T-Cell Lymphomas, Not Otherwise Specified

Peripheral T-cell lymphomas, not otherwise specified (PTCL, NOS) most often present in adults with generalized lymphadenopathy, hepatosplenomegaly, and frequent bone marrow involvement. It is a diagnosis of exclusion with most cases being nodal in origin. PTCLs are characterized by cytological and phenotypical heterogeneity. Most cases have a mature T-cell phenotype and express one of the major subset antigens, with CD4 expression seen more frequently than CD8. These are not clonal markers, and antigen expression can change over time. Loss of one of the pan-T-cell antigens (CD3, CD5, CD2, or CD7) is seen in two-thirds of cases, with loss of CD7 being most frequent.

Recently, GEP has shown a global signature close to one of the activated T lymphocytes and has identified at least three subtypes characterized by overexpression of *GATA3, TBX21,* and cytotoxic genes associated with a different clinical behavior and response to therapy. The clinical course is generally aggressive, especially with a high proliferation signature, with a lower response rate than that seen for aggressive B-cell lymphomas.

Anaplastic Large-Cell Lymphoma

Anaplastic large-cell lymphoma (ALCL) is most common in children and young adults, and more common in males than in females. Nodal presentations are most common; however, a variety of extranodal sites can be involved.

ALCL is characterized by pleomorphic or monomorphic cells that have a propensity to invade lymphoid sinuses. The cells of classic ALCL have large, often lobulated nuclei with small basophilic nucleoli. In some cases, the nuclei may be round. The cytoplasm is usually abundant and amphophilic and has distinct cytoplasmic borders, and a prominent Golgi region is generally visible.

The expression of the CD30 antigen is a hallmark of this disease. However, CD30 expression is not specific for ALCL and may also be seen in other forms of malignant lymphoma, including CHL. ALCL-ALK$^+$ is associated with a characteristic chromosomal translocation, t(2;5)(p23;q35), involving the *ALK* and *NPM* genes, respectively. A variety of other *ALK* partners have been identified and monoclonal antibodies to the ALK protein have been able to identify tumor cells regardless of the underlying translocation. The neoplastic cells show both nuclear and cytoplasmic staining in the majority of cases; the presence of cytoplasmic staining for ALK is suggestive of a variant translocation.

Immunohistochemistry is indispensable in the correct diagnosis of ALCL. The prominent Golgi region usually shows intense staining for CD30 and epithelial membrane antigen (EMA). The cells exhibit an aberrant phenotype with loss of many of the T cell–associated antigens. Both CD3 and CD45RO, the most widely used pan-T-cell markers, are negative in >50% of cases. CD2 and CD4 are positive in the majority, whereas CD8 is usually negative. ALCL cells, despite the CD4$^+$/CD8$^-$ phenotype, often express the cytotoxic-associated antigens (TIA-1, granzyme B, and perforin). In addition, clusterin is generally present in ALCL and represents another useful diagnostic marker. Molecular studies in most cases demonstrate TCR rearrangement, confirming a T-cell origin.

Improved criteria now exist for the recognition of ALK-negative ALCL as a separate entity. It occurs in an older age group compared with the ALK-positive cases. Recently, a unique form of ALK-negative ALCL arising in association with breast implants has been identified,[37] with a medial interval from time of implant to lymphoma of about 10 years. Confinement of the neoplastic cells to the seroma fluid without capsule invasion portends a favorable prognosis.

Recently GEP studies have shown that ALK-negative ALCLs have a signature close to that of ALK-positive counterparts, but distinct from other NK/T-cell lymphomas. Genetic studies have shown convergent mutations and kinase fusions leading to constitutive activation of the JAK/STAT3 pathway. The overall survival and disease-free survival are significantly better in ALK-positive cases than in ALK-negative cases. Clinical or prognostic variations exist in ALK-negative ALCLs—a subset with rearrangements at the *DUSP22* and *IFR4* locus on 6p25 has a superior prognosis, whereas a small subset with *TP63* rearrangements is very aggressive.

Primary Cutaneous ALCL

Primary cutaneous ALCL is closely related to lymphomatoid papulosis (LyP) and differs at the clinical, immunophenotypic, and molecular levels from the systemic form. Indeed, LyP and cutaneous ALCL represent a histological or clinical continuum of CD30$^+$ cutaneous lymphoproliferative diseases. Small lesions are likely to regress, whereas patients with large tumor masses may develop disseminated disease with lymph node involvement—a period of observation is usually warranted before the institution of any chemotherapy for isolated lesions. Most patients with primary cutaneous ALCL have multiple skin lesions, but it is a more indolent disease compared with other T-cell lymphomas of the skin. Cutaneous ALCL is CD30$^+$ but is usually ALK-1– and EMA-negative, and it also lacks the t(2;5) translocation. More recently, LyP variants D-E and LyP with 6p25 have been described; appreciation of these variants is important as they can mimic aggressive T-cell lymphoma histologically but clinically are similar to other forms of LyP A-C.

Subcutaneous Panniculitis-Like T-Cell Lymphoma

Subcutaneous panniculitis-like T-cell lymphoma (SPTCL) usually presents with subcutaneous nodules primarily affecting the extremities. In its early stages, the infiltrate may appear deceptively benign, and lesions are often misdiagnosed as panniculitis. However, histological progression usually occurs, and subsequent biopsies show more pronounced cytological atypia.

The neoplastic cells are CD8$^+$ cytotoxic α/β T cells. They are also positive for the cytotoxic proteins perforin, granzyme B, and TIA-1 that may be responsible for the cellular destruction seen in these tumors. EBV is negative. Some PTCL of γ/δ T-cell derivation may show similar features but differ from SPTCL in their clinical behavior (more aggressive) and histological pattern, as they are often not confined to the subcutis.

Some patients have a history of autoimmune disease, and in particular, the differential diagnosis of SPTCL with lupus profundus panniculitis can be challenging. Unlike SPTCL, lupus panniculitis usually contains abundant plasma cells, a mixture of CD4$^+$ and CD8$^+$ cells, and a relative increase of γ/δ T cells and plasmacytoid DCs.

A hemophagocytic syndrome can be a complication of SPTCL but is more often associated with γ/δ T-cell lymphomas involving the subcutaneous tissue. Patients present with fever, pancytopenia, and hepatosplenomegaly. It is most readily diagnosed in bone marrow aspirate smears, where histiocytes containing erythrocytes, platelets, and other blood elements may be observed. The hemophagocytic syndrome usually heralds a fulminant downhill clinical course.

Primary Cutaneous γ/δ T-Cell Lymphomas

Primary cutaneous γ/δ T-cell lymphomas are clinically aggressive tumors that can present with involvement of the subcutis and the dermis or with epidermal infiltration. Skin is the most common presenting site, however, similar lymphomas of γ/δ T-cell origin can present in other extranodal sites, including the GI tract and lungs. The cells have a cytotoxic phenotype and express cytotoxic molecules, and like normal γ/δ T cells, lack CD5. They may be CD8^{+} or, more often, double-negative for CD4 and CD8. Few cutaneous γ/δ T-cell lymphomas appear to be TCR silent (negative for both TCR-β and TCR-γ) but share the same aggressive clinical behavior. Recurrent mutations of *STAT5B* can be seen in T-cell lymphomas of γ/δ origin.[38]

Mycosis Fungoides and Sézary Syndrome

Mycosis fungoides (MF) and Sézary syndrome (SS) are closely related and often considered together from a clinical and biological standpoint but are now regarded as separate diseases. Both are primary cutaneous T-cell malignancies derived from mature skin homing CD4 T cells. Epidermotropism is the hallmark of MF; infiltration of the epidermis produces characteristic Pautrier microabscesses. The cutaneous lesions are categorized as patches, plaques, and tumors, based on the extent of the infiltrate. A γ/δ indolent variant of MF has also been described. SS presents with exfoliative erythroderma and circulating cerebriform lymphocytes known as *Sézary cells.* Clinically, SS is more aggressive than MF.

Enteropathy-Associated T-Cell Lymphoma

Enteropathy-associated T-cell lymphoma (EATL) is highly associated with celiac disease on a worldwide basis. This disease occurs in adults; the majority has either overt or clinically silent gluten-sensitive enteropathy (Chapter 75). Ulcerative jejunitis may precede the development of overt EATL and may share a common clonal T-cell population as with the subsequent lymphoma. The small bowel usually shows ulceration with frequent perforation, with or without a mass. EATL shows a cytological composition of variably sized or polymorphic atypical lymphoid cells. The adjacent small bowel usually shows villous atrophy associated with celiac disease. The neoplastic cells are CD3^{+}, CD7^{+} T cells that also express the homing receptor CD103. Anaplastic cells strongly positive for CD30 can be present. The cells express cytotoxic molecules, a feature shared by nearly all extranodal T-cell lymphomas. The majority belong to the α/β TCR subset, whereas only a minority expresses the γ/δ TCR. Other PTCLs can present with intestinal disease, including EBV^{+} extranodal T/NK-cell lymphomas and γ/δ T-cell lymphomas, and should be distinguished from EATL. The clinical course is aggressive.

The disease previously known as *EATL type II* has been now been formally designated as *monomorphic epitheliotropic intestinal T-cell lymphoma* (MEITL); it shows no association with celiac disease and appears to have an increased incidence in Asian and Hispanic populations. Unlike classic EATL, MEITL is monomorphic, usually positive for CD8, CD56, and MAPK.[39] *STAT5B* mutations are common in MEITL cases, many of which are of γ/δ T-cell origin.

Hepatosplenic T-Cell Lymphoma

Hepatosplenic T-cell lymphoma (HSTCL) shows a marked male gender predominance, and most patients are young adults. The clinical presentation is that of marked hepatosplenomegaly in the absence of lymphadenopathy.

The liver and the spleen show marked sinusoidal infiltration, with sparing of both portal triads and white pulp, respectively. The homing pattern manifested by the malignant cells is similar to that of normal γ/δ T cells, which also populate the sinusoidal areas of the spleen. The neoplastic cells have a phenotype that resembles normal γ/δ T cells, usually expressing neither CD4 nor CD8. CD56 is often positive. The cells are positive for the cytotoxic protein TIA-1, but are not activated and generally lack granzyme B and perforin. *In situ* hybridization for EBV is negative. The recognition of the atypical cells in bone marrow is greatly aided by immunohistochemical stains showing a sinusoidal pattern of infiltration.

Isochromosome 7 is a consistent cytogenetic abnormality, usually seen in conjunction with trisomy 8. Rare cases derived from α/β T cells but with similar morphological and biological features can be seen. Recurrent mutations of *STAT5B* and less often *STAT3* are seen in HSTCL of γ/δ origin.[38]

Clinically, HSTCL is aggressive. Although patients may respond initially to chemotherapy, relapse occurs in the vast majority of cases. The median survival is less than 3 years. Allogeneic bone marrow transplantation is required for sustained remission.

Adult T-Cell Leukemia/Lymphoma

Adult T-cell leukemia/lymphoma (ATLL) is a distinct form of T-cell lymphoma associated with the retrovirus human T-lymphotropic virus-1 (HTLV-1). The highest number of cases is seen in Southwestern Japan and the Caribbean basin. The disease has a long latency, and affected individuals are usually exposed to the virus very early in life. The virus may be transmitted in breast milk and through exposure to blood or blood products. The cumulative incidence of ATLL is estimated to be 2.5% among HTLV-1 carriers. The virus is monoclonally integrated into tumor DNA. Different clinical variants of ATLL, including the acute, lymphomatous, chronic, and smoldering types, have been recognized. Cutaneous involvement is seen in the majority of patients.

Peripheral blood involvement is very common, often without bone marrow disease. The atypical cells are often markedly polylobated and have been referred to as "flower" cells. The cells have a characteristic phenotype that resembles T-regulatory cells (Tregs)—CD3^{+}, CD4^{+}, and CD25^{+}. FoxP3 is expressed in a minority of tumor cells. The function of the tumor cells as Tregs may correlate with the associated immunodeficiency.

Somatic gain of function *CCR4* mutations have been implicated in the pathogenesis of ATLL.[40] Recent integrated molecular analysis has shown frequent intragenic deletions and mutations of key T-cell signaling molecules in ATLL.[41]

CLINICAL PEARLS

Hodgkin Lymphoma

- B-cell lineage established in nearly all cases
- Reed-Sternberg cell, the hallmark of the disease, represents a "crippled" germinal-center B cell
- Nodular lymphocyte-predominant Hodgkin lymphoma considered a related but distinct entity
- Eighty percent of patients are curable with current therapy
- Stage of disease guides the choice of therapy; even patients with advanced-stage disease may be cured
- Late complications from treatment include acute leukemia (alkylating agents with extended-field radiation therapy), second solid tumors (radiation therapy), and premature atherosclerotic coronary artery disease (radiation therapy)
- Cause of death in the first 5–10 years is mainly Hodgkin lymphoma; after 10 years, mainly secondary malignant tumors

HODGKIN LYMPHOMAS

HL and non-HL have long been regarded as distinct disease entities on the basis of their differences in pathology, phenotype, clinical features, and response to therapy. It is now accepted that the malignant cell of HL is an altered B cell. Therefore it is not surprising that biological and clinical overlap occurs between these two lymphoma groups, as also shown by GEP in PMBCL and classic Hodgkin cell lines. In spite of the close histogenetic relationship (hence the name *Hodgkin lymphoma*), these disorders are still treated with different modalities.

The diagnosis of classic HL (CHL) depends on the identification of Hodgkin/Reed–Sternberg (HRS) cells in an appropriate inflammatory background composed of small T lymphocytes, plasma cells, histiocytes, and granulocytes (often eosinophils). All cases of CHL share certain immunophenotypic and genotypic features. The phenotype is $CD30^+$, $CD15^{+/-}$, $CD45^-$, and EMA^-. Expression of B cell–associated antigens is seen in up to 75% of cases; however, when present, CD20 staining is usually weaker or more variable than that seen in normal B cells. CD79a is usually negative. Ig and TCR genes are usually germline because of the paucity of tumor cells in the inflammatory background; however, microdissection can enable amplification for clonal rearrangement of the Ig genes by polymerase chain reaction (PCR). In addition, the presence of somatic mutations indicates transit through the germinal center.

Sufficient evidence has emerged in recent years to warrant the recognition of nodular lymphocyte-predominant HL as a distinct entity. Although it resembles other types of HL in having a minority of putative neoplastic cells in a background of benign inflammatory cells, it differs morphologically, immunophenotypically, and clinically from CHL. The preferred term *Hodgkin lymphoma,* versus *Hodgkin disease,* reflects current knowledge concerning the nature of the neoplastic cell as a lymphocyte.

Nodular Lymphocyte-Predominant Hodgkin Lymphoma

Nodular lymphocyte-predominant Hodgkin lymphoma (NLPHL) occurs at all ages but is more common in adult males. It usually involves peripheral lymph nodes with sparing of the mediastinum and is localized at diagnosis, although rarely it may be disseminated. NLPHL usually has a nodular growth pattern, with or without diffuse areas. The number of infiltrating reactive T cells is variable, and various patterns have been described on the basis of the cellular composition and growth pattern.[42] The atypical cells have vesicular, polylobated nuclei and small nucleoli. These had been called *lymphocytic and/or histiocytic (L&H) cells,* or “popcorn” cells, but the term *lymphocyte-predominant (LP) cell* is now preferred. LP cells differ from classic HRS cells. The background is composed predominantly of lymphocytes with or without clusters of epithelioid histiocytes. Plasma cells are infrequent; eosinophils and neutrophils are also rare. The atypical cells are $CD45^+$, positive for B cell–associated antigens (CD19, CD20, CD22, CD79a), $CDw75^+$, $EMA^{+/-}$ $CD15^-$, $CD30^{-/+}$ and usually sIg^-, as shown by routine techniques. Small lymphocytes in the nodules are predominantly B cells with a mantle zone phenotype. However, numerous T cells are present, with $CD279^+$ T cells “rosetting” the LP cells. The proportion of T cells tends to increase over time in sequential biopsies. A prominent follicular DC meshwork is present within the nodules. LP cells, when microdissected, have shown to have clonally rearranged immunoglobulin genes with evidence of somatic hypermutation.

Survival for localized cases with or without treatment is long. However, cases with variant histology are more likely to have advanced stage disease with a higher relapse rate.[43] Moreover, patients with advanced-stage disease respond poorly to HL regimens, such as adriamycin, bleomycin, vinblastine, dacarbazine (ABVD), and benefit from treatment for aggressive B-cell lymphoma.[44] Although considered two separate entities with major differences, NLPHL shares immunomorphological similarities with THRLBCL; these similarities have also been shown by GEP and array comparative genomic hybridization. Two forms of histological progression can be seen, either to DLBCL or to a process resembling THRLBCL.

Classic Hodgkin Lymphoma, Nodular Sclerosis

This variant of HL, classic Hodgkin lymphoma, nodular sclerosis (CHLNS) is most commonly seen in adolescents and young adults but can occur at any age. Female cases equal or exceed those in males. The mediastinum is commonly involved; stage and bulk of disease have prognostic importance.

The tumor has at least a partially nodular pattern, with fibrous bands separating the nodules in most cases. Diffuse areas with necrosis may be present. The characteristic cell is the “lacunar-type” Reed-Sternberg (RS) cell, which may be very numerous. Diagnostic RS cells are usually also present. The background contains lymphocytes, histiocytes, plasma cells, eosinophils, and neutrophils. Grading (I and II) is based on the proportion of the tumor cells and the presence of necrosis but is considered optional. The immunophenotype and genotype are characteristic of CHL. However, EBV is infrequently positive— less than 15% of cases.

CHLNS is often curable; however, in long-term survivors, the risk of secondary malignancies is increased, especially in those receiving both chemotherapy and radiation. CHLNS of the mediastinum is thought to be closely related to PMBCL, and both types of tumors can be seen in the same patient, either as composite malignancy or sequentially.

Classic Hodgkin Lymphoma, Mixed Cellularity

Classic Hodgkin lymphoma, mixed cellularity (CHLMC) is predominantly seen in male adults. A bimodal age distribution, with peaks in young children and later in older adults can be present. Both CHLMC and the lymphocyte-depleted form (see below) can be associated with underlying HIV infection. The infiltrate is diffuse without band-forming sclerosis, although fine interstitial fibrosis may be present (Fig. 79.4). HRS cells are of the classic type. It is often EBV-positive, seen in up to 75% of cases. The stage is often advanced at diagnosis. The clinical course is moderately aggressive, but is often curable.

Classic Hodgkin Lymphoma, Lymphocyte Depletion

Classic Hodgkin lymphoma, lymphocyte depletion (CHLLD) is the least common variant of CHL and is most common in older people, in HIV-positive individuals, and in populations of nonindustrialized countries. It frequently presents with abdominal lymphadenopathy, and spleen, liver, and bone marrow involvement, but without peripheral adenopathy. The infiltrate is diffuse and often appears hypocellular because of the presence of diffuse fibrosis and necrosis. A large number of HRS cells and occasional bizarre “sarcomatous” variants are seen relative to the fewer number of normal lymphocytes and scarce other inflammatory cells. The immunophenotype is characteristic of CHL. Since the

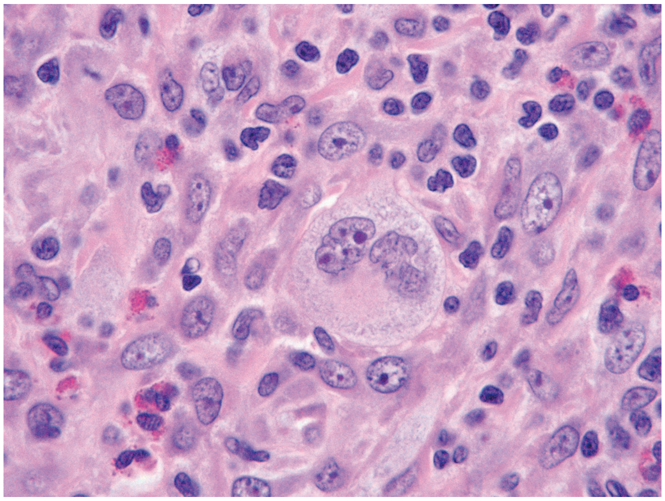


FIG 79.4 **Classic Hodgkin Lymphoma, Mixed Cellularity Subtype.** A classic Reed-Sternberg cell is shown in a mixed inflammatory background with eosinophils, plasma cells, histiocytes, and small lymphocytes.

histological differential diagnosis often includes B- or T-large-cell lymphoma or ALCL, immunohistochemistry should be performed in most cases. EBV is positive in the majority of cases. The stage is usually advanced at diagnosis.

Classic Hodgkin Lymphoma, Lymphocyte-Rich

Classic Hodgkin lymphoma, lymphocyte-rich (CHLLR) may be nodular or diffuse and contains relatively infrequent classic HRS cells. Eosinophils and plasma cells are rare. In the nodular form, the HRS cells are seen at the periphery of B cell–rich nodules, mainly in the marginal zone. CHLLR has some features that are intermediate between other CHL and NLPHL.[45] Immunophenotypically, the neoplastic cells resemble classic HRS cells, but morphological distinction from LP cells may be difficult in some cases. Thus, in the past, many cases were misdiagnosed as NLPHL. However, patients usually present with localized disease and tend to be older than patients with NLPHL. The genetic features are similar to other variants of CHL.

ON THE HORIZON

- In recent years, there has been a greater appreciation of early events in lymphoid neoplasia.
- These early lesions in some ways can be considered equivalent to benign neoplasms in the epithelial system.
- These are clonal proliferations of B or T cells that carry genetic aberrations associated with specific forms of lymphoid neoplasia: chronic lymphocytic leukemia, multiple myeloma, follicular lymphoma, mantle cell lymphoma.
- Examples include monoclonal gammopathy of undermined significance, monoclonal B-lymphocytosis, *in situ* mantle cell neoplasia, *in situ* follicular neoplasia, lymphomatoid papulosis, patch stage of mycosis fungoides.
- Early lesions appear to lack the secondary and tertiary "hits" seen in lymphoid neoplasms that are clinically significant, and most patients have a very low risk of clinical progression.
- Current focus is to define the precise genetic features that distinguish early lesions from lymphoma, assess the risk of clinical progression, and determine their clinical management.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Nogai H, Dorken B, Lenz G. Pathogenesis of non-Hodgkin's lymphoma. J Clin Oncol 2011;29:1803–11.
2. Swerdlow SH, Campo E, Harris NL, et al. WHO classification of tumours of haematopoietic and lymphoid tissues. 4th ed. Lyon, France: International Agency for Research on Cancer; 2008.
3. Swerdlow SH, Campo E, Pileri SA, et al. The 2016 revision of the World Health Organization (WHO) classification of lymphoid neoplasms. Blood 2016;pii: blood-2016-01-643569.
4. Campo E, Swerdlow SH, Harris NL, et al. The 2008 WHO classification of lymphoid neoplasms and beyond: evolving concepts and practical applications. Blood 2011;1175:19–32.
5. Treon SP, Xu L, Yang G, et al. MYD88 L265P Somatic mutation in Waldenström's macroglobulinemia. N Engl J Med 2012;367:826–33.
6. Fu K, Weisenburger DD, Greiner TC, et al. Cyclin D1-negative mantle cell lymphoma: a clinicopathologic study based on gene expression profiling. Blood 2005;106:4315–21.
7. Salaverria I, Royo C, Carvajal-Cuenca A, et al. CCND2 rearrangements are the most frequent genetic events in cyclin D1(-) mantle cell lymphoma. Blood 2013;121:1394–402.
8. Mozos A, Royo C, Hartmann E, et al. SOX11 expression is highly specific for mantle cell lymphoma and identifies the cyclin D1-negative subtype. Haematologica 2009;94:1555–62.
9. Fernandez V, Salamero O, Espinet B, et al. Genomic and gene expression profiling defines indolent forms of mantle cell lymphoma. Cancer Res 2010;70:1408–18.
10. Campo E, Rule S. Mantle cell lymphoma: evolving management strategies. Blood 2015;125:48–55.
11. Liu Q, Salaverria I, Pittaluga S, et al. Follicular lymphomas in children and young adults: a comparison of the pediatric variant with usual follicular lymphoma. Am J Surg Pathol 2013;37:333–43.
12. Louissaint A Jr, Ackerman AM, Dias-Santagata D, et al. Pediatric-type nodal follicular lymphoma: an indolent clonal proliferation in children and adults with high proliferation index and no BCL2 rearrangement. Blood 2012;120:2395–404.
13. Jegalian AG, Eberle FC, Pack SD, et al. Follicular lymphoma in situ: clinical implications and comparisons with partial involvement by follicular lymphoma. Blood 2011;118:2976–84.
14. Mamessier E, Song JY, Eberle FC, et al. Early lesions of follicular lymphoma: a genetic perspective. Haematologica 2014;99:481–8.
15. Karube K, Guo Y, Suzumiya J, et al. CD10-MUM1+ follicular lymphoma lacks BCL2 gene translocation and shows characteristic biologic and clinical features. Blood 2007;109:3076–9.
16. Salaverria I, Philipp C, Oschlies I, et al. Translocations activating IRF4 identify a subtype of germinal center-derived B-cell lymphoma affecting predominantly children and young adults. Blood 2011;118:139–47.
17. Katzenberger T, Kalla J, Leich E, et al. A distinctive subtype of t(14;18)-negative nodal follicular non-Hodgkin lymphoma characterized by a predominantly diffuse growth pattern and deletions in the chromosomal region 1p36. Blood 2009;113:1053–61.
18. Loeffler M, Kreuz M, Haake A, et al. Genomic and epigenomic co-evolution in follicular lymphomas. Leukemia 2015;29:456–63.
19. Swerdlow SH. Pediatric follicular lymphomas, marginal zone lymphomas, and marginal zone hyperplasia. Am J Clin Pathol 2004;122:S98–109.
20. Farinha P, Gascoyne RD. Molecular pathogenesis of mucosa-associated lymphoid tissue lymphoma. J Clin Oncol 2005;23:6370–8.
21. Watkins AJ, Huang Y, Ye H, et al. Splenic marginal zone lymphoma: characterization of 7q deletion and its value in diagnosis. J Pathol 2010;220:461–74.
22. Rosenwald A, Wright G, Chan WC, et al. The use of molecular profiling to predict survival after chemotherapy for diffuse large-B-cell lymphoma. N Engl J Med 2002;346:1937–47.

23. Coiffier B, Thieblemont C, Van Den Neste E, et al. Long-term outcome of patients in the LNH-98.5 trial, the first randomized study comparing rituximab-CHOP to standard CHOP chemotherapy in DLBCL patients: a study by the Groupe d'Etudes des Lymphomes de l'Adulte. Blood 2010;116:2040–5.
24. Dojcinov SD, Venkataraman G, Pittaluga S, et al. Age-related EBV-associated lymphoproliferative disorders in the Western population: a spectrum of reactive lymphoid hyperplasia and lymphoma. Blood 2011;117:4726–35.
25. Nicolae A, Pittaluga S, Abdullah S, et al. EBV-positive large B-cell lymphomas in young patients: a nodal lymphoma with evidence for a tolerogenic immune environment. Blood 2015;126:863–72.
26. Dojcinov SD, Venkataraman G, Raffeld M, et al. EBV positive mucocutaneous ulcer—a study of 26 cases associated with various sources of immunosuppression. Am J Surg Pathol 2010;34:405–17.
27. Taddesse-Heath L, Meloni-Ehrig A, et al. Plasmablastic lymphoma with MYC translocation: evidence for a common pathway in the generation of plasmablastic features. Mod Pathol 2010;23:991–9.
28. Rosenwald A, Wright G, Leroy K, et al. Molecular diagnosis of primary mediastinal B cell lymphoma identifies a clinically favorable subgroup of diffuse large B cell lymphoma related to Hodgkin lymphoma. J Exp Med 2003;198:851–62.
29. Traverse-Glehen A, Pittaluga S, Gaulard P, et al. Mediastinal gray zone lymphoma: the missing link between classic Hodgkin's lymphoma and mediastinal large B-cell lymphoma. Am J Surg Pathol 2005;29:1411–21.
30. Eberle FC, Rodriguez-Canales J, Wei L, et al. Methylation profiling of mediastinal gray zone lymphoma reveals a distinctive signature with elements shared by classical Hodgkin's lymphoma and primary mediastinal large B-cell lymphoma. Haematologica 2011;96:558–66.
31. Wilson WH, Grant C, et al. Therapy in primary mediastinal B-cell lymphoma. New Engl J Med 2013;369:283–4.
32. Dunleavy K, Pittaluga S, Maeda LS, et al. Dose-adjusted EPOCH-rituximab therapy in primary mediastinal B-cell lymphoma. New Engl J Med 2013;368:1408–16.
33. Dave SS, Fu K, Wright GW, et al. Molecular diagnosis of Burkitt's lymphoma. N Engl J Med 2006;354:2431–42.
34. Salaverria I, Martin-Guerrero I, Wagener R, et al. A recurrent 11q aberration pattern characterizes a subset of MYC-negative high-grade B-cell lymphomas resembling Burkitt lymphoma. Blood 2014;123:1187–98.
35. Au WY, Weisenburger DD, Intragumtornchai T, et al. Clinical differences between nasal and extranasal natural killer/T-cell lymphoma: a study of 136 cases from the International Peripheral T-Cell Lymphoma Project. Blood 2009;1133931–7.
36. Quintanilla-Martinez L, Ridaura C, Nagl F, et al. Hydroa vacciniforme-like lymphoma: a chronic EBV+ lymphoproliferative disorder with risk to develop a systemic lymphoma. Blood 2013;122:3101–10.
37. Thompson PA, Lade S, Webster H, et al. Effusion-associated anaplastic large cell lymphoma of the breast: time for it to be defined as a distinct clinico-pathological entity. Haematologica 2010;95:1977–9.
38. Nicolae A, Xi L, Pittaluga S, et al. Frequent STAT5B mutations in gammadelta hepatosplenic T-cell lymphomas. Leukemia 2014;28:2244–8.
39. Deleeuw RJ, Zettl A, Klinker E, et al. Whole-genome analysis and HLA genotyping of enteropathy-type T-cell lymphoma reveals 2 distinct lymphoma subtypes. Gastroenterology 2007;132:1902–11.
40. Nakagawa M, Schmitz R, Xiao W, et al. Gain-of-function CCR4 mutations in adult T cell leukemia/lymphoma. J Exp Med 2014;211:2497–505.
41. Kataoka K, Nagata Y, Kitanaka A, et al. Integrated molecular analysis of adult T cell leukemia/lymphoma. Nat Genet 2015;47:1304–15.
42. Fan Z, Natkunam Y, Bair E, et al. Characterization of variant patterns of nodular lymphocyte predominant Hodgkin lymphoma with immunohistologic and clinical correlation. Am J Surg Pathol 2003;27:1346–56.
43. Hartmann S, Eichenauer DA, Plutschow A, et al. The prognostic impact of variant histology in nodular lymphocyte-predominant Hodgkin lymphoma: a report from the German Hodgkin Study Group (GHSG). Blood 2013;122:4246–52.
44. Xing KH, Connors JM, Lai A, et al. Advanced-stage nodular lymphocyte predominant Hodgkin lymphoma compared with classical Hodgkin lymphoma: a matched pair outcome analysis. Blood 2014;123:3567–73.
45. Nam-Cha SH, Montes-Moreno S, Salcedo MT, et al. Lymphocyte-rich classical Hodgkin's lymphoma: distinctive tumor and microenvironment markers. Mod Pathol 2009;22:1006–15.

MULTIPLE-CHOICE QUESTIONS

1. In the diagnostic workup of a diffuse large B-cell lymphoma, which markers are helpful in the identification of the cell of origin?
   A. CD10/BCL6/MUM1
   B. CD10/P53/MIB-1
   C. CD10/BCL2/MUM1
   D. CD10/BCL6/MIB1

2. The presence of MYD88 L265P mutation is associated with the diagnosis of:
   A. Mantle cell lymphoma
   B. Follicular lymphoma
   C. Chronic lymphocytic leukemia
   D. Lymphoplasmacytic lymphoma
   E. Multiple myeloma

3. Which of the following lymphomas are typically associated with Epstein-Barr virus (EBV) positivity?
   A. Enteropathy-associated T-cell lymphoma
   B. Extranodal natural killer (NK)-/T-cell lymphoma, nasal type
   C. Hepatosplenic T-cell lymphoma
   D. Anaplastic large-cell lymphoma, ALK positive
   E. Anaplastic large-cell lymphoma, ALK negative

80

Monoclonal Gammopathies

Kazunori Murata, Samuel McCash, C. Ola Landgren

Monoclonal gammopathies are a group of disorders characterized by the proliferation of a single clone of plasma cells. These plasma cells produce an immunologically homogeneous immunoglobulin, or parts thereof, also known as an M (for monoclonal) protein.

This chapter covers 12 recognized types of monoclonal gammopathies: monoclonal gammopathy of undetermined significance (MGUS); multiple myeloma (MM); smoldering multiple myeloma (SMM); nonsecretory/oligosecretory myeloma; plasma cell leukemia (PCL); solitary plasmacytoma of bone (SPB); solitary extramedullary plasmacytoma (SEP); polyneuropathy, organomegaly, endocrinopathy, monoclonal protein, skin changes (POEMS) syndrome; Waldenström macroglobulinemia (WM); heavy chain disease (HCD); light chain amyloidosis (AL); and monoclonal immunoglobulin deposition disease. These disorders differ in the types of monoclonal proteins that are generated by the diseased cell. The proteins involved include immunoglobulin (Ig)-G, IgA, IgM, IgD, IgE, kappa or lambda light chains only, and immunoglobulin heavy chains only.

MONOCLONAL GAMMOPATHY OF UNDETERMINED SIGNIFICANCE

MGUS is a clinically asymptomatic premalignant clonal plasma cell or lymphoplasmacytic proliferative disorder. It is defined by the presence of the following characteristics: the presence of an M-protein in the serum at a concentration <3 g/dL, bone marrow with <10% monoclonal plasma cells, and the absence of end-organ damage related to the lymphoproliferative process. There are three distinct clinical types of MGUS: non-IgM MGUS, IgM MGUS, and light chain MGUS (LC-MGUS).

Epidemiology

MGUS occurs in approximately 1–2% of adults and >2% of the general Caucasian population above the age of 50. The mean age at diagnosis is 70 years with <2% of patients diagnosed before the age of 40. The incidence increases with age, and the prevalence of MGUS in persons ≥50, ≥70, and ≥85 years of age is 3.2%, 5.3%, and 7.5%, respectively.[1]

Men are at higher risk of developing MGUS compared with women (4.0% versus 2.7%). The incidence of MGUS in men is estimated to be 120/100 000 at age 50, increasing to 530/100 000 by the age of 90, whereas the incidence rate for women is estimated to be 60/100 000 at age 50, increasing to 370/100 000 by the age of 90. Race is also a risk factor, with the incidence of MGUS being 1.5- to 3-fold higher in Africans and African Americans compared with Caucasians.[1]

Most cases of MGUS are sporadic. However, relatives of persons with MGUS (or MM) have an increased risk of developing MGUS and related diseases. Acquired angioedema is closely associated with MGUS, being identified in approximately one-third of acquired angioedema patients.[1]

Clinical Presentation and Laboratory Findings

By definition, all patients with MGUS have no myeloma-associated symptoms attributable to the proliferation of plasma cells or the presence of the monoclonal protein. Non-IgM and IgM MGUS are characterized by the presence of an M-protein produced and secreted by clonal plasma cells. These proteins can be detected in serum using either protein electrophoresis or immunofixation electrophoresis as well as in a sample from a 24-hour urine collection using the same methods.

The quantity of monoclonal protein in all cases of MGUS is defined as <3 g/dL in the serum. Serum immunofixation electrophoresis (SIF) is used to confirm the presence of an M-protein and to determine its type. The clonal immunoglobulin distribution in MGUS varies according to isotype: IgG 69%, IgM 17%, IgA 11%, IgD <1%, and biclonal 3%.[2]

IgD MGUS is extremely rare with only two cases having been reported. Thus finding an IgD M-protein is almost always associated with the diagnosis of MM, light chain amyloid, or PCL. Biclonal MGUS is also uncommon, representing only 3% of patients with MGUS. Approximately one-third of MGUS patients will have a decrease in the concentration of the uninvolved immunoglobulin isotypes (e.g., IgM and IgA in the case of IgG MGUS), with many demonstrating a reduction in the normal polyclonal circulating immunoglobulins.[1]

Serum free light chain (sFLC) assays can detect low concentrations of monoclonal free light chain in patient serum. This technique can be used to diagnose light chain MGUS and to predict the risk of the progression of MGUS. A monoclonal light chain can also be found in the urine of approximately 20% of MGUS patients.[1]

The complete blood count and peripheral smear are typically normal. Infrequently, rouleaux formation (the phenomenon in which red cells take on the appearance of stacked coins in diluted suspensions of blood) is seen, and this is typically observed in those MGUS patients with elevated serum protein levels. Circulating plasma cells of the same isotype can be seen in peripheral blood of some patients with MGUS using a slide-based immunofluorescence assay or flow cytometry.[1]

By definition, a bone marrow aspirate and biopsy must demonstrate fewer than 10% clonal plasma cells. For IgM

MGUS, clonal lymphoplasmacytic cells must constitute <10% infiltration.

Conventional cytogenetics are normal in patients with MGUS due to technical limitations, a low proliferative rate, and the small number of plasma cells within the bone marrow samples. However, studies using more sensitive methods have shown that chromosomal abnormalities are very common in the clonal cells found in MGUS. Analysis by interphase fluorescent *in situ* hybridization (FISH) demonstrate chromosomal aneuploidy, in particular hyperdiploidy, in approximately 50% of patients with MGUS.[1]

Diagnosis

Diagnostic criteria for each of the individual types of MGUS are determined by the International Myeloma Working Group (IMWG) and are detailed in Table 80.1.[3]

Clinical Course

The clinical course of MGUS depends largely on the specific type.[4] In non-IgM MGUS a minority of cases will progress to the more advanced premalignant stage SMM and to symptomatic MM. Less frequently, these patients progress to light chain amyloidosis, light chain deposition disease (LCDD), or another lymphoproliferative disorder. IgM MGUS can progress to smoldering WM (SWM) and to symptomatic WM. IgM MGUS also infrequently progresses to IgM MM. LC-MGUS may progress to idiopathic Bence-Jones proteinuria, light chain MM, AL amyloidosis, or LCDD.

TABLE 80.1 IMWG Diagnostic Criteria for MGUS by Type

Non-IgM MGUS

1. The presence of a serum monoclonal protein (M-protein, whether IgA, IgG, or IgD) at a concentration <3 g/dL.
2. Fewer than 10% clonal plasma cells in the bone marrow.
3. The absence of lytic bone lesions, anemia, hypercalcemia, and renal insufficiency related to the plasma cell proliferative process.

IgM MGUS

1. The presence of a serum IgM monoclonal protein at a concentration <3 g/dL.
2. Fewer than 10% clonal lymphoplasmacytic/plasma cells in the bone marrow.
3. The absence of end-organ damage such as anemia, constitutional symptoms, hyperviscosity, lymphadenopathy, or hepatosplenomegaly related to the plasma cell proliferative process.

Light Chain MGUS

1. The presence of an abnormal free light chain ratio (*i.e.*, ratio of kappa to lambda free light chains <0.26 or >1.65).
2. Increased level of the appropriate involved light chain (*e.g.*, increased kappa FLC in patients with a ratio >1.65 and increased lambda FLC in patients with a ratio <0.26).
3. No monoclonal immunoglobulin heavy chain (IgG, IgA, IgD, or IgM).
4. Fewer than 10% clonal lymphoplasmacytic cells in the bone marrow.
5. The absence of lytic bone lesions, anemia, hypercalcemia, and renal insufficiency related to the plasma cell proliferative process.

FLC, free light chain; *Ig,* immunoglobulin; *IMWG,* International Myeloma Working Group; *MGUS,* monoclonal gammopathy of undetermined significance.
Source: International Myeloma Working Group. Criteria for the classification of monoclonal gammopathies, multiple myeloma and related disorders: a report of the International Myeloma Working Group. Br J Haematol 2003;121(5):749–57. PubMed PMID: 12780789.

Patients with MGUS progress to a symptomatic plasma cell proliferative disorder or lymphoproliferative disorder at a rate of 1% per year. However, not all persons with MGUS have the same risk of disease progression. The risk of progression to a more serious disease ranges widely from 0.6–3.4%/year based on the initial value of serum monoclonal protein as well as from 0.25–2.9%/year according to a risk stratification model.[4] Three adverse risk factors have been combined to create a risk stratification model that is useful in predicting progression of MGUS (non-IgM and IgM) to MM or a related malignancy[5]: The findings that generate a less favorable outcome include serum monoclonal protein level ≥1.5 g/dL, non-IgG MGUS, and abnormal serum free light chain ratio (kappa/lambda free light chain ratio <0.26 or >1.65).

The stratified absolute chance of disease progression over 20 years is 58% for high-risk MGUS associated with all three risk factors, 37% for high- or intermediate-risk MGUS based on two risk factors, 21% for low- or intermediate-risk MGUS associated with one risk factor, and 5% for low-risk MGUS based on no risk factors.[5] Risk factors for progression of LC-MGUS have yet to be defined.

KEY CONCEPTS

MGUS

- A premalignant clonal plasma cell/lymphoplasmacytic proliferative disorder.
- Asymptomatic by definition.
- Defined by the presence of a monoclonal protein (M-protein) in the serum or bone marrow with <10% monoclonal plasma cells and absence of end-organ damage related to the proliferative process.
- Three types of MGUS: non-IgM MGUS, IgM MGUS, and light chain MGUS.
- Progresses to a symptomatic plasma cell proliferative disorder or lymphoproliferative disorder at a rate of 1% per year.
- Patients with MGUS should not be treated but instead monitored for signs of progressive disease.

Management and Prognosis

Patients with MGUS should be followed over time with history and physical examination for signs and symptoms of progressive disease. Treatment of MGUS has not been shown to have any effect on mortality, and, as noted above, only a minority of patients will progress to symptomatic disease over a set time interval.[4] All other patients are followed with annual serum and urinary M-protein, complete blood count, creatinine, and serum calcium.[4] It has been observed that patients with MGUS are at increased risk of fracture and thromboembolic disease. Patients with MGUS should be evaluated for osteoporosis with a dual-energy X-ray absorptiometry (DEXA) scan and have their vitamin D and calcium intake optimized.[4]

MULTIPLE MYELOMA

MM is the neoplastic proliferation of a single clone of plasma cells producing an M-protein. The difference between MM and MGUS lies in the presence of detectable end-organ damage caused by the proliferation of plasma cells and/or the overproduction of M-protein. The proliferation of plasma cells in MM causes skeletal bone destruction, primarily in the axial skeleton, and may also result in anemia. Myeloma patients often present with pain and hypercalcemia caused by osteolytic lesions and/or pathological fractures. Overproduction of M-protein over time can contribute to renal failure, hyperviscosity syndrome, and/or recurrent bacterial infections.

Epidemiology

MM accounts for approximately 1% of all cancers and slightly more than 10% of hematological malignancies in the United States. The annual incidence in the United States is approximately 4 to 5 per 100 000. The incidence of MM in people of African background is two to three times that in Caucasians, whereas the risk is lower in Japanese and in Mexicans. MM is also slightly more frequent in men than in women with an incidence ratio of approximately 1.4 : 1. The median age at diagnosis is 66 years; only 10% and 2% of patients are diagnosed before 50 and 40 years, respectively. The risk of developing MM is approximately 3.7-fold higher for persons with a first-degree relative diagnosed with MM.[6]

Clinical Presentation

Most patients with MM present with signs and symptoms related to the infiltration of plasma cells into the bone or other organs as well as kidney damage from excess light chains. Patients often complain of fatigue and display pallor upon physical examination.[6]

Bone pain, particularly in the back or chest, and less often in the extremities, is present at the time of diagnosis in approximately 60% of patients. The pain is usually induced by movement and does not occur at night except with change of position. The patient's height may be reduced by several inches as a result of vertebral collapse. Plasmacytomas of the ribs occur and can present either as expanding costal lesions or soft-tissue masses.[6] Radiculopathy, usually in the thoracic or lumbosacral area, is the most common neurological complication of MM and is due to compression of the nerve by a paravertebral plasmacytoma or, in rare instances, by the bone collapse itself.[6]

Extramedullary plasmacytomas (EP) are seen in approximately 7% of patients with MM at the time of diagnosis. An additional 6% of patients will develop EP later during the course of disease, and these can present as large, purplish subcutaneous masses. Spinal cord compression from an extramedullary plasmacytoma or a bone fragment due to fracture of a vertebral body occurs in approximately 5% of patients. Intracranial plasmacytomas are rare and almost always represent extensions of myelomatous lesions of the skull or plasmacytomas involving the clivus or base of the skull.[6] Patients with MM are at increased risk for infection with *Streptococcus pneumoniae* and gram-negative organisms representing the most frequent pathogens.[6]

Laboratory Findings

Biochemical Tests

The serum creatinine concentration is increased in almost one-half of patients at diagnosis. Hypercalcemia was found in 28% of one series of patients with MM at the time of diagnosis. Serum calcium greater than 11 mg/dL has been found in 13% of patients and can require emergent treatment. Elevation of the serum calcium may be due to binding of the monoclonal protein to calcium. A low anion gap may be present due to severe hypercalcemia and the presence of a cationic IgG molecule.[6]

The vast majority (97%) of patients with MM will have an M-protein produced and secreted by the malignant plasma cells. M-proteins can be detected by protein electrophoresis of the serum (SPEP) and/or of urine (UPEP) from a 24-hour collection combined with SIF and urine (UIF).[6] SPEP will demonstrate a localized band or peak in 82% of patients with myeloma (Figs. 80.1 and 80.2).

Addition of serum protein immunofixation electrophoresis increases the sensitivity to 93%. Adding either the sFLC assay or urine monoclonal protein studies (UPEP and UIF) increases the diagnostic sensitivity to 97% or more. The very rare patient who lacks detectable M-protein by any of these tests is considered to have "nonsecretory myeloma." Among the 20% with no localized band on SPEP, hypogammaglobulinemia is seen in approximately one-half, with no apparent abnormality in the remainder.[6]

SIF confirms the presence of an M-protein and determines its type. The malignant plasma cells can produce a complete clonal immunoglobulin molecule consisting of heavy and light chains, light chains alone, or neither, with the following frequencies on serum immunofixation electrophoresis: IgG 52%, IgA 21%, IgD 2%, IgM 0.5%, biclonal 2%, kappa or lambda light chain only 16%, and negative 17%.[7]

Kappa is the predominant light chain compared with lambda by a ratio of 2 to 1, with the exception that lambda light chains are more common in IgD myeloma and myeloma associated with amyloidosis. The level of one of the major uninvolved immunoglobulins (*i.e.*, IgM or IgA in the case of IgG myeloma) is reduced in 91% of patients overall, and both isotypes are reduced in 73%. Up to 20% of myeloma patients have only monoclonal light chains in the serum or urine. Approximately 3% of patients with MM have no M-protein in the serum or urine on immunofixation at the time of diagnosis. In approximately 60% of patients with myeloma who have a normal serum and urine immunofixation, monoclonal sFLC can be detected in the serum using FLC assays.[6] M-proteins can increase the serum viscosity and erythrocyte sedimentation rate (ESR). The ESR is >20 mm/h in 84% and >100 mm/h in one-third of patients with MM.

The sFLC assay measures kappa and lambda light chains that are unbound to heavy chains in the serum. These assays typically report out three values: quantitation of free kappa chains, quantitation of free lambda chains, and the kappa/lambda ratio. The kappa/lambda ratio is obtained by dividing the free kappa value by the free lambda quantitation (For example, the kappa/lambda ratio of a patient with a kappa value of 6.32 mg/dL and a lambda value of 0.51 mg/dL would have a kappa/lambda ratio of 12.4). The kappa/lambda ratio currently plays a key role in the diagnosis and management of patients with monoclonal gammopathies. The normal kappa/lambda FLC ratio is 0.26 to 1.65. Abnormal FLC ratios outside this range are usually seen in clonal plasma cell disorders when there is excess production of one type of light chain (kappa or lambda). Abnormal FLC ratios are seen in approximately 90% of patients with MM. Patients with otherwise asymptomatic myeloma who have an involved/uninvolved FLC ratio of 100 or greater have a risk of progression to end-organ damage in the next 2 years of approximately 80%. For this reason, an involved/uninvolved FLC ratio of 100 or more is now considered diagnostic of MM.[6]

Hematology

A normocytic, normochromic anemia (hemoglobin ≤12 g/dL) is present in 73% at diagnosis and in 97% at some time during the course of MM. Macrocytosis (mean corpuscular volume >100 fL) has been shown to be present in 9% of patients. The most frequent findings on peripheral smear are rouleaux formation,

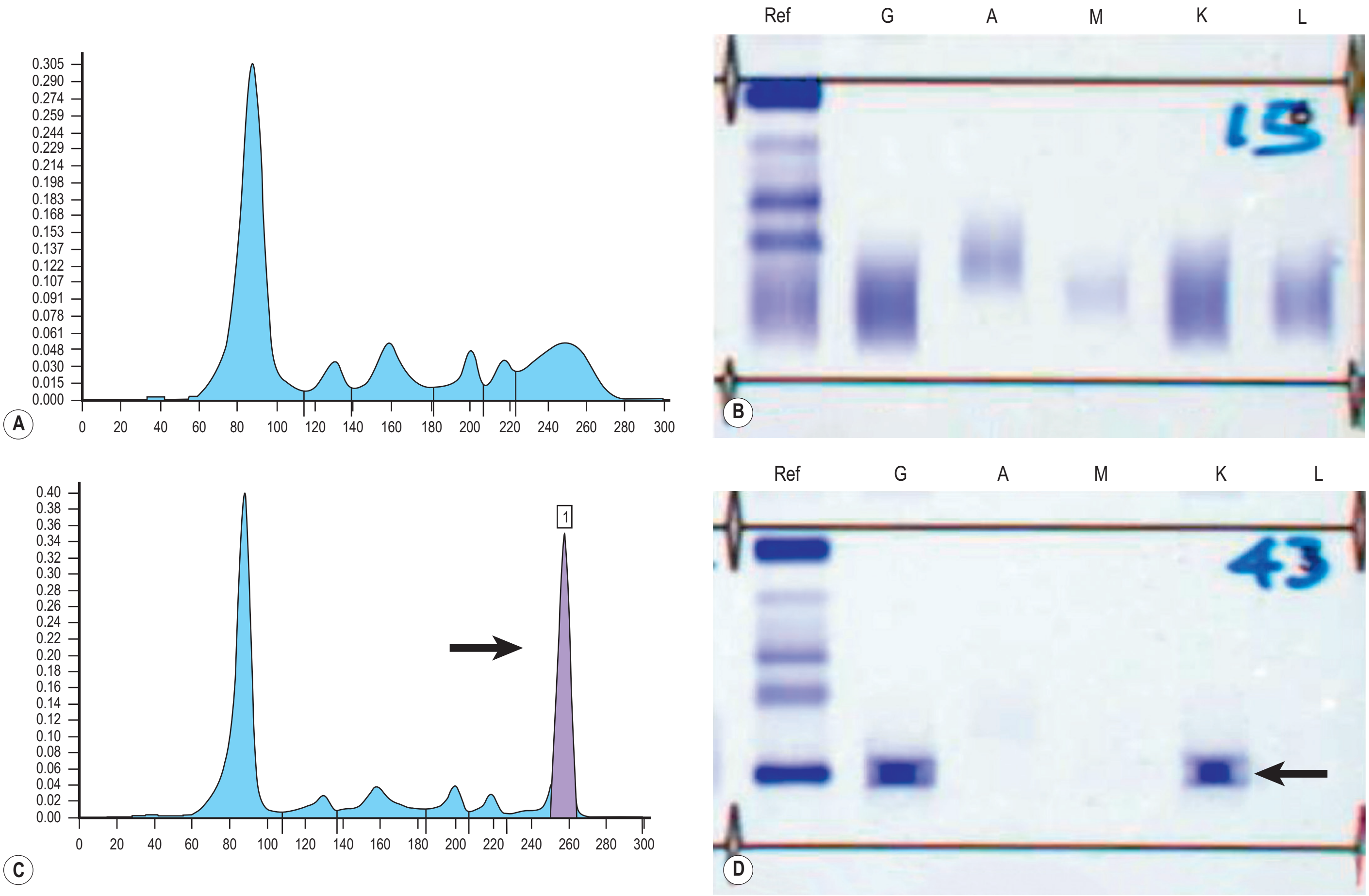


FIG 80.1 Electropherogram (A) and Immunofixation Electrophoresis (B) of Serum from a Normal Patient and a Myeloma Patient (C) and (D). Note the M-spike in the gamma region of the electropherogram and the immunoglobulin G (IgG) kappa band visible in the immunofixation of the myeloma patient specimen.

leukopenia, and thrombocytopenia. Rouleaux formation is particularly common in patients with elevated serum protein levels, as noted in the discussion of MGUS. Monoclonal plasma cells are rarely seen in the peripheral smear in patients with MM; a detectable absolute peripheral blood plasma cell count ≥100 cells/μL (≥0.1 × 10/L) is found in approximately 10%. Evaluation for circulating monoclonal plasma cells utilizes a slide-based immunofluorescence assay, a two-color immunoassay technique (ELISPOT), or flow cytometry by gating on CD38$^+$/CD45$^-$ cells.[6]

Bone Marrow Evaluation

Bone marrow aspirates and biopsies are key to the diagnosis of MM. The bone marrow plasma cell percentage should be estimated from a core biopsy specimen. However, if the percentage of plasma cells in the aspirate and core biopsy differ, the higher value should be used. Flow cytometry is not used to determine bone marrow plasma cell percentage for diagnostic purposes. Clonality can be established by demonstrating kappa/lambda light chain restriction using flow cytometry, immunohistochemistry, or immunofluorescence.[6]

The bone marrow of the vast majority of patients contains 10% or more clonal plasma cells. However, due to patchy bone marrow involvement, a bone marrow aspirate and biopsy may show less than 10% plasma cells in approximately 4% of patients.[6]

Mature plasma cells are oval with abundant basophilic cytoplasm. The nucleus is round and eccentrically located with a marked perinuclear cytoplasmic clearing, termed a hof. 80. The nucleus contains “clock face“ or “spoke wheel“ chromatin without nucleoli. Immature plasma cells have dispersed nuclear chromatin, prominent nucleoli, and a high nuclear to cytoplasmic ratio (Fig. 80.3).[6] The cytoplasm of myeloma cells may contain condensed or crystallized cytoplasmic immunoglobulin resulting in the following unusual findings, which are not limited to MM: multiple pale bluish-white, grapelike accumulations, cherry-red refractive round bodies, vermilion staining glycogen-rich IgA, overstuffed fibrils, and crystalline rods.

Immunohistochemical staining, immunofluorescent studies, and flow cytometry detect either kappa or lambda light chains, but not both, in the cytoplasm of bone marrow plasma cells in patients with MM while surface immunoglobulin is absent. The normal kappa/lambda ratio in the bone marrow is 2:1. A ratio of more than 4:1 or less than 1:2 is considered to meet the definition of kappa or lambda monoclonality, respectively.[6] As with normal plasma cells, myeloma cells express CD79a, p63, CD138, and CD38. In contrast to normal plasma cells, myeloma cells infrequently express CD19. Approximately 70% of myeloma cells express CD56, which is typically negative in normal plasma cells and in PCL.[6]

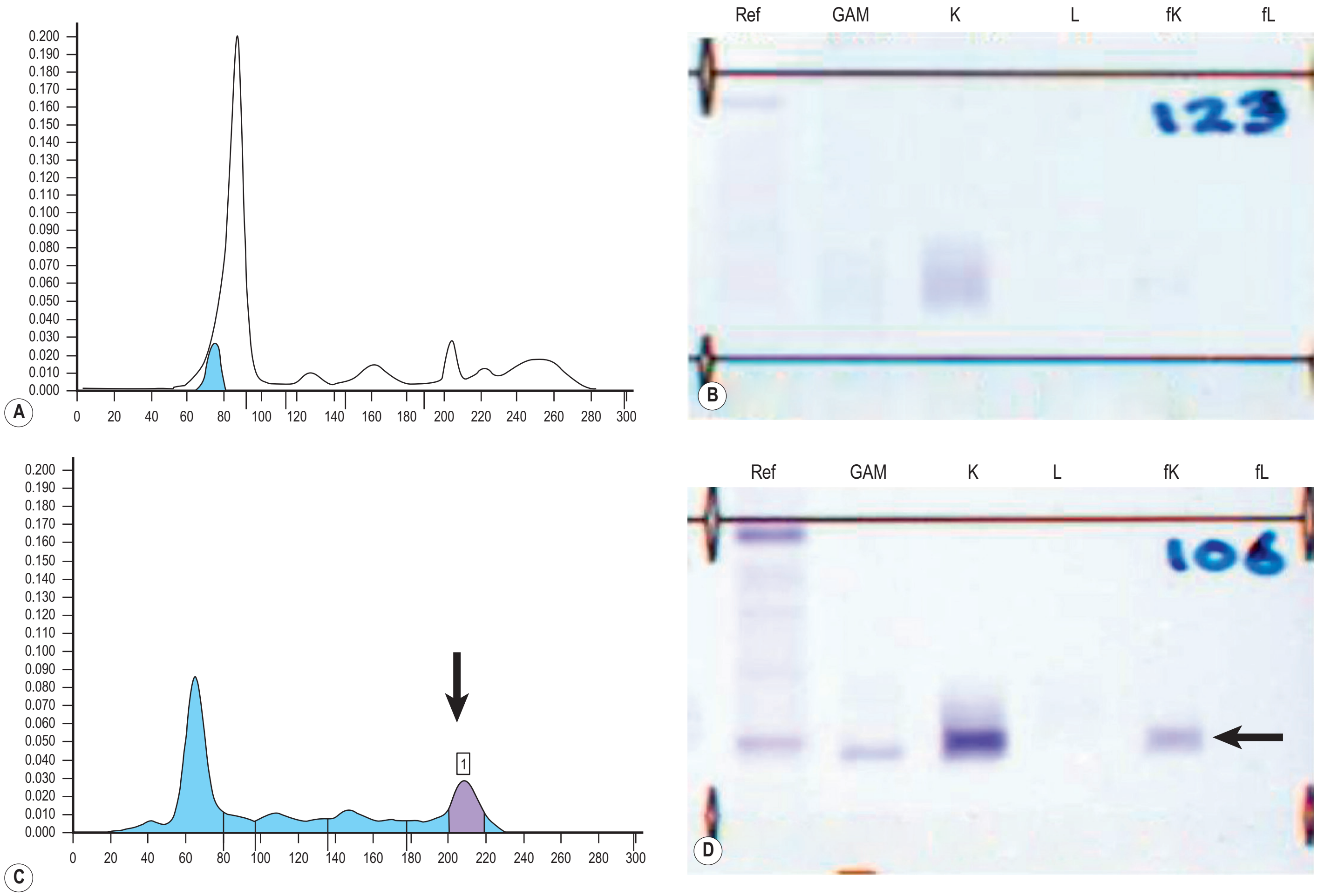


FIG 80.2 Electropherogram (A) and Immunofixation Electrophoresis (B) of Urine from a Normal Patient and a Myeloma Patient (C and D). Note the M-spike in the gamma region seen in the myeloma patient specimen. Both a monoclonal free kappa band as well as a monoclonal intact immunoglobulin band is visible in the urine immunofixation of the myeloma patient, indicating the presence of both monoclonal free kappa light chains (Bence-Jones protein) and monoclonal immunoglobulin M-protein in the patient's urine.

There is no specific cytogenetic abnormality that is typical or diagnostic of MM. The majority of myeloma tumors have genetic abnormalities that can be detected with sensitive molecular genetic techniques, such as interphase FISH. In contrast, only 20–30% of patients will have cytogenetic abnormalities detected in bone marrow plasma cells by conventional karyotyping; this is due to the low number of metaphases in myeloma cells.[6]

Radiography

A metastatic bone survey with plain radiographs including the humeri and femurs is a key component to the evaluation of a patient suspected of having MM (Fig. 80.4). The skeletal survey for patients with MM includes a posteroanterior view of the chest; anteroposterior and lateral views of the cervical spine, thoracic spine, lumbar spine, humeri, and femurs; anteroposterior and lateral views of the skull; and anteroposterior view of the pelvis. Symptomatic areas are also imaged. Conventional skeletal surveys reveal punched-out lytic lesions, diffuse osteopenia, or fractures in nearly 80% of patients with MM at the time of diagnosis.

Focal lytic lesions are found in nearly 60% of patients; osteoporosis, pathological fractures, or compression fractures of the spine each occur in approximately 20%. The most frequent sites of involvement include areas with active hematopoiesis, such as the vertebral bodies, skull, thoracic cage, pelvis, and proximal humeri and femurs.[6]

Low-dose whole-body computed tomography (CT), magnetic resonance imaging (MRI), and positron emission tomography (PET)/CT scans are helpful in patients who have bone pain but no abnormalities on routine roentgenograms (Fig. 80.5). MRI can detect diffuse and focal bone marrow lesions in patients with MM without osteopenia or focal osteolytic lesions on standard metastatic bone surveys, and bone surveys can detect lesions not found on MRI of the axial skeleton. PET/CT scanning using fluorine-18-labeled fluorodeoxyglucose (FDG) appears to correlate with areas of active lytic bone disease.[6]

Diagnosis

The physical examination, laboratory tests, and imaging studies recommended by the IMWG for the diagnosis of MM are outlined in Table 80.2.[8] The 2015 IMWG criteria for the diagnosis of MM require the fulfillment of criteria as outlined in Table 80.3.[8]

KEY CONCEPTS

Multiple Myeloma

- A cancer of plasma cells.
- Organ damage due to plasma cell proliferation and M-protein overproduction.
- Pathophysiology of multiple myeloma consists mainly of CRAB: hyperCalcemia, Renal, Anemia, and Bone.
- History and physical examination, laboratory tests, bone marrow testing, and imaging all play key roles in the diagnosis and management of myeloma.
- Prognosis and treatment vary depending on risk stratification.

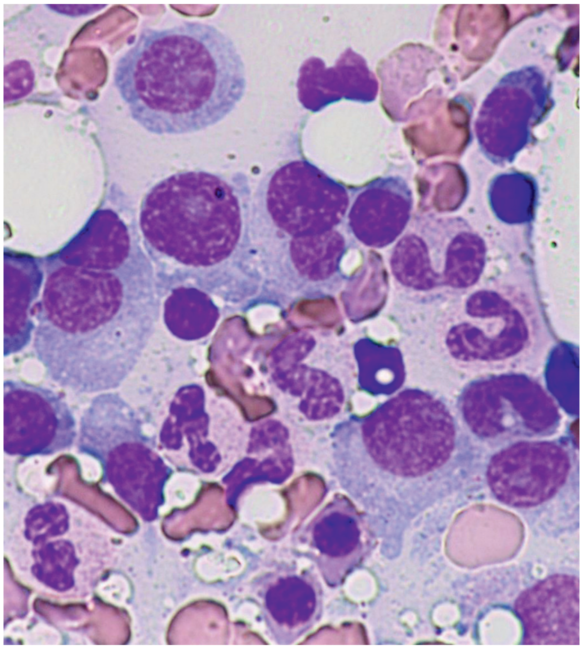


FIG 80.3 Photomicrograph of Plasma Cells in a Bone Marrow Biopsy From a Myeloma Patient. Plasma cells show atypical features, including visible nucleoli, multinucleation, high nuclear/cytoplasmic ratios, and increased size. (Courtesy Dr. Mikhail Roshal, Department of Pathology, Memorial Sloan Kettering Cancer Center.)

Management

Patients who have end-organ damage attributable to the underlying plasma cell disorder have MM and require therapy. The following findings, if attributable to the underlying plasma cell disorder, are clear indications for treatment: anemia, hypercalcemia, renal insufficiency, lytic bone lesions or severe osteopenia, and extramedullary plasmacytoma (excluding SEP).[9] Asymptomatic patients who have one of the following biomarkers are also considered to have MM that requires therapy: ≥60% clonal plasma cells in the bone marrow, involved/uninvolved FRL ratio of ≥100, more than one focal bone lesion on MRI.[9]

High-dose chemotherapy with autologous hematopoietic stem cell transplantation (HSCT) is the preferred therapy for patients with standard- or intermediate-risk myeloma. Patients eligible for HSCT receive induction therapy for 2–4 months prior to stem cell collection to reduce the number of tumor cells in the bone marrow and peripheral blood, to lessen symptoms, and to mitigate end-organ damage.[9] The preferred induction chemotherapy regimen depends on the risk stratification.

Patients ineligible for HSCT receive induction lenalidomide plus dexamethasone. In contrast, those receiving alkylator- or bortezomib-based regimen are treated for approximately 12 to 18 months and then observed until progression.[9] Patients with high-risk MM do poorly with all conventional treatment options. All patients should be evaluated before each treatment cycle to determine the degree of disease response.

Patients with one or more lesions on skeletal radiographs and those with osteopenia should be given bisphosphonate therapy. Pneumococcal and influenza vaccines should also be given to all patients. Patients with MM frequently develop complications related to their disease, including hypercalcemia, renal insufficiency, infection, and skeletal lesions, which require specific treatment.[9]

Almost all patients with MM who survive initial treatment will eventually relapse and require further therapy. Treatment options for patients with relapsed or refractory MM include HSCT, a reinitiation of the previous chemotherapy regimen, or a trial of a new regimen.[9]

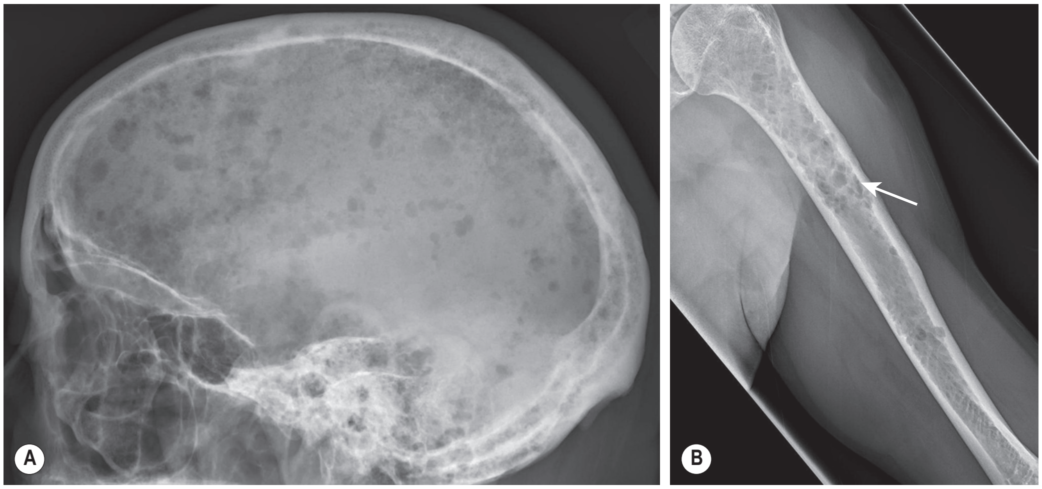


FIG 80.4 Conventional Radiographs Showing Myeloma Osteolytic Lesions in the (A) Skull and (B) Humerus. (Courtesy Dr. Jonathan Landa, Department of Radiology, Memorial Sloan Kettering Cancer Center.)

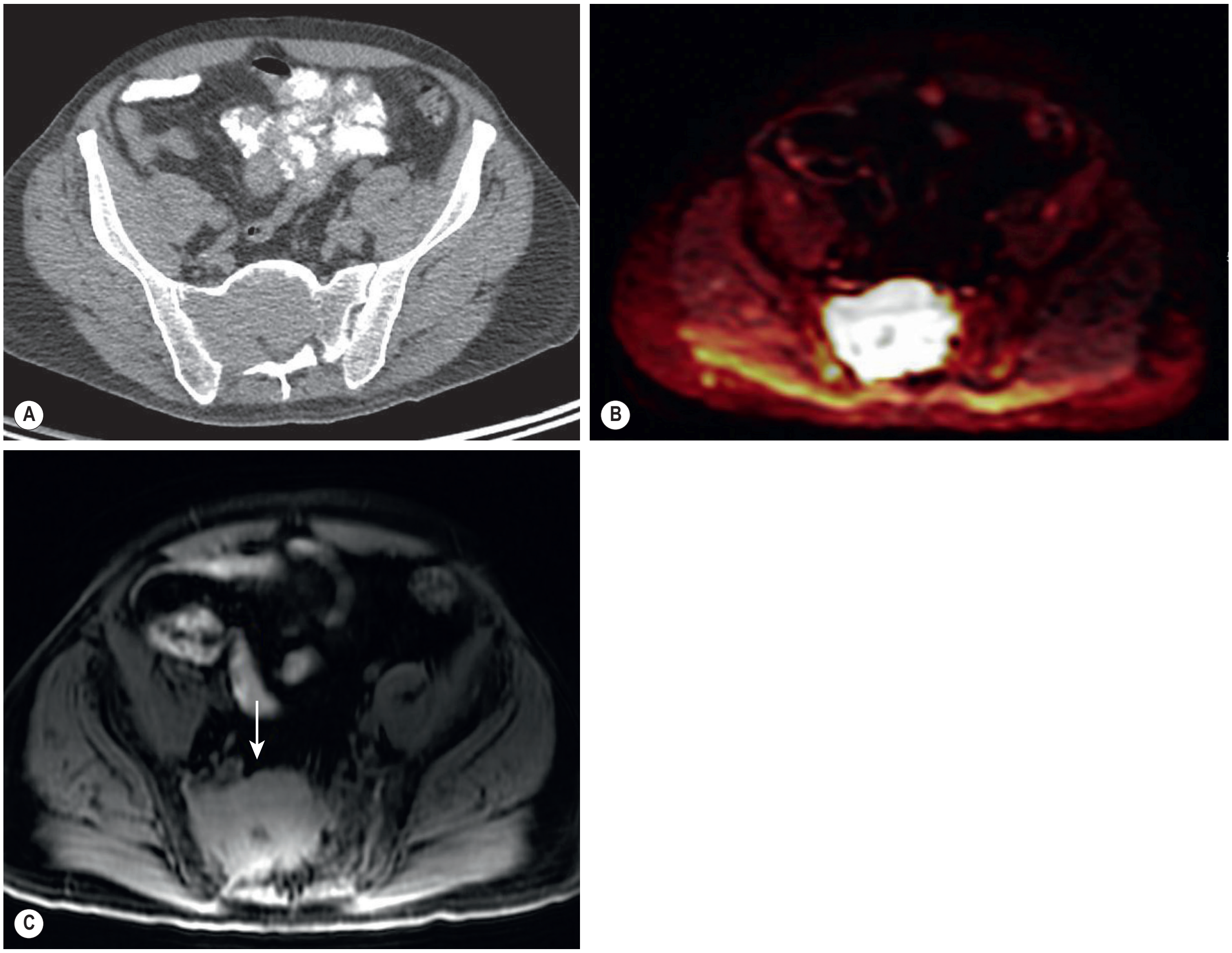


FIG 80.5 Myeloma lytic bone lesions visible in the sacral vertebra by (A) conventional computed tomography (CT), (B) positron emission tomography (PET)/CT scans, and (C) magnetic resonance imaging (MRI) demonstrate hypermetabolism in the sacrum. (Courtesy Dr. Jonathan Landa, Department of Radiology, Memorial Sloan Kettering Cancer Center.)

Prognosis

MM is a heterogeneous disease. Some patients progress rapidly despite treatment, and others remain stable without therapy for a number of years. Myeloma patients are typically staged using the Revised International Staging System (R-ISS). This system stratifies patients with MM into three main risk categories at time of diagnosis.[10]

- High-Risk Myeloma: Patients with t(14;16), t(14;20), or del17p13 by FISH are considered to have high-risk myeloma. These patients have a median survival of approximately 2–3 years despite standard treatment. Patients with lactate dehydrogenase (LDH) ≥2 times the institutional laboratory upper limit of normal and those with features of primary are also considered to have high-risk MM.
- Intermediate-Risk Myeloma: Patients with t(4;14) or 1q+ by FISH or deletion 13/hypodiploidy are considered to have intermediate-risk MM.
- Standard-Risk Myeloma: All other patients with MM who lack any of the high- or intermediate-risk cytogenetic abnormalities or features are considered to have standard-risk MM; this includes patients with trisomies, t(11;14) and t(6;14).

THERAPEUTIC PRINCIPLES

Current Treatment of Multiple Myeloma (MM)

- Patients with standard- or intermediate-risk myeloma: High-dose chemotherapy with autologous hematopoietic stem cell transplantation (HSCT) with induction chemotherapy regimen dependent on risk stratification.
- Patients ineligible for HSCT: induction lenalidomide plus dexamethasone.
- Patients with high-risk MM do poorly with all conventional treatment options.
- Patients should be evaluated before each treatment cycle to determine degree of disease response.
- Patients with one or more lesions on skeletal radiographs and those with osteopenia should be given bisphosphonate therapy.
- Almost all patients with MM who survive initial treatment will eventually relapse and require further therapy.

SMOLDERING MULTIPLE MYELOMA

SMM is defined as follows: M-protein ≥3 g/dL in serum and/or 10–60% bone marrow plasma cells, no end-organ damage (lytic lesions, anemia, renal disease, or hypercalcemia) attributable to

TABLE 80.2 Exams, Tests, and Imaging Studies Recommended by the International Myeloma Working Group for the Diagnosis of Multiple Myeloma

1. History and Physical Examination
2. Routine Testing
   - Complete blood count with differential and peripheral blood smear review.
   - Chemistry panel including calcium and creatinine.
   - Serum protein electrophoresis, immunofixation electrophoresis.
   - Nephelometric quantitation of immunoglobulins.
   - Routine urinalysis, 24-hour urine collection for proteinuria, electrophoresis, and immunofixation electrophoresis.
   - Quantification of both urine M-component level and albuminuria.
3. Bone Marrow Testing: Obtain an aspirate plus trephine biopsy with testing for cytogenetics, fluorescent in situ hybridization (FISH), and immunophenotyping.
4. Imaging
   - Bone survey including spine, pelvis, skull, humeri, and femurs.
   - The IMWG now recommends the use of low-dose whole-body CT (LDWBCT) or MRI in the workup of smoldering multiple myeloma (SMM) and solitary plasmacytoma.
   - The IMWG now recommends that one of PET/CT, LDWBCT, or MRI of the whole body or spine be done in all patients with suspected smoldering myeloma, with the exact imaging modality determined by availability and resources.
   - Clear evidence of one or more sites of osteolytic bone destruction (≥5 mm in size) seen on CT (including LDWBCT) or PET/CT does fulfill the criteria for bone disease in multiple myeloma and should be regarded as meeting the CRAB (hyperCalcemia, Renal, Anemia, and Bone) requirement, whether or not the lesions can be visualized on skeletal radiography.
   - Increased uptake on PET/CT alone is not adequate for the diagnosis of multiple myeloma; evidence of underlying osteolytic bone destruction is needed on the CT portion of the examination.
   - Bone densitometry studies are not sufficient to determine the presence of multiple myeloma.
   - The IMWG no longer recommends the presence of osteoporosis or vertebral compression fractures in the absence of lytic lesions as being sufficient evidence of bone disease for purposes of the diagnostic criteria.

Source: Rajkumar SV, Dimopoulos MA, Palumbo A, Blade J, Merlini G, Mateos MV, et al. International Myeloma Working Group updated criteria for the diagnosis of multiple myeloma. Lancet Oncol 2014;15(12):e538–48. doi: 10.1016/S1470-2045(14)70442-5. PubMed PMID: 25439696.

TABLE 80.3 IMWG Criteria for the Diagnosis of Multiple Myeloma

Clonal bone marrow plasma cells ≥10% or biopsy-proven bony or extramedullary plasmacytoma and any one or more of the following CRAB (hyperCalcemia, Renal, Anemia, and Bone) features and myeloma defining events (MDEs):

1. Evidence of end-organ damage that can be attributed to the underlying plasma cell proliferative disorder, specifically:
   - Hypercalcemia: serum calcium >0.25 mmol/L (>1 mg/dL) higher than the upper limit of normal or >2.75 mmol/L (>11 mg/dL)
   - Renal insufficiency: creatinine clearance <40 mL per minute or serum creatinine >177 µmol/L (<2 g/dL [20 g/L])
   - Anemia: hemoglobin value of >20 g/L below the lowest limit of normal, or a hemoglobin value <10 g/dL (100 g/L)
   - Bone lesions: one or more osteolytic lesions on skeletal radiography, CT, or PET/CT. If bone marrow has <10% clonal plasma cells, more than one bone lesion is required to distinguish from solitary plasmacytoma with minimal marrow involvement
2. Any one or more of the following biomarkers of malignancy (MDEs):
   - 60% or greater clonal plasma cells on bone marrow examination
   - Serum-involved uninvolved free light chain ratio of 100 or greater, provided the absolute level of the involved light chain is at least 100 mg/L (a patient's "involved" free light chain—either kappa or lambda—is the one that is above the normal reference range; the "uninvolved" free light chain is the one that is typically in, or below, the normal range)
   - More than one focal lesion on MRI that is at least 5 mm or greater in size

Source: Rajkumar SV, Dimopoulos MA, Palumbo A, Blade J, Merlini G, Mateos MV, et al. International Myeloma Working Group updated criteria for the diagnosis of multiple myeloma. Lancet Oncol. 2014;15(12):e538–48. doi: 10.1016/S1470-2045(14)70442-5. PubMed PMID: 25439696.

the underlying plasma cell disorder, and no amyloidosis. SMM is differentiated from MGUS by the quantity of M-protein in serum (≥3 g/dL), urine (≥500 mg/24 h), and/or the percentage of clonal bone marrow plasma cells (10–60%).

Prognosis and Management

Patients with SMM will progress to symptomatic MM or AL amyloidosis at an approximate rate of 10%/year for the first 5 years, 3%/year for the next 5 years, and 1–2%/year for the following 10 years. The majority of patients with SMM will progress, with a median time to progression of 4.8 years.[11] The following factors are associated with an increased risk of progression: abnormal free light chain κ/λ ratio (<0.125 or >8.0), bone marrow clonal plasma cells 50–60%, serum M-protein ≥3 g/dL (or progressive increase in M-protein level), IgA SMM, immunoparesis, high-risk genetic features, increased circulating plasma cells, abnormal plasma cell immunophenotype, and abnormalities on imaging (MRI or PET/CT).[11,12] Patients with SMM should not be treated but instead closely monitored with follow-up every 3–4 months with treatment initiated upon progression to MM.[11]

NONSECRETORY MYELOMA

Approximately 3% of patients with MM have no M-protein in the serum or urine based on immunofixation electrophoresis at the time of diagnosis. Patients with myeloma who have normal serum and urine immunofixation electrophoresis as well as a normal serum FLC ratio are considered to have true nonsecretory myeloma. Of these, approximately 85% will have M-protein that can be detected in the cytoplasm of the neoplastic plasma cells by immunohistochemistry; however, they have impaired secretion of this protein. The other 15% do not have immunoglobulin detectable in the plasma cells.[6]

Patients with true nonsecretory myeloma need to be monitored mainly on the basis of imaging tests and bone marrow studies. Patients with nonsecretory MM are not at risk for myeloma kidney as long as light chains cannot be detected in the urine, but they are at risk for other complications of MM.[6]

OLIGOSECRETORY MYELOMA

Approximately 5–10% of patients with MM have oligosecretory myeloma at the time of diagnosis as defined by the absence of measurable disease in serum or urine using the following parameters: serum M-protein <1 g/dL and urine M-protein <200 mg/24 hours.[6]

Monitoring these patients is difficult using the standard serum and urine electrophoretic tests. In most of these patients, the serum FLC assay can be used to monitor the disease, provided that the serum FLC ratio is abnormal and the involved (affected) FLC level is ≥10 mg/dL. Patients with oligosecretory disease may need to be monitored with imaging and bone marrow studies, particularly if the baseline FLC levels are unmeasurable (<10 mg/dL).

PLASMA CELL LEUKEMIA

PCL is a very rare and aggressive variant of MM. It is characterized by high levels of plasma cells circulating in peripheral blood. PCL may originate de novo (primary PCL) or as a secondary transformation of MM (secondary PCL). The incidence of PCL in Europe has been estimated at 4 cases/10 000 000 persons/year. Secondary PCL occurs as a progression of disease in 1–4% of all cases of MM.[13]

Clinical Presentation and Diagnosis

Patients with PCL typically present with signs and symptoms similar to those seen in MM as well as in other leukemias.[13] The diagnosis of PCL is based upon the evaluation of peripheral blood smear, bone marrow aspirate and biopsy, and protein electrophoresis. PCL diagnosis is confirmed when a monoclonal population of plasma cells is present in the peripheral blood with an absolute plasma cell count exceeding 2000/μL and 20% of the peripheral blood white cells.[13]

Management and Prognosis

Treatment generally consists of induction therapy followed by HSCT, if eligible for HSCT, and chemotherapy alone for those ineligible for HSCT.[13] The prognosis of PCL is poor with a median survival of 7–11 months. Median survival for PCL occurring after refractory or relapsing MM is 2–7 months.[13]

SOLITARY PLASMACYTOMA OF BONE

SPB is a localized tumor in the bone comprised of a single clone of plasma cells in the absence of other features of MM. Approximately 5% of all cases of plasma cell disorders are SPB. The incidence of SPB is approximately 0.15 cases/100 000 person-years with approximately 450 new cases per year in the United States. The incidence is highest in patients of African background and lowest in Asians and Pacific Islanders. Men are diagnosed twice as frequently as women. The median age at diagnosis is 55–65 years compared with a median age at diagnosis of 71 years for patients with MM.[14]

Clinical Presentation

Most patients present with skeletal pain or a pathological fracture of the affected bone. Patients with vertebral involvement may have severe back pain or neurological compromise. Less commonly, SPB can extend into the surrounding soft tissue, resulting in a palpable mass. The most common bones involved are those with active hematopoiesis; the axial skeleton is more commonly involved than is the appendicular skeleton, and involvement of the distal appendicular skeleton below the knees or elbows is extremely rare.[14] By definition, patients with SPB do not have anemia or renal insufficiency attributable to the underlying plasma cell disorder.

Diagnosis

The diagnosis of SPB requires the following: biopsy-proven solitary tumor of bone with evidence of clonal plasma cells, metastatic bone survey and either CT (PET/CT) or MRI of the spine and pelvis showing no other lytic lesions, bone marrow aspirate, and biopsy that contains no clonal plasma cells as well as no findings attributable to clonal plasma cell proliferative disorder (anemia, hypercalcemia, or renal insufficiency).[8] The presence of an M-protein does not exclude the diagnosis of SPB, as a low level of M-protein may be present in 30–75% of cases. This M-protein may or may not disappear with treatment.

Management and Prognosis

The primary treatment for patients with SPB is localized radiation therapy (RT). Surgery may be required for patients with structural instability of the bone, retropulsed bone, or rapidly progressive symptoms from cord compression. Bisphosphonates are not recommended for patients with SPB, except in the setting of underlying osteopenia.[14]

The median overall survival of patients with SPB is approximately 10 years. Overall survival rates at 5 and 10 years are approximately 75% and 45%, respectively, with corresponding disease-free survival rates of 45% and 25%. A little more than half of patients with SPB will eventually develop overt MM. Patients with true SPB who meet the strict criteria listed above have a recurrence/progression rate of approximately 10% within 3 years.[14]

SOLITARY EXTRAMEDULLARY PLASMACYTOMA

SEPs are plasma cell tumors that arise outside the bone marrow. SEP refers to a solitary nonosseous plasma cell neoplasm in the absence of any other sign of MM. The SEP lesions are most often located in the head and neck region, mainly in the upper aerodigestive tract, but may also occur in the gastrointestinal tract, urinary bladder, central nervous system, thyroid, breast, testes, parotid gland, lymph nodes, and skin. The median age at diagnosis for SEP is 55 to 60 years; approximately two-thirds of patients are male, and SEP accounts for approximately 3% of plasma cell malignancies.[15]

Clinical Presentation

Most patients present with symptoms related to the location of the mass. Approximately 80% involve the upper respiratory tract (*i.e.*, oronasopharynx and paranasal sinuses) and cause epistaxis, nasal discharge, or nasal obstruction. Less common sites of involvement include the gastrointestinal tract, liver, lymph nodes, testes, skin, and central nervous system. Primary plasmacytoma of the lung often presents as a pulmonary nodule or hilar mass with or without hemoptysis.[15]

Diagnosis

By definition, patients with SEP do not have anemia, hypercalcemia, renal insufficiency, or bone lesions attributable to the underlying plasma cell disorder. The diagnosis of SEP requires the following: biopsy-proven extramedullary tumor with evidence of clonal plasma cells, metastatic bone survey and either MRI or PET/CT of the spine and pelvis that show no lytic lesions, bone marrow aspirate and biopsy without clonal plasma cells, and absence of anemia, hypercalcemia, or renal insufficiency attributable to a clonal plasma cell proliferative disorder.[16]

Some patients may have a small amount of M-protein, usually IgA, in the serum or urine, which often disappears after treatment. The bone marrow of patients with SEP does not have clonal plasma cells, whereas some patients with a solitary extramedullary lesion may demonstrate up to 10% clonal plasma cells, in which case the diagnosis is SEP with minimal marrow involvement.[15]

Management and Prognosis

The treatment of choice for SEP is RT. Small lesions may be cured with surgery alone; no adjuvant RT is indicated unless there is suspicion of residual local disease.[15]

Less than 7% of patients with SEP will develop a local recurrence after tumoricidal radiation. Approximately 10–15% of patients ultimately will develop MM; the progression rate is higher (20%) in patients with SEP with minimal marrow involvement. Five-year overall survival rates range from 40–85%.[15]

POEMS SYNDROME

POEMS (polyneuropathy, organomegaly, endocrinopathy, monoclonal protein, skin changes) syndrome is characterized by the presence of a monoclonal plasma cell disorder, peripheral neuropathy, and one or more of the following features: osteosclerotic myeloma, Castleman disease, increased levels of serum vascular endothelial growth factor (VEGF), organomegaly, endocrinopathy, edema, typical skin changes, and papilledema. The cause of POEMS syndrome is unknown, although chronic overproduction of proinflammatory and other cytokines such as VEGF appears to be a major feature of this disorder. POEMS syndrome is a rare disorder, but the exact incidence is unknown. POEMS syndrome commonly presents in the fifth to sixth decade.[17]

Clinical Presentation

By definition, all patients have peripheral neuropathy and a monoclonal plasma cell disorder, almost always of the lambda light chain type. In addition, almost all patients have osteosclerotic lesions, whereas Castleman disease is present in approximately 15%.[17]

Diagnosis

The IMWG has developed a series of criteria required for the diagnosis of POEMS syndrome; other signs and symptoms that do not fall into these criteria may also be present, as outlined in Table 80.4.[8]

TABLE 80.4 IMWG Diagnostic Criteria for the Diagnosis of POEMS Syndrome

Both Mandatory Criteria:

1. Polyneuropathy
2. Monoclonal plasma cell proliferative disorder

Plus at Least One Major Criterion:

3. Osteosclerotic or mixed sclerotic/lytic lesion visualized on plain films or computed tomography (CT)
4. Castleman disease
5. Elevated serum or plasma vascular endothelial growth factor (VEGF) levels (at least three to four times the upper limit of normal)

Plus at Least One Minor Criterion:

1. Organomegaly (splenomegaly, hepatomegaly, or lymphadenopathy)
2. Extravascular volume overload (peripheral edema, ascites, or pleural effusion)
3. Endocrinopathy (adrenal, thyroid, pituitary, gonadal, parathyroid, or pancreatic disorder excluding diabetes mellitus or hypothyroidism)
4. Skin changes (hyperpigmentation, hypertrichosis, glomeruloid hemangiomata, plethora, acrocyanosis, flushing, white nails)
5. Papilledema
6. Thrombocytosis or polycythemia

POEMS, polyneuropathy, organomegaly, endocrinopathy, monoclonal protein, skin changes.

Source: Rajkumar SV, Dimopoulos MA, Palumbo A, Blade J, Merlini G, Mateos MV, et al. International Myeloma Working Group updated criteria for the diagnosis of multiple myeloma. Lancet Oncol 2014;15(12):e538–48. doi: 10.1016/S1470-2045(14)70442-5. PubMed PMID: 25439696.

Management and Prognosis

There is no standard treatment for POEMS syndrome. Those with limited disease may be treated with RT. Patients with widespread bone lesions may be treated with therapy similar to that used for MM.[17] POEMS syndrome is chronic, and patients with POEMS syndrome survive three times longer than those with MM. The natural history is one of progressive peripheral neuropathy until the patient is bedridden. Death usually occurs from inanition or a terminal bronchopneumonia.[17]

WALDENSTRÖM MACROGLOBULINEMIA

WM is a rare clinicopathological entity demonstrating 10% or greater infiltration of the bone marrow by clonal lymphoplasmacytic cells and a monoclonal IgM gammopathy in the blood. WM is a rare disorder with an incidence of approximately 3/million people/year, with 1400 new cases diagnosed in the United States each year. The median age at diagnosis is 64 years; less than 1% of patients are diagnosed under 40 years of age, and approximately 60% are males. WM is more common in Caucasians than in other ethnic groups, and familial predisposition is present in up to 20%.[18]

Clinical Presentation

Patients usually present in their seventh decade with symptoms related to the infiltration of the hematopoietic tissues or to the effects of monoclonal IgM in the blood. WM most commonly presents with pallor, oronasal bleeding, systemic complaints (weakness, fatigue, weight loss, fever, night sweats), and organomegaly (enlarged lymph nodes, spleen, and/or liver). Involvement of the bone or kidneys is uncommon.[18]

An important presentation includes central nervous system signs and symptoms due to the hyperviscosity syndrome (blurring or loss of vision, headache, ataxia, dementia, stroke, or coma). A classic finding in WM associated with hyperviscosity is the presence of dilated, segmented, and tortuous retinal veins. Another major presentation involves neurological symptoms such as paresthesias and weakness as well as cranial nerve palsies and sudden deafness. Patients with WM who are asymptomatic are considered to have SWM.[18]

Laboratory Findings and Diagnosis

Patients with WM must have 10% or greater infiltration of the bone marrow by clonal lymphoplasmacytic cells and a monoclonal

IgM gammopathy in the blood. The diagnosis of WM is made when the following two criteria are met: (i) presence of an IgM monoclonal paraprotein on serum immunofixation and (ii) 10% or more of the bone marrow biopsy sample demonstrates infiltration by small lymphocytes that exhibit plasmacytoid or plasma cell differentiation (lymphoplasmacytic features or lymphoplasmacytic lymphoma) with an intertrabecular pattern. This infiltrate should express a typical immunophenotype (*e.g.*, surface IgM^{+}, CD5$^{+/-}$, CD10^{-}, CD19^{+}, CD20^{+}, CD22^{+}, CD23^{-}, CD25^{+}, CD27^{+}, FMC7^{+}, CD103^{-}, CD138^{-}).[19] WM must be differentiated from IgM MGUS, MM, chronic lymphocytic leukemia (CLL), and mantle-cell lymphoma.

Management and Prognosis

Many patients with WM are asymptomatic and can be observed for months to years after the diagnosis is established before requiring treatment. Asymptomatic patients with adequate hemoglobin and platelet levels are followed every 6 months with complete blood counts and monoclonal protein levels.[20] Treatment is indicated in those who meet one or more of the conditions presented in Table 80.5.[21]

Symptomatic WM is treated with a regimen that incorporates rituximab in combination with other agents. Symptomatic patients with low tumor burden and minimally symptomatic patients may be treated with the single agent rituximab. Patients who present with symptoms due to hyperviscosity should undergo immediate therapeutic plasmapheresis, followed by chemotherapy for control of the malignant clone.[21] Median survival in WM is approximately 5–8 years from the time of diagnosis. However, outcomes are highly variable. A staging system for WM has been developed based on a prospective multicenter observational study outlined in Table 80.6.[22]

TABLE 80.5 Clinical and Laboratory Indications for Initiation of Treatment for Waldenström Macroglobulinemia (WM)

Clinical Indications for Therapy Initiation:

- Recurrent fever, night sweats, weight loss, fatigue
- Hyperviscosity
- Lymphadenopathy that is either symptomatic or bulky (≥5 cm in maximum diameter)
- Symptomatic hepatomegaly and/or splenomegaly
- Symptomatic organomegaly and/or organ or tissue infiltration
- Peripheral neuropathy due to WM

Laboratory Indications for Therapy Initiation:

- Symptomatic cryoglobulinemia
- Cold agglutinin anemia
- Immune hemolytic anemia and/or thrombocytopenia
- Nephropathy related to WM
- Amyloidosis related to WM
- Hemoglobin ≤10 g/dL
- Platelet count <100 × 10^{9}/L

Source: Dimopoulos MA, Kastritis E, Owen RG, Kyle RA, Landgren O, Morra E, et al. Treatment recommendations for patients with Waldenstrom macroglobulinemia (WM) and related disorders: IWWM-7 consensus. Blood 2014;124(9):1404–11. doi: 10.1182/blood-2014-03-565135. PubMed PMID: 25027391; PubMed Central PMCID: PMC4148763.

HEAVY CHAIN DISEASES

The HCDs are rare B-cell proliferative disorders characterized by the production of an M-protein consisting of a portion of the immunoglobulin heavy chain without a bound light chain. Three types of HCD are recognized, based upon the class of immunoglobulin heavy chain produced by the malignant cell:

1. Alpha HCD is a form of mucosa-associated lymphoreticular tissue (MALT) lymphoma that is also called immunoproliferative small intestinal disease (IPSID), Mediterranean lymphoma, or Seligmann disease.
2. Gamma HCD (Franklin disease) is typically associated with the presence of a systemic lymphoma, often of mixed lymphoid-plasmacytic character.
3. Mu HCD has clinical features resembling small lymphocytic lymphoma/CLL, often with distinctive vacuolated lymphocytes/plasma cells in the bone marrow.

Clinical Presentation

The clinical presentation of the HCDs is that of a patient with a low-grade B-cell malignancy. Alpha HCD is a form of MALT lymphoma, with the same histological features of MALT-type gastrointestinal lymphomas with marked plasma cell differentiation. It is the most common form of HCD and occurs in patients from the Mediterranean region or Middle East, usually young males, and is often associated with relatively poor sanitation. The gastrointestinal tract is most commonly involved in alpha HCD, resulting in abdominal pain, malabsorption with chronic diarrhea, steatorrhea, and loss of weight. Growth retardation, digital clubbing, and mesenteric lymphadenopathy may also be present.[23]

Patients with gamma HCD typically present with systemic symptoms, lymphadenopathy, splenomegaly, and/or anemia, and occasionally with palatal and uvular swelling. The median age of patients with gamma HCD is 60 to 70 years, although the condition has been noted in persons younger than age 20. Autoimmune manifestations are seen in about one-third of patients.[23]

Mu HCD is the least common of the HCDs and has features resembling CLL/small lymphocytic lymphoma, although peripheral adenopathy is less common than in CLL. Osteolytic lesions or pathological fractures occasionally have been reported in patients. Unlike those with alpha and gamma HCD, some patients

TABLE 80.6 Staging System for Waldenström Macroglobulinemia

Stage	Risk	Characteristics	Five-Year Survival	Five-Year Progression-Free Survival
A	Low	B2M <3 mg/L and Hgb ≥12.0 g/dL	87%	83%
B	Medium	B2M <3 mg/L and Hgb <12.0 g/dL	63%	55%
C	Medium	B2M ≥3 mg/L and serum IgM ≥4.0 g/dL	53%	33%
D	High	B2M ≥3 mg/L and IgM <4.0 g/dL	21%	12%

B2M, Beta 2 microglobulin.

Source: Dhodapkar MV, Jacobson JL, Gertz MA, Rivkin SE, Roodman GD, Tuscano JM, et al. Prognostic factors and response to fludarabine therapy in patients with Waldenstrom macroglobulinemia: results of United States intergroup trial (Southwest Oncology Group S9003). Blood 2001;98(1):41–8. PubMed PMID: 11418461.

with mu HCD have increased free light chain secretion, and they may develop cast nephropathy or amyloidosis.[23]

Diagnosis

The diagnosis of the HCDs can be made on histopathological examination of affected tissues and/or analysis of serum or urine samples by electrophoresis and immunofixation electrophoresis. A typical M-spike in the serum or urine may not always be present. On serum or urine immunofixation electrophoresis or immunohistochemical staining of affected tissues, the abnormal protein consists of a heavy chain without an associated light chain. Immunohistochemical staining of affected tissue shows the presence of a clonal population of cells staining positively for a heavy chain but negative for both kappa and lambda light chains.[23]

Alpha HCD most commonly involves infiltration of the jejunal mucosa with plasmacytoid cells, whereas the duodenum and ileum are affected less often. The diagnosis depends upon the recognition of a monoclonal alpha heavy chain without an associated light chain in the serum, urine, intestinal secretions, or the cells infiltrating the intestinal mucosa. The serum protein electrophoretic pattern is normal in one-half of cases, and in the remainder an unimpressive broad band may appear in the alpha-2 or beta mobility regions. The amount of alpha heavy chain in the urine is small.[23]

Gamma HCD is typically associated with a polymorphous infiltrate, including admixtures of lymphocytes, plasmacytoid lymphocytes, plasma cells, immunoblasts, and eosinophils. The electrophoretic pattern often shows a broad-based band more suggestive of a polyclonal increase, but the gamma chains are monoclonal. The diagnosis is established by immunofixation electrophoresis, which shows that the abnormal protein consists only of gamma heavy chains without associated light chains.[23]

Mu HCD is characterized by vacuolated plasma cells or lymphoid cells observed in the bone marrow, together with panhypogammaglobulinemia. The serum protein electrophoretic pattern is usually normal, except for hypogammaglobulinemia.[23]

Management and Prognosis

Initial treatment of alpha HCD consists of eradication of any concurrent infection with appropriate antibiotics. Those with symptomatic disease who do not respond adequately to antibiotics may be treated with chemotherapy similar to that used for non-Hodgkin lymphoma, such as the CHOP regimen.[23] Treatment of gamma HCD is indicated only for symptomatic patients. Appropriate chemotherapy includes melphalan plus prednisone or regimens used to treat B-cell lymphoma, such as the combination of cyclophosphamide, vincristine, and prednisone, with or without doxorubicin, or rituximab if the abnormal cells express CD20.[23] Treatment for symptomatic patients with mu HCD is similar to that employed in patients with CLL (*e.g.*, glucocorticoids, alkylating agents, fludarabine).[23]

In the absence of therapy, alpha HCD typically is progressive and fatal. The prognosis of gamma HCD is variable and ranges from the asymptomatic presence of a stable monoclonal heavy chain in the serum or urine (*e.g.*, MGUS) to a rapidly progressive downhill course over a short duration (*e.g.*, highly aggressive lymphoma). Median survival is 7.4 years with an extremely wide range from 1 month to more than 21 years.[23] The course of mu HCD is variable, and survival ranges from a few months to many years.[23]

LIGHT CHAIN AMYLOIDOSIS

Amyloidosis refers to the deposition of fibrils composed of low-molecular-weight subunits of normal serum proteins. AL amyloidosis is a specific form of amyloidosis caused by the deposition of monoclonal light chains. These light chains undergo transformation to a beta-pleated fibrillar configuration. Deposition of these light chains in various organs results in organ failure.

AL amyloidosis is an uncommon disorder of older adults, and the exact incidence is unknown. In the United States, the incidence is approximately 6–10 cases per million person-years. Age-specific incidence rates increase in each decade of life after age 40 years. The median age at diagnosis is 64 years, and less than 5% of patients are under the age of 40. There is a male predominance, with men accounting for 65–70% of patients. AL amyloidosis occurs in all races and all geographic locations.[24]

Clinical Presentation

AL amyloidosis is a systemic disorder that can present with a variety of symptoms, including significant proteinuria, edema, hepatosplenomegaly, otherwise unexplained heart failure, and carpal tunnel syndrome. Although virtually all patients have multisystem amyloid deposition, it is not uncommon for a patient to present with evidence of mainly one organ involvement.

Nonspecific systemic symptoms, including fatigue and unintentional weight loss, are common in patients with AL amyloidosis. Renal involvement occurs in approximately 70% of patients and most often presents as asymptomatic proteinuria or nephrotic syndrome. Cardiac involvement is seen in approximately 60% of patients and typically is characterized by thickening of the interventricular septum and ventricular wall. This can lead to systolic or diastolic dysfunction with symptoms of heart failure. Other manifestations that may be seen include sudden death or syncope due to arrhythmia or heart block and, rarely, angina or infarction due to accumulation of amyloid in the coronary arteries.[24]

Hepatomegaly with or without splenomegaly is seen in as many as 70% of patients. Potential gastrointestinal manifestations include bleeding, gastroparesis, constipation, bacterial overgrowth, malabsorption, and intestinal pseudo-obstruction resulting from dysmotility.[24]

Mixed sensory and motor peripheral neuropathy and/or autonomic neuropathy is a prominent feature in AL amyloidosis. Symptoms of numbness, paresthesia, and pain are frequently noted, as in peripheral neuropathy of many other causes. Compression of peripheral nerves, especially the median nerve within the carpal tunnel, can cause more localized sensory changes. Symptoms of bowel or bladder dysfunction and findings of orthostatic hypotension may be due to autonomic nervous system damage.[24]

Amyloid infiltration of skeletal muscles may cause visible enlargement. Macroglossia, or lateral scalloping of the tongue from impingement on the teeth, is characteristic of AL amyloid. Arthropathy may be due to amyloid deposition in joints and surrounding structures. The “shoulder pad“ sign is visible enlargement of the anterior shoulder due to fluid in the glenohumeral joint and/or amyloid infiltration of the synovial membrane and surrounding structures.[24]

Purpura, characteristically elicited in a periorbital distribution (raccoon eyes) by a Valsalva maneuver or minor trauma, is present in only a minority of patients, but it is highly characteristic of AL amyloidosis. Other signs of skin involvement include waxy thickening, easy bruising, and subcutaneous nodules or plaques. Amyloidosis may also be directly associated with a bleeding diathesis. Approximately 10% of patients have coexisting MM.[24]

KEY CONCEPTS

Light Chain Amyloidosis

- A specific form of amyloidosis caused by tissue deposition of monoclonal light chains.
- A systemic disorder with multisystem amyloid deposition causing a variety of nonspecific symptoms.
- Demonstration of amyloid deposition in affected organs is key to diagnosis.
- Poor long-term prognosis due to organ dysfunction caused by amyloid deposition.

Laboratory Findings and Diagnosis

Demonstration of amyloid fibrils upon histological evaluation of an affected organ or a surrogate site is required for the diagnosis of AL amyloidosis. Kidney or liver biopsy is positive for amyloid deposition in over 90% of cases; however, a high success rate can also be achieved by less invasive procedures, such as abdominal fat pad aspirate, rectal biopsy, bone marrow biopsy, or skin biopsy.[24]

Amyloid appears as a pink, amorphous, waxy substance with a characteristic "cracking" artifact on hematoxylin and eosin–stained biopsies (Fig. 80.6a). The presence of amyloid fibrils can be confirmed by their characteristic appearance on electron microscopy and by their ability to bind Congo red or thioflavin-T (Fig. 80.6b and 80.6c).[24]

Staining for kappa and lambda light chains, transthyretin, and serum amyloid A can determine the type of amyloidosis. Positive staining for kappa or lambda indicates AL amyloidosis.[24] Immunofluorescence microscopy using light chain–specific antisera and other specialized techniques may show deposition of a monoclonal (lambda or kappa) light chain in the kidney or other affected tissues. Laser microdissection of tissue followed by analysis by tandem mass spectrometric proteomic analysis

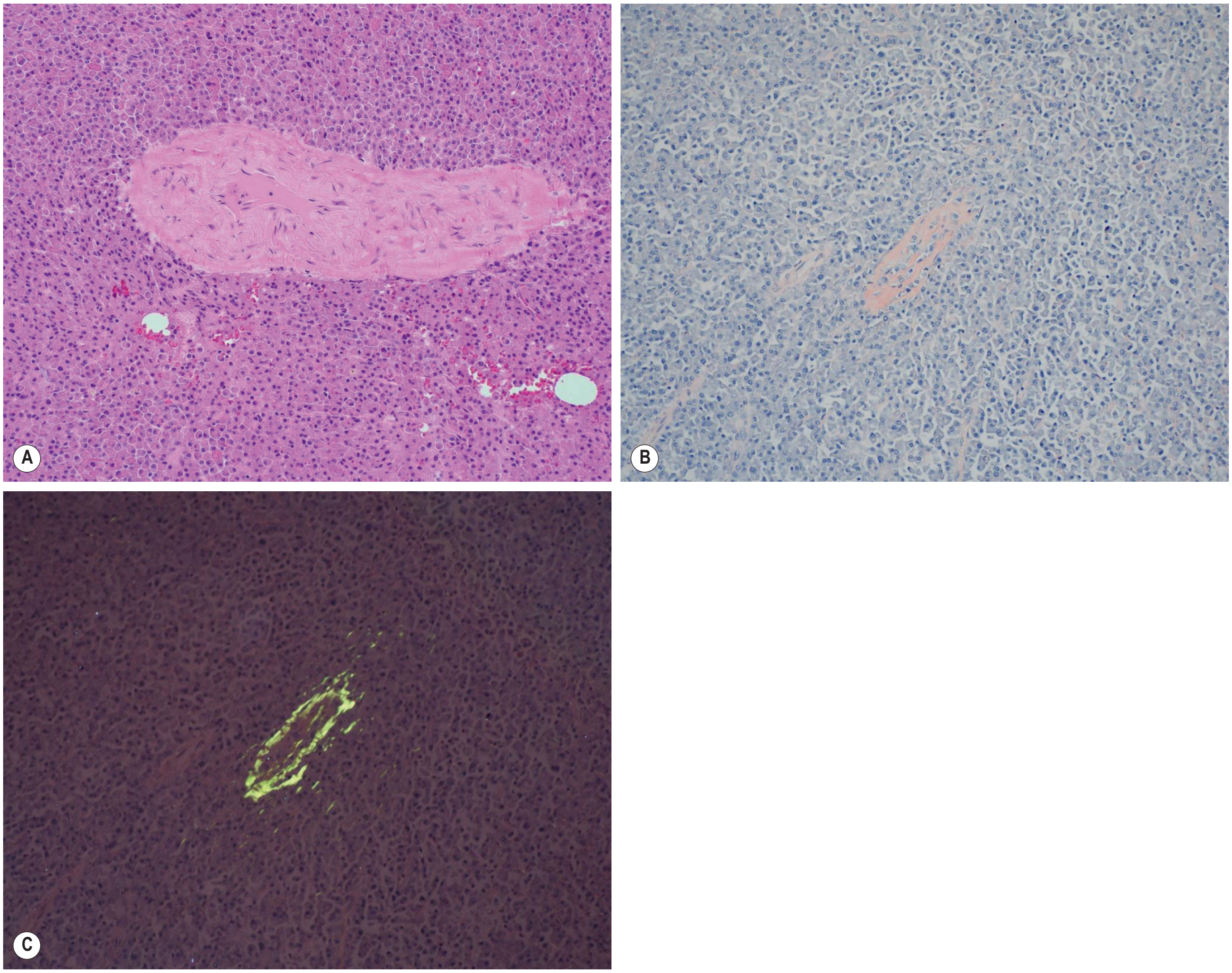


FIG 80.6 **Photomicrograph of Plasmacytoma With AL Amyloid Deposition.** (A) Hematoxylin and eosin stain, (B) Congo red staining without polarization, (C) Congo red staining under polarized light. (Courtesy Dr. Mikhail Roshal, Department of Pathology, Memorial Sloan Kettering Cancer Center.)

may be helpful in identifying the type of amyloid protein (*e.g.*, monoclonal light chain versus other forms of amyloid proteins). An alternative to mass spectrometry is to use immunoelectron microscopy, if available.[24]

Patients with AL amyloidosis frequently have chromosomal abnormalities, but there is no single chromosomal change that is diagnostic of this disease. Chromosomal abnormalities reported with moderate to high frequency in patients with AL amyloidosis include t(11;14)(q13;q32), del(13q14), and gain of 1q21.[24]

Most patients with AL amyloidosis have little or no intact monoclonal immunoglobulin, but are characterized by the presence of monoclonal free light chain. The monoclonal light chain type is lambda in approximately 70% of cases, kappa in 25%, and biclonal in 5%. SPEP demonstrates a localized band or peak in less than 50% of patients with AL amyloidosis. Immunofixation electrophoresis detects a serum or urinary monoclonal protein in nearly 90% of cases of AL amyloidosis. When serum and urine immunofixation electrophoresis is combined with serum free light chain ratio analysis, a monoclonal protein can be detected in virtually all cases.[24]

Bone marrow biopsy specimens typically demonstrate a slightly increased percentage of plasma cells that may appear morphologically normal. Less commonly, the bone marrow demonstrates overt myeloma or lymphoplasmacytic lymphoma. A clonal excess of plasma cells (lambda or kappa) can also be demonstrated by immunoperoxidase staining or flow cytometric analysis of specimens of involved bone marrow.[24] Diagnostic criteria for AL amyloidosis have been developed by the Mayo Clinic and the IMWG that require the presence of the four criteria outlined in Table 80.7.[8] Approximately 2–3% of patients with AL amyloidosis will not meet the requirement for evidence of a monoclonal plasma cell disorder listed.

Management and Prognosis

Amyloidosis patients meeting the criteria in Table 80.8 are considered to be eligible for autologous HSCT.[25] Those eligible for HSCT may be treated with melphalan followed by autologous HSCT. Those not eligible for HSCT may be treated with melphalan/dexamethasone or a bortezomib-based regimen.[25]

AL amyloidosis typically has poor long-term prognosis. Death is typically due to organ dysfunction, with major causes of death due to cardiac or hepatic failure as well as infection. There is, however, significant variability in median survival, depending on the nature, number, and extent of organ involvement. Median survival may range from 4 to 6 months in patients diagnosed at an advanced stage, to in excess of 5 years in patients with limited organ involvement. Two key prognostic factors include the presence and severity of cardiac involvement and the presence of concurrent myeloma.[26]

TABLE 80.7 IMWG Diagnostic Criteria for the Diagnosis of Light Chain Amyloidosis

1. Presence of an amyloid-related systemic syndrome (*e.g.*, renal, liver, heart, gastrointestinal tract, or peripheral nerve involvement). To be included as a diagnostic criterion, the organ damage must be felt to be related to amyloid deposition and not to another common disease, such as diabetes or hypertension.
2. Positive amyloid staining by Congo red in any tissue (*e.g.*, fat aspirate, bone marrow, or organ biopsy) or the presence of amyloid fibrils on electron microscopy.
3. Evidence that the amyloid is light chain–related is established by direct examination of the amyloid using spectrometry-based proteomic analysis or immunoelectron microscopy.
4. Evidence of a monoclonal plasma cell proliferative disorder (*e.g.*, presence of a serum or urine M-protein, abnormal serum free light chain ratio, or clonal plasma cells in the bone marrow).

Source: Rajkumar SV, Dimopoulos MA, Palumbo A, Blade J, Merlini G, Mateos MV, et al. International Myeloma Working Group updated criteria for the diagnosis of multiple myeloma. Lancet Oncol 2014;15(12):e538–48. doi: 10.1016/S1470-2045(14)70442-5. PubMed PMID: 25439696.

MONOCLONAL IMMUNOGLOBULIN DEPOSITION DISEASES

Monoclonal immunoglobulin deposition diseases are a nonamyloid monoclonal immunoglobulin chain deposition disease that is caused by a clonal plasma cell proliferative disorder. These diseases are characterized by the deposition of immunoglobulin light or heavy chain fragments, leading to organ dysfunction. There are two types of immunoglobulin deposition diseases: heavy chain deposition disease (HCDD) and LCDD. They are differentiated from light chain and heavy chain amyloidosis in that the light chain fragments do not have the necessary biochemical characteristics to form amyloid fibrils.

Clinical Presentation

HCDD is extremely rare, with 37 cases reported in the literature from 1992 to 2011. The majority of patients present with hypertension, progressive renal dysfunction, anemia, and nephrotic syndrome with microhematuria.[27] Patients with LCDD usually present with renal, cardiac, or hepatic involvement and may have underlying MM or a lymphoproliferative disorder.[24]

Laboratory Findings

In LCDD, a population of clonal plasma cells produces monoclonal light chain fragments that are deposited as granules in the tissues. The deposits in LCDD are almost always composed of kappa light chains; they are granular, not fibrillar, and do not bind Congo red, thioflavin-T, or serum amyloid P component. Tissue deposits of fragments of monoclonal light chains are commonly seen and can occur in a variety of organs, including the kidney, heart, liver, and small intestine. Routine electrophoretic techniques may not demonstrate a monoclonal protein in the serum or urine in some patients, but it may be detectable by serum free light chain analysis.[26]

TABLE 80.8 Eligibility Criteria for Autologous HSCT in Light Chain Amyloidosis Patients

1. Physiological age ≤70
2. Troponin T <0.06 ng/mL
3. N-terminal pro-brain natriuretic peptide (NT proBNP) <5000 ng/L
4. Creatinine clearance ≥30 mL/min (unless on chronic stable dialysis)
5. Eastern Cooperative Oncology Group (ECOG) performance status ≤2
6. New York Heart Association functional status Class I or II
7. No more than two organs significantly involved (liver, heart, kidney, or autonomic nerve)
8. No large pleural effusions
9. No dependency on oxygen therapy

HSCT, hematopoietic stem cell transplantation.
Source: Gertz MA. How to manage primary amyloidosis. Leukemia. 2012;26(2):191–8. doi: 10.1038/leu.2011.219. PubMed PMID: 21869840.

In HCDD, a population of clonal plasma cells produces both light and heavy monoclonal immunoglobulin chains or only short (truncated) heavy chains that are unable to form amyloid fibrils but are deposited as granules in the tissues. HCDD can be identified by immunofluorescence staining of biopsy tissue with anti–heavy chain antibodies.[26]

Management and Prognosis

The therapeutic approaches for both LCDD and HCDD are similar to those employed for patients with MM. Serial free light chain assays can be used to monitor response to treatment in LCDD.[26] The prognoses of LCDD and HCDD vary considerably, depending on the nature, number, and extent of organ involvement.[26]

ON THE HORIZON

- Two monoclonal antibody drugs (mAb) have recently been approved by the FDA for the treatment of multiple myeloma, with more on the way.
- Caution is warranted when monitoring myeloma patients on mAb drugs, as these biologicals may cause false positive results by protein electrophoresis of the serum and immunofixation electrophoresis (IFE).

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Rajkumar SV. Diagnosis of monoclonal gammopathy of undetermined significance. UpToDate: Wolters Kluwer; 2016.
2. Kyle RA, Therneau TM, Rajkumar SV, et al. Prevalence of monoclonal gammopathy of undetermined significance. N Engl J Med 2006;354(13):1362–9. doi:10.1056/NEJMoa054494. PubMed PMID: 16571879.
3. International Myeloma Working Group. Criteria for the classification of monoclonal gammopathies, multiple myeloma and related disorders: a report of the International Myeloma Working Group. Br J Haematol 2003;121(5):749–57. PubMed PMID: 12780789.
4. Rajkumar SV. Clinical course and management of monoclonal gammopathy of undetermined significance. UpToDate: Wolters Kluwer; 2016.
5. Rajkumar SV, Kyle RA, Therneau TM, et al. Serum free light chain ratio is an independent risk factor for progression in monoclonal gammopathy of undetermined significance. Blood 2005;106(3):812–17. doi:10.1182/blood-2005-03-1038. PubMed PMID: 15855274, PubMed Central PMCID: PMC1895159.
6. Rajkumar SV. Clinical features, laboratory manifestations, and diagnosis of multiple myeloma. UpToDate: Wolters Kluwer; 2016.
7. Kyle RA, Gertz MA, Witzig TE, et al. Review of 1027 patients with newly diagnosed multiple myeloma. Mayo Clin Proc 2003;78(1):21–33. doi:10.4065/78.1.21. PubMed PMID: 12528874.
8. Rajkumar SV, Dimopoulos MA, Palumbo A, et al. International Myeloma Working Group updated criteria for the diagnosis of multiple myeloma. Lancet Oncol 2014;15(12):e538–48. doi:10.1016/S1470-2045(14)70442-5. PubMed PMID: 25439696.
9. Rajkumar SV. Overview of the management of multiple myeloma. UpToDate: Wolters Kluwer; 2016.
10. Palumbo A, Avet-Loiseau H, Oliva S, et al. Revised International Staging System for Multiple Myeloma: A Report From International Myeloma Working Group. J Clin Oncol 2015;33(26):2863–9. doi:10.1200/JCO.2015.61.2267. PubMed PMID: 26240224, PubMed Central PMCID: PMC4846284.
11. Rajkumar SV. Smoldering multiple myeloma. UpToDate: Wolters Kluwer; 2016.
12. Kristinsson SY, Holmberg E, Blimark C. Treatment for high-risk smoldering myeloma. N Engl J Med 2013;369(18):1762–3. doi:10.1056/NEJMc1310911#SA1. PubMed PMID: 24171527.
13. Rajkumar SV. Plasma cell leukemia. UpToDate: Wolters Kluwer; 2016.
14. Rajkumar SV. Diagnosis and management of solitary plasmacytoma of bone. UpToDate: Wolters Kluwer; 2016.
15. Rajkumar SV. Diagnosis and management of solitary extramedullary plasmacytoma. UpToDate: Wolters Kluwer; 2016.
16. Kyle RA, Rajkumar SV. Criteria for diagnosis, staging, risk stratification and response assessment of multiple myeloma. Leukemia 2009;23(1):3–9. doi:10.1038/leu.2008.291. PubMed PMID: 18971951, PubMed Central PMCID: PMC2627786.
17. Rajkumar SV. POEMS syndrome. UpToDate: Wolters Kluwer; 2016.
18. Rajkumar SV. Epidemiology, pathogenesis, clinical manifestations and diagnosis of Waldenstrom macroglobulinemia. UpToDate: Wolters Kluwer; 2016.
19. Owen RG, Treon SP, Al-Katib A, et al. Clinicopathological definition of Waldenstrom's macroglobulinemia: consensus panel recommendations from the Second International Workshop on Waldenstrom's Macroglobulinemia. Semin Oncol 2003;30(2):110–15. doi:10.1053/sonc.2003.50082. PubMed PMID: 12720118.
20. Rajkumar SV. Treatment and prognosis of Waldenstrom macroglobulinemia. UpToDate: Wolters Kluwer; 2016.
21. Dimopoulos MA, Kastritis E, Owen RG, et al. Treatment recommendations for patients with Waldenstrom macroglobulinemia (WM) and related disorders: IWWM-7 consensus. Blood 2014;124(9):1404–11. doi:10.1182/blood-2014-03-565135. PubMed PMID: 25027391, PubMed Central PMCID: PMC4148763.
22. Dhodapkar MV, Jacobson JL, Gertz MA, et al. Prognostic factors and response to fludarabine therapy in patients with Waldenstrom macroglobulinemia: results of United States intergroup trial (Southwest Oncology Group S9003). Blood 2001;98(1):41–8. PubMed PMID: 11418461.
23. Rajkumar SV. The heavy chain diseases. UpToDate: Wolters Kluwer; 2016.
24. Rajkumar SV. Clinical presentation, laboratory manifestations, and diagnosis of immunoglobulin light chain (AL) amyloidosis (primary amyloidosis). UpToDate: Wolters Kluwer; 2016.
25. Gertz MA. How to manage primary amyloidosis. Leukemia 2012;26(2):191–8. doi:10.1038/leu.2011.219. PubMed PMID: 21869840.
26. Rajkumar SV. Prognosis and treatment of immunoglobulin light chain (AL) amyloidosis and light and heavy chain deposition diseases. UpToDate: Wolters Kluwer; 2016.
27. Oe Y, Soma J, Sato H, et al. Heavy chain deposition disease: an overview. Clin Exp Nephrol 2013;17(6):771–8. doi:10.1007/s10157-013-0812-x. PubMed PMID: 23652830.

MULTIPLE-CHOICE QUESTIONS

1. Which of the following is a "myeloma defining event" according to the 2014 International Myeloma Working Group (IMWG) guidelines?
   A. One focal 4-mm lesion on magnetic resonance imaging
   B. M-protein >0.5 g/dL
   C. Back pain
   D. Clonal bone marrow plasma cells ≥60%

2. Which one of these is a risk factor for progression from monoclonal gammopathy of undetermined significance (MGUS) to multiple myeloma or a related malignancy?
   A. Immunoglobulin G (IgG) MGUS
   B. Abnormal serum free light chain ratio
   C. M-protein ≥0.3 g/dL
   D. Patient age >50

3. Which one of these differentiates Waldenström macroglobulinemia (WM) from IgM MGUS?
   A. IgM M-protein ≥1 g/dL
   B. ≥10% plasma cells in bone marrow
   C. Renal damage
   D. Patient age <50

81

Concepts and Challenges in Organ Transplantation: Rejection, Immunosuppression, and Tolerance

Kathryn J. Wood, Sushma Shankar, Joanna Hester, Fadi Issa

The clinical era of transplantation began on December 23, 1954, when Dr. Joseph Murray and colleagues performed the first successful renal transplant on the genetically identical Herrick twins.[1] Solid organ transplantation (SOT) has since transformed the landscape of modern medicine, offering dramatic improvements in patient survival and quality of life in many end-stage diseases. The development of powerful immunosuppressive regimens and cutting-edge biological agents represents an elegant proof of concept, translating seminal work from the laboratory bench to the patient's bedside. However, this ground-breaking field within clinical immunology is not free of setbacks. Rejection of transplanted organs and tissues can result in devastating problems for the patient, and potent immunosuppression is associated with substantial comorbidity, including exposure to infection, malignancy, and cardiovascular risks that may be fatal. Theoretically, immunosuppression withdrawal and allograft-specific protection against host responses is the ultimate treatment to offer a transplant recipient. Although this is undoubtedly a challenging goal, it is being realized in defined subgroups of recipients. This chapter examines the rejection response, outlines the mainstay of current immunosuppressive therapy, and discusses the latest advances in the search for transplantation tolerance.

REJECTION

The immune system is complex and has evolved to protect the individual from harm. This harm may be either in the form of foreign pathogenic microorganisms or premalignant mutations in the individual's own cells. To achieve this safely, the immune system must have the ability to distinguish "self" from "non-self" or "altered-self" to avoid damaging the host itself. Any immune response that is generated must also be proportional to the threat, and thus antigens encountered in the context of inflammation will prime T cells and evoke a more aggressive immune response. Although an effective immune response is essential to survival in the context of infection or malignancy, it can represent significant management challenges in the transplantation setting.

Some elements of the immune system respond to the general trauma associated with organ retrieval, perfusion, and surgery, whereas others respond after specific recognition of antigenic differences between the donor and the recipient. If this process continues in an uncontrolled manner, it inevitably leads to destruction of the graft. Thus isografts or syngeneic grafts (grafting of tissue between two genetically identical individuals such that there are no antigenic differences) do not typically result in rejection, whereas allografts (transplantation of tissues from genetically different individuals) usually evoke an aggressive immune response as a result of the antigenic differences between the donor and the recipient. Tissue transplanted between species, or xenografts, are also possible but not yet in clinical use in cell and organ transplantations because these usually trigger a more rapid and aggressive immune response compared with an allograft. This section considers the processes involved in allorecognition and graft destruction.

Innate Immunity and Ischemia-Reperfusion Injury

Trauma to the graft is initiated during, or even before, retrieval because death of the donor brainstem induces hemodynamic and neuroendocrine responses; in addition, organs obtained from donors after their cardiac death suffer a period of warm ischemia. During cold storage, loss of intracellular potassium and cell injury continues. This leads to graft damage through cell swelling and a buildup of toxic metabolites, leaving the graft susceptible to further ischemia–reperfusion injury on rapid warming following revascularization. Efforts are made during storage and organ implantation surgery to reduce the metabolic rate through storage on ice and perfusion with specialized fluids. This approach aims to minimize the accumulation of toxic metabolites and the resultant pH changes to reduce the effect of cold ischemic time on posttransplantation outcomes, such as delayed graft function.[2]

Tissue injury leads to the expression of damage-associated molecular patterns (DAMPs), such as heparin sulfate, heat shock proteins (HSPs), nucleic acids, and high-mobility group box-1 (HMGB1) protein. These are identified by invariant pattern-recognition receptors (PRRs) of the innate immune system, for example, Toll-like receptors (TLRs) (Chapter 3). This process results in the local production of inflammatory mediators, such as interleukin (IL)-1 and -6, chemokines, and the expression of adhesion molecules within the graft. These early innate responses are seen in autografts as well as isografts and are generally not alloantigen specific, although a subcategory of innate cells does display a limited degree of alloantigen reactivity. In turn, this triggers activation of macrophages and dendritic cells (DCs) via cell-surface and internal PRRs,[3] causing both cell types to show greater antigen-presenting capacity and to enter a cytocidal state. Endogenous signals, such as stress, can also activate the complement cascade to generate several products, including complement component C3 in the graft, which promotes DC maturation and subsequently their ability to activate T cells. DCs are also activated

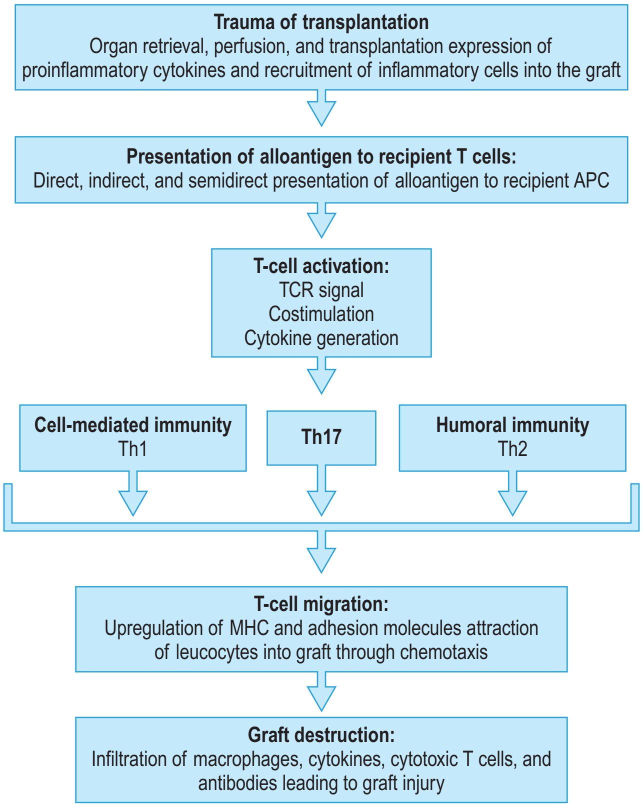

Trauma of transplantation
Organ retrieval, perfusion, and transplantation expression of proinflammatory cytokines and recruitment of inflammatory cells into the graft
Presentation of alloantigen to recipient T cells:
Direct, indirect, and semidirect presentation of alloantigen to recipient APC
T-cell activation:
TCR signal
Costimulation
Cytokine generation
Cell-mediated immunity
Th1
Th17
Humoral immunity
Th2
T-cell migration:
Upregulation of MHC and adhesion molecules attraction of leucocytes into graft through chemotaxis
Graft destruction:
Infiltration of macrophages, cytokines, cytotoxic T cells, and antibodies leading to graft injury

FIG 81.1 **Mechanisms Leading to Graft Injury.** Interplay of innate and adaptive immune responses results in eventual graft destruction.

by interferon-γ (IFN-γ) produced by natural killer (NK) cells. In one or more of these ways the innate immune system, activated by local tissue injury, promotes the initiation of adaptive immune responses when there are antigenic differences between the donor and the recipient. Activation of the adaptive immune system results in a series of effector mechanisms, both cell and antibody mediated, that lead to further graft injury (Fig. 81.1).

A severely damaged graft will initiate a more aggressive immune response, and this may explain the superior outcomes in living donor grafts compared with deceased donor grafts (as the former are subject to a lower degree of ischemic injury), even in the presence of significant major histocompatibility complex (MHC) mismatches.[4] The immune response evoked by the damaged tissue also accounts for the higher rates of rejection observed in individuals with delayed graft function.[5]

Initiation of the Adaptive Immune System

Recognition of differences between donor and recipient major and/or minor histocompatibility (miH) molecules or antigens by T cells is central to the adaptive immune response. The extensive polymorphism of MHC genes (Chapter 5) makes complete matching of unrelated individuals rare, and therefore, almost inevitably, transplantation of cells or organs between genetically unrelated individuals occurs across an MHC mismatch. Knowledge of the MHC and of the structure of MHC molecules has contributed significantly to our understanding of how rejection is triggered and has facilitated the development of novel immunosuppressive drugs. Additionally, transplantation between MHC-identical siblings has shown that differences in miH antigens can also be significant, particularly in bone marrow transplantation (BMT), where activation of the adaptive immune system of the donor may result in graft-versus-host disease (GvHD). These miH antigens may be derived from a wide variety of polymorphic proteins and are encoded for a vast array of genes distributed throughout the genome.

Signal 1: Recognition of Alloantigen

Class I MHC molecules are cell surface glycoproteins expressed on most nucleated cells and are recognized by the T-cell receptors (TCRs) of CD8 T cells. Class II MHC molecules are not expressed by every cell in the body; rather they are found on DCs, B lymphocytes, macrophages, and, in humans, endothelial cells. MHC class II molecules are recognized by T cells bearing the CD4 surface glycoprotein. Expression of both class I and II MHC molecules can be increased or, in the latter case, induced during inflammation, in particular by IFN-γ.

Class I and II MHC proteins contain a highly polymorphic groove or cleft, which can bind peptides, formed on a platform of beta strands, flanked by two α helices. Antigen processing within antigen-presenting cells (APCs) results in the production of peptides that can bind in these grooves producing a MHC–peptide complex that can be recognized by T cells (Chapter 6). These peptides may be of self origin or derived from foreign molecules (*e.g.*, from an allograft after transplantation or from a virus after an infection). In general, peptides derived from molecules present inside the cell are processed and loaded into MHC class I molecules, whereas molecules present outside the cell in the extracellular environment are processed into peptides that load into class II molecules. However, cross-presentation can also occur, whereby some APCs, notably DCs, process extracellular proteins such that they can be loaded onto MHC class I molecules and presented to CD8 T cells, leading to "cross-priming."[6]

The TCR is composed of two chains that confer MHC–peptide specificity and is associated with a complex of polypeptides referred to collectively as CD3 (Chapter 4). On antigen recognition by a TCR, CD3 delivers intracellular signals to the T cell (Chapter 12). Recognition of antigen by the TCR–CD3 complex is the first step in T-cell activation and commonly referred to as "signal 1." This initial meeting between host naive and memory T cells and alloantigen from the foreign transplant is believed to take place in secondary lymphoid organs rather than in the transplanted graft itself.[7] The inflammatory response triggered in the allograft by retrieval and implantation of the organ or tissue, as outlined above, not only initiates the migration of donor-derived passenger leukocytes but also their maturation into functional APCs expressing high levels of donor MHC molecules.

The presentation of intact allogeneic MHC molecules by donor-derived passenger leukocytes to T cells is known as the direct pathway of allorecognition. This is the dominant pathway through which the immune reaction to the graft is initiated, and T cells responding to direct antigen presentation constitute a vast majority of the alloreactive immune repertoire, estimated at 10% of T cells.[8] After a time, the donor lymphocytes transplanted with the graft are depleted; however, this does not abrogate rejection. The immune response against an allograft is maintained by recipient APCs, largely consisting of DCs and B cells, which process and present peptides derived from allogeneic MHC molecules shed from the graft (soluble MHC molecules or apoptotic cells) as well as miH antigens. This route of allorecognition is known as the *indirect pathway.* Experimental models have

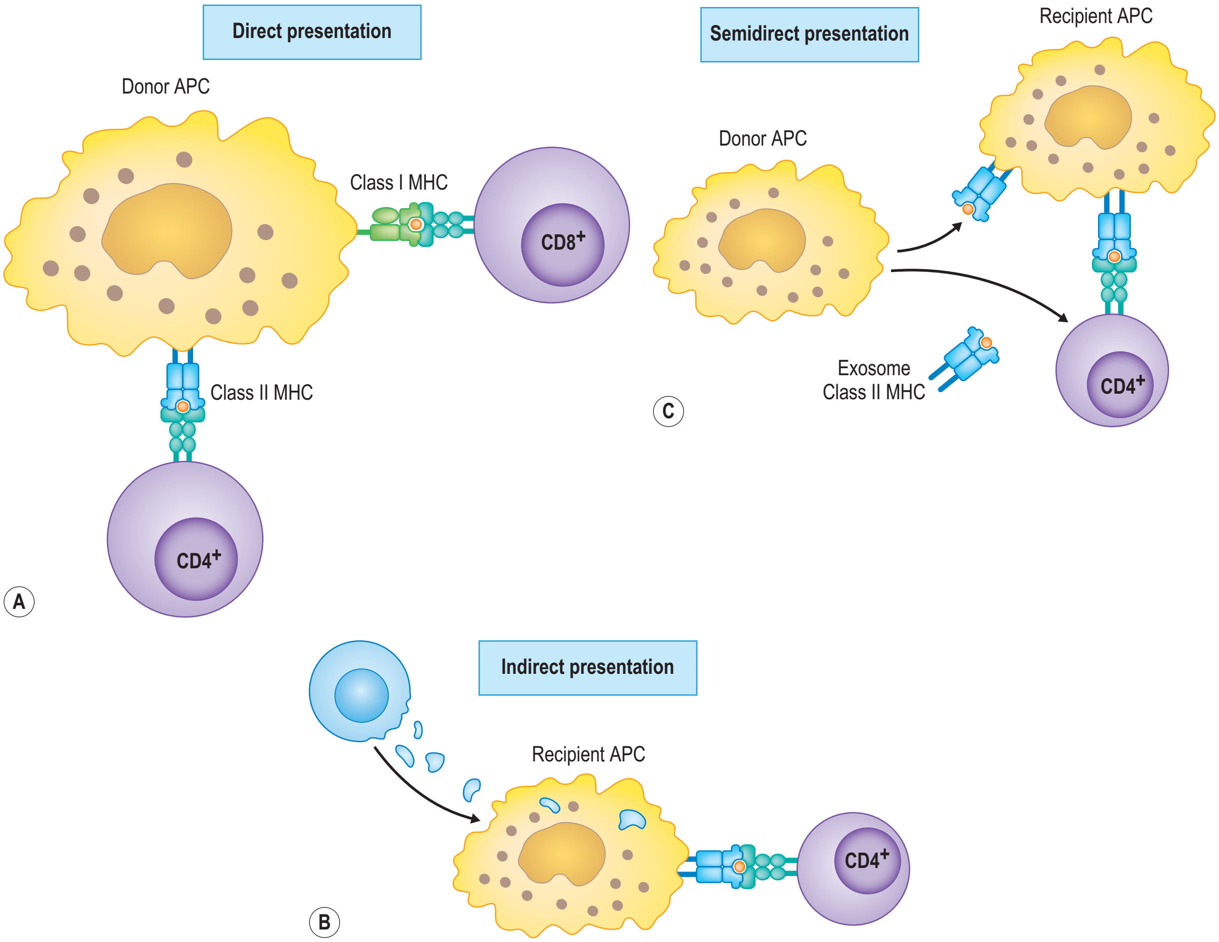


FIG 81.2 Antigen Presentation. (A) Direct presentation: passenger donor antigen-presenting cells (APCs) present alloantigen to recipient T cells in lymphoid tissue. (B) Indirect presentation: alloantigen from donor cells is processed and presented by recipient APCs via major histocompatibility complex (MHC) class II to recipient $CD4^+$ cells. (C) Semidirect presentation: donor MHC class I and class II may be transferred to the surface of recipient APCs enabling presentation of alloantigen to recipient T cells.

suggested that indirect presentation of donor antigens may play a greater role in rejection than direct presentation overall, as it continues for as long as the graft remains *in situ*.[9] It is also now known that the direct and indirect pathways of allorecognition interact as a third pathway referred to as the *semidirect pathway* of antigen presentation. In the semidirect pathway of allorecognition, donor MHC proteins are transferred intact to recipient APCs (through membrane transfer or the exosomal route), enabling them to present allogeneic MHC–peptide complexes to recipient T cells. This MHC transfer is temperature and energy dependent and requires close cell-to-cell contact. Both MHC class I and class II may be transferred, although class II MHC appears to be transferred more efficiently.[10] These three modes of antigen presentation are illustrated in Fig. 81.2.

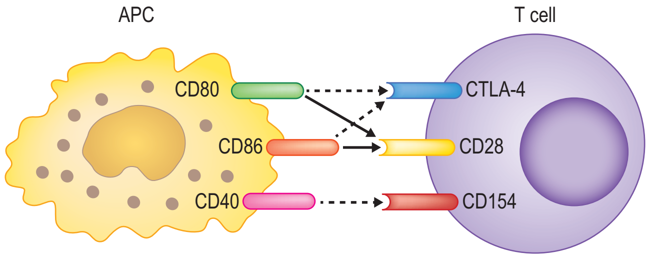


FIG 81.3 Costimulation. Costimulation follows binding to major histocompatibility complex (MHC) class I and class II on antigen-presenting cells (APCs) and involves signal transduction via intracellular proteins and increased interaction affinity via the binding of several cell surface proteins. Several costimulatory molecule pairs have been identified.

Signal 2: Costimulation

T lymphocyte activation relies not only on contact of the TCR–CD3 complex with its specific MHC–peptide complex (“signal 1”) but also on signals delivered by the interaction of several costimulatory receptors and their ligands, known collectively as “signal 2” (Fig. 81.3). During T-cell activation, the membrane of the T cell reorganizes such that the TCR–CD3 complex and costimulatory molecules are brought together in the cell membrane to form the immunological synapse. Signal 1 is specific for the antigen involved, whereas costimulation is not.

Costimulatory receptors fall into two major families, the B7 family (*e.g.,* CD28 and CD152) and the tumor necrosis factor (TNF) family (*e.g.,* CD154 and CD70) (Chapter 12). Although several costimulatory molecule pairs have been identified, the CD28 and CD154 pathways have been the most clearly defined to date. CD80 and CD86 are receptor ligands on the surface of DCs and other cells that can bind to CD28 on T cells, resulting in the activation of additional signal–transduction pathways within the T cell. This has a number of effects: lowering the threshold for T-cell activation; increasing glucose metabolism, as well as cytokine and chemokine expression, including IL-2 production; reducing T-cell death through apoptosis; and expanding the number of T cells that respond through proliferation. CD80 and CD86 are also linked to an inhibitory receptor, CD152 or cytotoxic T lymphocyte antigen-4 (CTLA-4), which inhibits T-cell activation possibly by competing with CD28 to bind with CD80 or CD86. CTLA-4 has a 10-fold higher receptor affinity compared with CD28 and is rapidly upregulated following T-cell activation. Additionally, CD152 activates tryptophan catabolism in DCs, which results in the inhibition of proliferation and promotes apoptosis of the responding T cells. Related to this pathway are other pathways involving members of the B7–CD28 family of molecules, including ICOS/B7h, PD1/PD-L1/PD-L2, and B7-H3. Each of these can play a role in T-cell costimulation and later coinhibition, but for the early phase of the response, the CD28 pathway dominates. CD40 belongs to the TNF receptor family, expressed on all APCs, and binds to CD154 (CD40L), which is present on activated CD4 cells, a subset of CD8 cells and NK cells. CD40 stimulation causes triggering signals for antibody production and induces MHC expression on APCs, thus amplifying antigen presentation.[11]

In clinical transplantation, treatment with calcineurin inhibitors (CNIs) to block signal 1 and/or the use of T cell–depleting agents has been highly successful in the prevention and reversal of rejection episodes. More recently, blockade of costimulatory pathways has also been shown to be effective. These approaches are discussed in the “Immunosuppression” section.

Signal 3: Proliferation and Differentiation of Effector T Cells

Following alloantigen recognition by the TCR–CD3 complex and costimulation, a process begins resulting in “signal 3.” This is the complex process involving the recruitment and phosphorylation of a series of signaling molecules, which induce intracellular biochemical processes that lead to the activation of three signal–transduction pathways: the calcium–calcineurin pathway, the Ras- and Rac-mitogen–activated protein kinase pathways, and the protein kinase C nuclear factor κB (NF-κB) pathway. These pathways interact with inositol triphosphate (IP3) and diacylglycerol (DAG), formed from the hydrolysis of phosphatidylinositol 4,5-bisphosphonate, to activate transcription factors: nuclear factor of activated T cells (NFAT), activating protein-1 (AP-1), and NF-κB, respectively. These transcription factors cause the expression of many genes leading to upregulation of growth factors and cytokines, in particular, IL-2 and CD25 (IL-2Rα). Growth signals are then delivered through the phosphoinositide-3-kinase (PI3K) and mechanistic target of rapamycin (mTOR) pathways to promote cell cycle progression, and initiation of clonal expansion and differentiation of activated T cells to express effector functions. These mechanisms are targeted by various immunosuppressive agents discussed in the “Immunosuppression” section and illustrated in Fig. 81.5.

The pattern of cytokine production through the interaction of the combination of signaling processes determines the nature of the response, in which either cell-mediated or antibody-mediated immunity is seen to dominate (Chapter 16). The T cells known to promote a cell-mediated response, and the cytokines they produce, are termed T-helper 1 (Th1) cells. These cells are associated with the production of IFN-γ. T cells promoting a humoral response are referred to as Th2 and are associated with the generation of IL-4, -5, and -6. Additionally, a Th17 population that has been identified is characterized by the production of IL-17 and promoting the infiltration of neutrophils.[12] Th22 cells, which express IL-13, IL-22, and TNF-α, have also been described. A subset of CD4 cells called *regulatory T cells* (Tregs) can also be induced following antigen exposure in the periphery (pTregs; Chapter 18).[13] These cells secrete IL-10 or transforming growth factor (TGF) and have suppressive or regulatory functions against effector cells and APCs. The balance of these responses results in either graft injury or the induction of tolerance (discussed below).

Memory T Cells

Following exposure to an antigen, antigen-specific memory T and B cells are generated. These memory cells are then able to produce a more rapid and intense immune reaction if the antigen is encountered on a second occasion because they have a lower activation threshold and are less dependent on costimulation. Transplant recipients, particularly older patients or those who have had previous antigen exposure through previous transplantation, blood transfusion, or pregnancy, may therefore have specific anti-donor memory cells. Memory-type responses may also occur as a result of antigen receptor cross-reactivity known as *heterologous immunity.*

Cell Migration

After activation, chemotactic cytokines, or chemokines, generated at the site of inflammation attract leukocytes that have been primed in secondary lymphoid tissues to the graft (Chapter 10). Once activated, leukocytes upregulate chemokine receptors, thus enabling their migration along a chemokine gradient. Vessels in proximity to the graft become dilated, allowing increased blood flow, and the endothelium becomes activated. Activated endothelial cells externalize Weibel-Palade bodies containing P-selectin, and chemokines generated in the graft, such as IL-8, RANTES, and MCP-1, stick to the vascular endothelium (Chapter 11). This signals the passing leukocytes in postcapillary venules to leave laminar flow of blood and move closer to the vessel edge so that they can interact with the activated endothelial cells. Extravasation then occurs in a multistep process, involving molecules on the surface of activated leukocytes called *Sialyl-Lewis moieties,* which interact with endothelial P-selectin, forming temporary bonds that form and break causing the leukocytes to “roll.” Chemokines on the endothelial surface then induce a conformational change in leukocyte integrin molecules, allowing high-affinity binding to endothelial adhesion molecules, such as intercellular adhesion molecule-1 (ICAM-1). When this occurs, the leukocytes stop rolling, and this allows them to extravasate and follow the chemokine gradient further into the graft, where damage then occurs.

The Effector Response and Graft Destruction

The immune system generates many different effector mechanisms, depending both on the challenge it meets and the microenvironments that are present. In transplantation, the type of transplanted tissue, graft site, and immune status of the recipient at time of transplantation may modify this response. Although initiation of rejection in a nonsensitized recipient is principally T-cell dependent, many components of the immune system contribute to the subsequent destruction of the transplanted tissue. Graft destruction may be alloantigen specific, or there may be bystander tissue destruction. The nature of the immune cells involved in the effector response is reflected in the characteristics of the resulting damage and in the speed of onset of tissue destruction.

Acute Antibody-Mediated Rejection

Alloantigen-specific antibodies, or alloantibodies, are secreted by plasma cells. Alloantibodies are produced after alloantigen-driven B-cell activation in the presence of T-cell help, such as can occur during rejection or following a blood transfusion. Antibodies that cross-react with alloantigens can also be generated as a result of infections that result in heterologous immunity. In the former, in addition to DCs, the B cells themselves may act as APCs. MHC class II molecules are presented to and bind immunoglobulins (Igs) on the surface of B cells, enabling the B cell to internalize the alloantigen and process it into peptides that are presented at the cell surface within MHC class II molecules (Chapter 6). The subsequently activated T cells produce cytokines that activate B cells enabling them to differentiate into alloantibody-producing plasma cells. When allospecific antibodies encounter their specific antigen, antibody-mediated rejection occurs. Antibodies function through a number of effector mechanisms. Antibody binding activates endothelial cells within the graft, resulting in the expression of adhesion molecules, cytokines, and chemokines as well as the synthesis of tissue factor. Antibody binding can trigger complement activation, which can result in cell lysis and graft damage directly or indirectly by the binding of complement components to the graft and the recruitment of macrophages and neutrophils. Additionally damage is caused by a mechanism whereby effector cells, such as NK cells and macrophages, bind to the Fc non–antigen-specific portion of the antibody. This encourages NK cells and macrophages to kill any target cells with antibody bound to their surface. This is a nonspecific process termed *antibody-dependent cellular cytotoxicity* (ADCC) and can contribute to graft rejection.[14]

When donor-specific alloantibodies (DSAs) exist in a recipient prior to transplantation (*e.g.,* as a result of previous transplantation, blood transfusion, or pregnancy), a dramatic response is seen upon perfusion of the transplanted graft. Coagulation and complement cascades are activated, resulting in extensive thrombosis and graft infarction within minutes. This so-called hyperacute rejection is now very rarely seen in clinical practice because of the advances in screening and cross-matching techniques as well as desensitization techniques, such as plasmapheresis and intravenous immunoglobulin (IVIG).

Acute Cellular Rejection

As described above, following the non–antigen-specific innate response to organ retrieval and implantation, an inflammatory environment is created within the graft, thus promoting an adaptive cellular response over the ensuing days. Naïve cytotoxic T cells, activated by CD4 cells clustering with APCs, migrate to the graft, where they recognize allogeneic class I MHC molecules. This causes them to release cytotoxic molecules, such as perforin and granzyme B; upregulate surface Fas ligand; and secrete soluble mediators, such as TNF-α. Perforins insert into the target cell membrane to form pores, allowing granzyme to enter the cell, and this causes proteolysis and activates the caspase cascade. Fas ligand binds to Fas on the target cell and also initiates the caspase cascade. These processes induce target cell apoptosis and acute cellular rejection and typically occur 1 week to 3 months after transplantation.

Additionally a nonspecific delayed-type hypersensitivity response occurs, usually mediated by CD4 cells that are attracted to the graft and involving the release of multiple proinflammatory cytokines, including IL-1, IFN-γ, and TNF-α. This leads to the recruitment and activation of further leukocytes, affects graft cell permeability and vascular smooth muscle tone, thus affecting graft physiology and contributing to acute and chronic rejection. CD4 alloreactive T cells responding to donor-derived peptides bound to recipient MHC class II molecules have also been correlated with chronic allograft dysfunction.[15]

As acute allograft rejection is initiated by the recognition of polymorphic donor MHC molecules by recipient T cells, it follows that transplantation of MHC incompatible tissues will elicit a strong, T cell–dependent immune response to donor tissues. In general, mismatches for only class I MHC result in slower rates of rejection compared with grafts with differences in class II MHC or both class I and II.[16] Rejection can still occur when transplantation has taken place between MHC-matched siblings because of T-cell recognition of minor histocompatibility antigens.

Delayed Allograft Rejection and Dysfunction

Rejection later after transplantation, which is often referred to as chronic rejection or chronic allograft dysfunction, comprises a number of mechanisms. It is characterized by a cellular infiltrate constituted of macrophages, eosinophils, NK cells, and T cells, where the primary target is the vascular endothelium.[17] Damaged endothelium allows platelet deposition on the arterial wall and growth factor production. This, in turn, leads to smooth muscle proliferation in the arterial wall media. Proliferating muscle cells invade the intima and contribute to intimal fibrosis. These changes are concentric, affecting all graft arteries. In heart transplants, this may appear as transplant arteriosclerosis. Chronic rejection therefore manifests in an organ-specific manner. For example, within renal glomeruli glomerulosclerosis and interstitial fibrosis may develop, whereas in lungs, bronchiolitis obliterans develops. The underlying mechanistic processes remain similar.

Clinical Implications

To prevent hyperacute rejection, patients on transplant waiting lists are monitored routinely for the development of anti–human leukocyte antigen (HLA) antibodies as a marker of sensitization. Immediately before transplantation, a further check is performed by mixing recipient serum and donor splenocytes and observing for cell lysis.[18] The extent of the risk of hyperacute rejection is determined by the antibody target and quantitative titer of the anti-donor antibodies.[19] Highly sensitized recipients can remain on waiting list for extended periods but may have successful transplantations after desensitization, as mentioned previously, by removal of existing antibodies through plasma exchange or adsorption or by depletion of B cells with immunosuppressive agents, such as rituximab or proteasome inhibitors. Despite the extensive screening performed to detect donor HLA antibodies before transplantation, preformed antibodies against donor non-MHC antigens may exist and have the ability to induce

1. Normal

2. Antibody-mediated rejection
Acute antibody-mediated rejection

I. ATN-like – C4d+, minimal inflammation
II. Capillary - margination and/or thromboses, C4d+
III. Arterial – v3, C4d+

Chronic active antibody-mediated rejection
Glomerular double contours and/or peritubular capillary basement membrane multilayering and/or interstitial fibrosis/tubular atrophy and/or fibrous intimal thickening in arteries, C4d

3. Borderline changes: 'suspicious' for acute T-cell-mediated rejection
This category is used when no intimal arteritis is present, but there are foci of tubulitis

4. T-cell-mediated rejection
Acute T-cell-mediated rejection

IA. Cases with significant interstitial infiltration (>25% of parenchyma affected, i2 or i3) and foci of moderate tubulitis (t2)
IB. Cases with significant interstitial infiltration (>25% of parenchyma affected, i2 or i3) and foci of severe tubulitis (t3)
IIA. Cases with mild to moderate intimal arteritis (v1)
IIB. Cases with severe intimal arteritis comprising >25% of the luminal area (v2)
III. Cases with 'transmural' arteritis and/or arterial fibrinoid change and necrosis of medial smooth muscle cells with accompanying lymphocytic inflammation (v3)

Chronic active T-cell-mediated rejection
'Chronic allograft arteriopathy' (arterial intimal fibrosis with mononuclear cell infiltration in fibrosis, formation of neo-intima)

5. Interstitial fibrosis and tubular atrophy, no evidence of any specific etiology

I. Mild interstitial fibrosis and tubular atrophy (<25% of cortical area)
II. Moderate interstitial fibrosis and tubular atrophy (26–50% of cortical area)
III. Severe interstitial fibrosis and tubular atrophy/loss (>50% of cortical area) (may include nonspecific vascular and glomerular sclerosis, but severity graded by tubulointerstitial features)

6. Other: Changes not considered to be due to rejection-acute and/or chronic

FIG 81.4 **Banff Diagnostic Criteria for Renal Allografts.** Histological criteria have been developed and are regularly reviewed to aid diagnosis of the cause of chronic allograft dysfunction. Determining the cause of dysfunction aids decision making with regard to pathology and management. C4d, complement degradation product C4d.

vascular rejection. There is increasing interest in determining the impact of such antibodies on graft outcome.

Acute graft rejection is suspected when there is a sudden deterioration in allograft function. Biopsy of the transplanted tissue and histological evaluation are performed, resulting in diagnosis of rejection and numerical grading of severity. Grading schemes, such as the Banff criteria shown in Fig. 81.4, which provide semiquantitative measures for histopathological assessment of the inflammatory response, have been developed for specific organs and form the basis on which further clinical management is guided; decisions to treat with high-dose steroids or other immunosuppressive medications can thus be made after structured analysis of a biopsy specimen of the transplant.

In addition to the diagnosis of cell-mediated rejection, modern staining techniques have enabled the identification of complement component 4d (C4d) in biopsy specimens from rejecting tissues, thus providing indirect evidence of antibody deposition and complement fixation in not only late rejection but also early rejection. Antibody-mediated rejection (ABMR) is frequently seen following heart, lung, and kidney transplantations; however, liver allografts are relatively protected.[20] Nevertheless ABMR is increasingly recognized as a cause of graft damage in liver transplantation. This is explored further in the discussion below on tolerance in the liver (Fig. 81.6).

Although improved immunosuppression regimens have led to a reduction in the occurrence of acute rejection, chronic rejection has become more evident and remains a significant contributor to late graft loss. Nonimmunological factors, including CNI-induced toxicity, advanced donor age, ischemic injury during implantation, hypertension, and infection, also contribute to chronic allograft dysfunction, and modifying these factors can improve graft outcome.[21] However, immunological processes play a significant role with increased levels of pretransplantation anti-HLA antibodies, *de novo* posttransplantation donor-specific antibodies and antibodies against non-HLA antigen MHC class

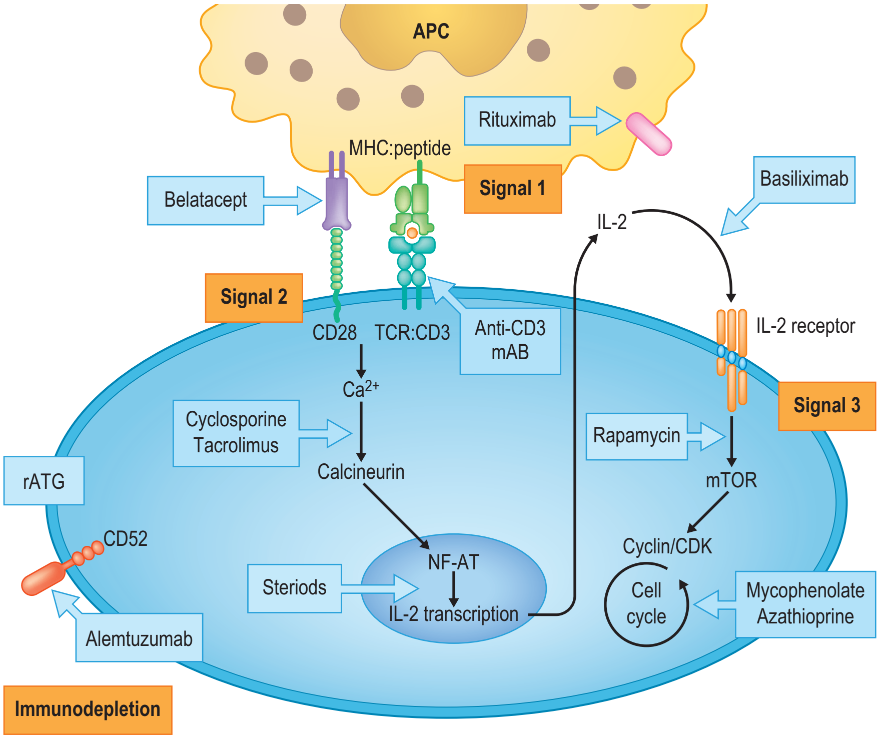


FIG 81.5 **Immunosuppressive Drugs and Their Targets.** Signal 1 results from major histocompatibility complex (MHC): antigen recognition through the T-cell receptor (TCR)–CD3 complex, a process blocked by anti-CD3 mAbs and indirectly by rituximab. Signal 2 results in costimulation, a process which can be blocked by belatacept. Costimulation results in full activation of the TCR–CD3 complex, initiating signal transduction, Signal 3: Downstream signaling pathways result in calcineurin activation, a stage which can be inhibited by tacrolimus and cyclosporine. Activated calcineurin dephosphorylates nuclear factor of activated T cells (NFAT), allowing interleukin-2 (IL-2) transcription, a process that can be inhibited by steroids. IL-2 receptor stimulation, a step which can be blocked by basiliximab, activates the mechanistic target of rapamycin (mTOR) signaling cascade, which can be inhibited by sirolimus. This pathway induces the T cell to enter cell cycle and proliferate, which, in turn, can be blocked by mycophenolate and azathioprine. Rabbit antithymocyte globulin (rATG) exerts polyclonal effects, and alemtuzumab binds to CD52, both resulting in immunodepletion,

I polypeptide-related sequence A (MICA) being seen to affect allograft survival.

The correlation among late acute rejection, chronic allograft dysfunction, and graft loss has been consistently reported, and there is increasing evidence that subclinical rejection may represent an important factor in predicting graft loss.[22] Establishing a clear link between subclinical rejection and development of chronic allograft dysfunction may be difficult, however, because of the interplay of other factors, including CNI toxicity.[21]

IMMUNOSUPPRESSION

In recent years, substantial advances in immunosuppressive strategies and their translation to routine clinical practice have revolutionized management and outcomes in solid organ transplantation (SOT)—an option that has become the therapy of choice for many end-stage organ diseases. Short-term outcomes, such as patient and allograft survival at 1 year, acute rejection rates, and time-course of disease progression and symptom control have steadily improved since the first successful transplantations performed over 50 years ago.

Clinical regimens implemented today are focused on targeting the steps involved in the T cell–mediated immune response to alloantigen. As outlined earlier, there are three main stages in this pathway: recognition of alloantigen or self antigen; costimulation; and proliferation/differentiation of effector T cells. The current clinical paradigm is based on blockade of at least one of these stages and/or by total immunodepletion therapy as well as by alterations in lymphocyte trafficking (see Fig. 81.5).

The risk of acute graft rejection is greatest during the initial three months following transplantation; therefore current immunosuppression strategies are primarily based on a potent induction regimen using a monoclonal antibody (mAb) or polyclonal antibody, followed by "maintenance immunosuppression" often consisting of calcineurin inhibitors (CNIs, cyclosporine or tacrolimus), an antiproliferative agent (mycophenolate mofetil, MMF), and low-dose corticosteroids (prednisolone). Provided there are no episodes of acute rejection, the doses of these agents are gradually reduced, and then maintenance immunosuppression is continued indefinitely (Fig. 81.6). This dose reduction is based on the concept that as graft inflammation subsides and

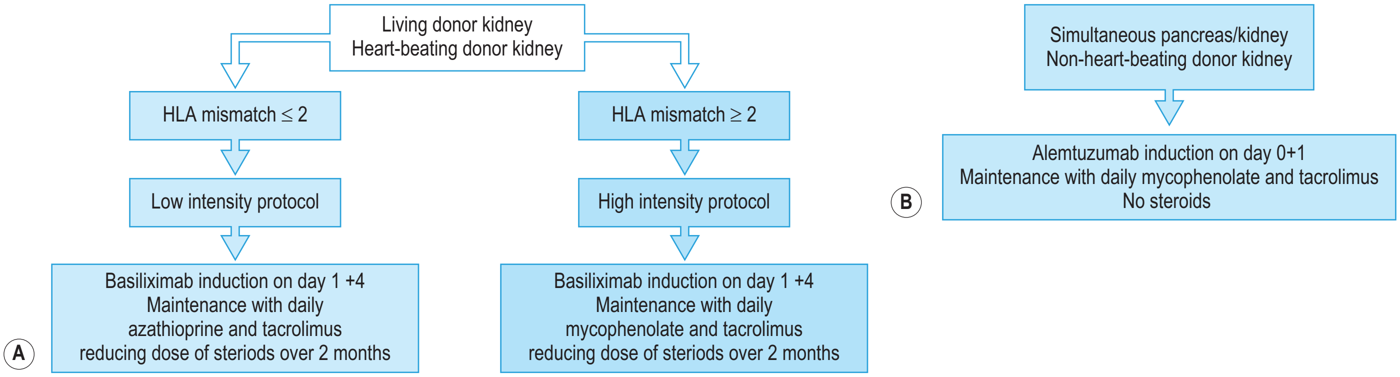

Living donor kidney
Heart-beating donor kidney
HLA mismatch ≤ 2
HLA mismatch ≥ 2
Low intensity protocol
High intensity protocol
Basiliximab induction on day 1 +4
Maintenance with daily
azathioprine and tacrolimus
reducing dose of steroids over 2 months
Basiliximab induction on day 1 +4
Maintenance with daily
mycophenolate and tacrolimus
reducing dose of steroids over 2 months
A
Simultaneous pancreas/kidney
Non-heart-beating donor kidney
Alemtuzumab induction on day 0+1
Maintenance with daily mycophenolate and tacrolimus
No steroids
B

FIG 81.6 **Examples of Immunosuppressive Protocols in Clinical Use.** (A) Patients receiving a renal transplant from a living donor or heart-beating donor are stratified according to human leukocyte antigen (HLA) mismatch. When HLA mismatch ≤2, the patient is treated with a low intensity protocol; when HLA mismatch ≥2, the patient is treated with a high intensity protocol. (B) Pancreas and non–heart-beating donor kidneys have higher immunogenicity, and so induction is with a potent monoclonal antibody, alemtuzumab; therefore steroids are not required.

TABLE 81.1 Immunosuppressive Therapies in Transplantation: Maintenance and Induction Agents

Drugs	Mechanisms	Adverse Effects
Maintenance Agents		
Azathioprine	Inhibits purine and DNA synthesis, inhibits cell proliferation	Bone marrow depression, opportunistic infection, macrocytosis, liver toxicity
Cyclosporine	Binds to cyclophilin, inhibits calcineurin-phosphatase, blocks NFAT dephosphorylation, blocks IL-2 transcription and & T-cell activation	Hypertension, hyperlipidemia, nephrotoxicity, hepatotoxicity, pancreatitis, peptic ulcers, thrombotic microangiopathy, opportunistic infection, neurotoxicity, tremor, gingival hyperplasia, hirsutism
Mycophenolate mofetil	Inhibits inosine-monophosphate-dehydrogenase, inhibits purine synthesis and blocks cell proliferation	Gastrointestinal symptoms, bone marrow depression, opportunistic infection in particular CMV and BK nephropathy
Rapamycin	Binds to FKBP12, inhibits mTOR and blocks IL-2–driven cell proliferation	Delayed graft function, delayed wound healing, mouth ulcers, pneumonitis, increased proteinuria, peripheral edema, hyperlipidemia
Steroids	Induces phospholipase A_2 inhibitory proteins, inhibits arachidonic acid synthesis, inhibits prostaglandins and leukotrienes	Diabetes, delayed wound healing, peptic ulcers, psychosis, osteoporosis, infection, blurred vision, fluid retention, weight gain, acne, constipation
Tacrolimus	Binds to FKBP12, inhibits calcineurin-phosphatase and blocks T-cell activation	Posttransplantation diabetes mellitus, nephrotoxicity, thrombotic microangiopathy, neurotoxicity
Induction Agents		
Antithymocyte globulin	Polyclonal effects not well characterized; immunosuppressive efficacy attributed to T-cell depletion through apoptosis, antibody-dependent cytolysis and complement-dependent lysis	Polyclonal effects: cytokine release syndrome, serum sickness, leukopenia, thrombocytopenia. *De novo* tumors and opportunistic infection: CMV and HSV.
Alemtuzumab	Binds to CD52 antigen, (expressed on 95% of peripheral blood lymphocytes, NK cells, macrophages, and thymocytes) Results in profound lymphopenia	Opportunistic infection: disseminated *Candida* and CMV infection
Basiliximab	Binds to IL-2R with similar affinity as IL-2, thereby inhibiting IL-2-driven T-cell proliferation	Occasional hypersensitivity reactions, inadequate immunosuppression in immunologically high-risk recipients

IL-2, interleukin-2; IL-2R, interleukin-2 receptor; NFAT, nuclear factor of activated T cells; DHFR, dihydrofolate reductase; mTOR, mechanistic target of rapamycin; CMV, cytomegalovirus; HSV, herpes simplex virus.

donor-derived APCs are cleared, the immunogenicity and risk of rejection decreases.

Chronic immunosuppression is associated with several undesirable sequelae (Table 81.1), in particular an increased relative risk of infections and malignancy, cardiovascular morbidity (the leading cause of death with a functioning graft in transplant recipients), and target organ damage.

Many different immunosuppressive regimens for SOT are in clinical use and are dependent on local factors and guidelines. An introduction to the immunosuppressive agents currently in widespread use is provided below.

Immunodepletion

The more recent therapeutic strategies are based on induction therapies that concentrate on profound immune cell depletion at the time of transplantation, when immune activation is most intense.

CLINICAL PEARLS

Treatment Options for Rejection

- Acute cellular rejection
  - High-dose steroids
- Rejection refractory to steroid treatment
  - Antithymocyte globulin
  - Alemtuzumab
- Acute humoral rejection
  - Plasmapheresis
  - Intravenous immunoglobulin
  - Rituximab

Antithymocyte Globulin

Rabbit-derived antithymocyte globulin (rATG) is a lymphocyte-depleting polyclonal IgG preparation with specificity toward human thymocytes. It binds primarily to the peripheral blood lymphocytes as well as to those present in the lymphoid organs, including lymph nodes, spleen, and thymus, as demonstrated by *in vivo* studies in monkeys. The agent's polyclonal nature enables it to display specificity toward a wide variety of molecules expressed on the surface of T cells, B cells, DCs, NK cells, and endothelial cells, including those involved in T-cell activation, proliferation, apoptosis, signal transduction, cell adhesion, and trafficking.

The precise mechanism of action underlying the immunosuppressive efficacy of rATG in SOT recipients is unknown at present, although has been primarily attributed to T-cell depletion. Data from *in vitro* studies have suggested that rATG modulates the expression of various lymphocyte surface antigens resulting in apoptosis, antibody-dependent cytolysis, or complement-dependent lysis. Lymphocyte depletion with rATG has been further demonstrated in adult recipients of renal transplants in several randomized, comparative clinical studies, with repopulation reported to take at least 3 months.[23] More recently, data from preclinical and clinical studies have suggested that rATG therapy may induce the expansion and enrichment of certain Treg subsets, such as $CD4^+CD25^+$ Forkhead box $P3^+$ ($FOXP3^+$), CTLA-4^+, and glucocorticoid-induced TNF receptor $(GITR+)^+$ Tregs from human peripheral blood lymphocytes or peripheral blood mononuclear cells *in vitro*[24]; these Treg subsets are key to immune response modulation and are further discussed in the "Tolerance" section.

In terms of clinical use, rATG induction in combination with tacrolimus-based immunosuppressive therapy is more effective in preventing episodes of acute renal graft rejection than with tacrolimus-based therapy without induction, as reported by primary endpoint data from two, randomized, open-label, multicenter trials.[25,26] Moreover, the median time to biopsy-proven acute rejection (BPAR) was >1 week longer in rATG induction than in noninduction; however, it should be noted that there was no significant difference between induction and noninduction regimens in terms of patient or graft survival.

Several randomized, multicenter studies have reported that the efficacy of rATG induction therapy is generally no different from that of basiliximab or low-dose daclizumab (IL-2 receptor [IL-2R] mAbs, which are an alternative induction agent; discussed later in this section) with regard to the incidence of BPAR, graft loss, or death at 6 and 12 months after transplantation, in the context of renal transplant recipients receiving triple immunosuppressive maintenance therapy. These trials, however, were not powered to show overall superiority of one agent over the other and were powered to address safety rather than efficacy parameters or used a limited daclizumab dosage regimen.[24] More robust studies that are designed to evaluate the efficacy of rATG relative to IL-2R mAbs specifically, are needed to definitively establish the use of rATG with respect to these monoclonal agents.

Anti-CD52 Monoclonal Antibody (Alemtuzumab)

Anti-CD52 mAb (alemtuzumab) is a humanized rat IgG2b directed against the CD52 antigen, which is expressed on 95% of peripheral blood lymphocytes, NK cells, macrophages, and thymocytes[27]; thus almost all mononuclear cells are affected. The profound and long-lasting lymphopenia produced after the administration of one or two doses of alemtuzumab is likely explained by such abundance on monocyte cell surfaces. Examination of the peripheral blood lymphocytes from recipients after alemtuzumab induction has identified a subset of T cells, predominantly CD4 central memory cells that survive despite alemtuzumab induction and appear largely resistant to depletion; these memory T cells express lower CD52 levels compared with naïve T cells.[28,29] CD52 is not present on granulocytes, platelets, erythrocytes, or hematopoietic stem cells (HSCs).

After binding to CD52, alemtuzumab causes cell death through several mechanisms: complement-mediated cytolysis, antibody-mediated cytotoxicity, and apoptosis. The plasma elimination half-life is approximately 12 days, whereas its clinical effects are far more persistent. Lymphocyte depletion of >99% can be seen after a single dose, with lymph node depletion taking up to 3–5 days compared with <1 hour seen in the peripheral lymphocytes. Different subpopulations display varying rates of recovery, depending on the subpopulation of interest: NK cells are almost unaffected and decrease only transiently (a population of $CD52^-$ NK cells has also been identified); monocyte and B-cell recovery can be seen at 3 and 12 months, respectively; T cells levels recover to only 50% of baseline at 36 months.[30]

Alemtuzumab is currently being used for induction in SOT with the aim of minimizing maintenance immunosuppression. It was first used in transplantation as an induction agent in 1998,[31] in 13 renal transplant recipients who received low-dose cyclosporine alone as maintenance therapy. At 6–12 months follow-up, patient and graft survival were 100%, and there were two episodes of acute rejection. The 5-year results of the initial series,[32] together with another 20 patients who were subsequently enrolled, showed no significant difference in graft or patient survival or in acute rejection rates compared with a retrospective, contemporaneous control group of 66 renal transplant recipients who had received no induction, but triple immunosuppression therapy alone (cyclosporine, azathioprine, and prednisolone). The study did also find, however, more episodes of late rejection in the alemtuzumab-treated group. Indeed, severe lymphopenia and homeostatic cytokines are known to drive the rapid homeostatic proliferation of naïve and memory T cells, and lymphocytes generated under such conditions have been previously found to be potent alloreactive cells, inevitably triggering rejection in animal models.[33] Investigation of this phenomenon in human recipients[28] illustrated that the recovery of the immune system after alemtuzumab induction differed with respect to CD4 and CD8 T cells. CD8 T cells recovered to pretransplantation levels by 6 months after alemtuzumab induction, whereas the number of CD4 T cells remained low, even at 1-year after transplantation. Repopulating CD8 T cells were mainly of the immunosenescent $CD28^-CD8^+$ phenotype, and the percentage of this population was lower in those recipients who experienced rejection compared with those who did not mount an immune response to their allograft. A series of *in vitro* experiments demonstrated that the expanded $CD28^-CD8^+$ T cells compete for immune space with CD4 T cells, suppressing their proliferation and thus delaying CD4 T-cell recovery. This delay might be associated with the clinical outcome because CD4 T cells, notably CD4 T effector memory cells, have been shown to be associated with rejection.[34]

Despite considerable experience with this agent over many years, largely in renal transplantation, until recently few randomized controlled trials had been conducted. Two randomized controlled clinical trials, the INTAC and 3C studies, investigating alemtuzumab induction in renal transplantation have now provided further insight.[35,36] The data from these studies support the use of alemtuzumab as an induction agent in low-risk patients, and the incidence of biopsy-confirmed acute rejection at 1 year was lower in patients treated with alemtuzumab compared with those treated with basiliximab induction. Similar to ATG, alemtuzumab cannot be used as a solitary immunosuppressive therapy and requires long-term maintenance immunosuppression to prevent allograft rejection. Importantly, in the 3C study, alemtuzumab induction therapy was found to result in reduced CNI and mycophenolate exposure and steroid avoidance, reducing the risk of biopsy-proven acute rejection in a broad range of patients receiving a kidney transplant, and this supported earlier data from smaller studies, including a study in lung transplant recipients.[37] However, patients with preformed DSAs despite a negative cross-match are at high risk of adverse outcomes when receiving a minimal immunosuppressive regimen incorporating alemtuzumab induction; these patients may thus benefit from augmented immunosuppression.[38]

Review of rejection episodes that have occurred with alemtuzumab induction, particularly in the absence of CNIs, has highlighted some interesting findings. Several authors have reported that a number of these rejections demonstrate positive staining for the complement component C4d, indicative of acute ABMR. In contrast, in acute rejection, examination of biopsy specimens taken from patients who had received both alemtuzumab induction and tacrolimus revealed none that was humoral in origin.[27] These data suggest that tacrolimus may prevent this early ABMR. Immune profiles of renal transplant recipients receiving alemtuzumab induction, 60 days of a CNI, and subsequent sirolimus monotherapy displayed a general increase in the naïve B-cell population.[39] Bloom et al. subsequently showed that B cell–activating factor (BAFF), a B-cell survival cytokine influencing the threshold of B-cell activation, substantially increased in renal transplantation following treatment with alemtuzumab.[40] This observation may help explain and prevalent the development of alloantibody in patients treated with depleting antibody therapy at the time of SOT.

Signal 1: Blockade of Antigen Recognition

Activation of the rejection response to an allograft hinges on the recognition of antigen by the host immune system. Targeting signal 1 through the use of mAbs has been used in both transplantation and autoimmunity.

Anti-CD3 Monoclonal Antibody

Muromonab-CD3 (OKT3), a mouse mAb binding to the CD3 component of the TCR signal–transduction complex, was used successfully as an induction agent for high-risk patients or for the treatment of corticosteroid-resistant acute rejection episodes as long ago as 1980. However, severe adverse effects of muromonab-CD3, such as cytokine-release syndrome (owing to the propensity of muromonab-CD3 to initially activate T cells, releasing TNF-α, and IL-2), resulted in preference for rATG and anti-CD52 agents, as outlined above. Furthermore, OKT3, an mAb of mouse origin, was limited by host production of human antimouse antibodies (HAMAs) that bind to circulating reagent molecules, neutralizing their activity. Nonactivating, therefore less toxic, humanized Fc-receptor nonbinding anti-CD3 mAbs, such as teplizumab and visilizumab, have been tested in clinical trials in autoimmune disease settings (type 1 diabetes, Crohn disease) as well as in renal and pancreatic islet transplantation.[41] Indeed, these studies have suggested a further mechanism of action with regard to anti-CD3 antibodies, and this could be helpful in the setting of transplantation: Circulating CD8 T cells isolated from patients with type I diabetes, after treatment with anti-CD3 mAb, may have regulatory function *ex vivo;* recent data have demonstrated functional human CD8 Tregs *in vivo* after administration of anti-CD3 mAb.[42] Further work is needed to elucidate the action of such agents on the immune response to both self and alloantigens to realize their full potential.

Anti-CD20 Monoclonal Antibody (Rituximab)

Rituximab is an anti-CD20, chimeric mAb that eliminates most B cells from the circulation. Originally approved to treat B-cell lymphoproliferative diseases in patients other than transplant recipients as well as to treat posttransplantation lymphoproliferative disease (PTLD), it is also now used in SOT as a treatment of antibody-mediated rejection and to desensitize patients who are receiving ABO-incompatible transplants or retransplants. However, depletion of antibody-producing cells may be incomplete because rituximab cannot target $CD20^+$ plasma cells. It has been hypothesized that the action of rituximab in allograft rejection is caused by not only depletion of plasma cells but also the effective depletion of APCs, which thus limits indirect pathway T-cell activation and a sustained immune response. Clinical administration is in conjunction with maintenance immunosuppressive agents, plasmapheresis, intravenous immunoglobulin (IVIG), and even splenectomy in some protocols.[43]

Recent work with nonhuman primate (NHP) models has shown that preemptive B-cell depletion with rituximab prevents both acute cellular rejection and antibody-mediated rejection of cardiac allografts; alloantibody was prevented from developing in blood, and complement-mediated graft injury was avoided. Not only was the B-cell response inhibited, but acute cellular rejection was also reduced by the combination of cyclosporine and rituximab, suggesting that B cells may play a role in acute cellular rejection as well. There may be a role for targeting B cells preemptively in SOT, and additional strategies to downregulate the B-cell response are under evaluation. Promising targets include BAFF and its receptor, as well as costimulatory molecules, such as CD28, CD154, and CD40.[44]

Signal 2: Blockade of Costimulation

In the absence of appropriate costimulation, partially activated T cells either become hyporesponsive to subsequent antigen-specific TCR signals (donor-specific anergy) or die by apoptosis. It has been hypothesized that the inhibition of full T-cell activation by costimulatory blockade rather than total T-cell depletion might more selectively target effector T cells and spare beneficial Tregs while avoiding the many adverse effects of nonspecific immunodepletion.

CD28:B7 (CD80/CD86) Blockade

CTLA-4 (CD152) is an inducible, T-cell surface antigen that when bound to CD80/86 receptor:ligand (B7 molecules) on APCs, delivers inhibitory signals to the activated T cell. Belatacept (LEA29Y) is a fusion protein that combines a mutated version of the extracellular binding domain of CTLA-4 with the Fc portion of IgG1, with specificity for CD80/86 expressed on APCs. Ligation of CD80/86 by CD28 (a surface antigen constitutively expressed on T cells) usually lowers the activation threshold of T cells. Belatacept has a higher affinity and slower dissociation rate from human B7 molecules (*i.e.,* CD80/86) compared with CD28, resulting in inhibition of the costimulation required for effective T-cell activation.[43]

The complexities of the human immune system present significant challenges to the translation of such agents into clinical practice: *In vivo* work indicates memory and cytotoxic $CD8^+$ T cells have different costimulation requirements for complete activation compared with $CD4^+$ naïve T cells and may therefore be more resistant to some tolerance induction strategies. Such blockade may also affect the function of Tregs as the CD28:B7 pathway has been shown to be important for their expansion and survival. However, reassuringly, data from clinical trials of belatacept in renal transplantation, albeit with small patient numbers, suggest that costimulation blockade does not interfere with Treg homeostasis.[45,46] In non-human primate models, CTLA-4 Ig prolongs pancreatic islet survival and, in combination with CD154:CD40 pathway blockade, induces indefinite acceptance of renal and heart allografts.

BENEFIT (Belatacept Evaluation of Nephroprotection and Efficacy as First-line Immunosuppression Trial), a 3-year, phase III clinical trial, randomized patients to three groups: cyclosporine, less-intensive belatacept, and more-intensive belatacept. Patients received induction with basiliximab and were maintained on mycophenolate mofetil (MMF) and corticosteroids. Patient and graft survival rates were similar across the three groups at both 1 and 2 years after transplantation. At the end of 1 year, although the incidence of acute rejection was greater for more-intensive and less-intensive belatacept compared with cyclosporine, no apparent impact on graft survival was demonstrated. At the end of 2 years, however, glomerular filtration rates (GFRs) continued to be significantly higher (15–17 mL/min) in the belatacept-treated patients. Encouragingly, belatacept-treated patients also had sustained benefits in their cardiovascular and metabolic risk profiles compared with those treated with cyclosporine, supporting the potential role of belatacept as a nonnephrotoxic alternative to CNIs.[47,48] Patients enrolled in this trial have now been followed up for 7 years, and the long-term data show that patient and graft survival and the mean eGFR are significantly higher with belatacept (both the more-intensive regimen and the less-intensive regimen) than with cyclosporine.[49]

Because of concerns about an increased post-transplant lymphoproliferative disease (PTLD) risk in Epstein-Barr virus (EBV)–seronegative patients, current belatacept trial protocols have now been modified to enroll EBV-seropositive patients only. Unfortunately, this exclusion will complicate the introduction of this drug in young patients, who might derive the maximum long-term benefit from nonnephrotoxic regimens. A significant limitation to note is that administration of belatacept requires intravenous infusion; this is not an ideal characteristic for a maintenance immunosuppressant and raises issues of patient compliance. Furthermore, the agent has a long half-life of 8–10 days and so may be difficult to appropriately titrate in transplant recipients suffering from infections.

Signal 3: Blockade of Proliferation/Differentiation

Activated T cells produce IL-2, which, in turn, binds to IL-2Rs, which are expressed only on the surface of activated cells and are not present on resting T cells. IL-2R is composed of three high-affinity transmembrane protein subunits: α (CD25), β (CD122), and γ (CD132) subunits, which are covalently linked. The α subunit is specific to IL-2R only, and it is the binding of α and β subunits that is crucial to the IL-2 signal transduction and T-cell activation that subsequently leads to proliferation and clonal expansion of T and B cells specific to alloantigen or self antigen. These cells are also stimulated to release more IL-2, further magnifying the immune response.

Anti-IL-2R Monoclonal Antibody (Basiliximab and Daclizumab)

Anti-IL-2R (anti-CD25) mAbs specifically target activated T cells but do not cause significant lymphocyte depletion and are not associated with major adverse effects compared with lymphocyte-depleting agents. However, other T-cell subtypes, including Tregs, also express CD25, and therefore the use of these agents may impact some of the natural mechanisms of immunoregulation. Basiliximab (chimeric form) and daclizumab (humanized form) have been used in renal transplantation in low-risk recipients, as defined as first allografts with living-related donors.[50,51] Basiliximab is currently the agent of choice and binds to IL-2R with similar affinity as IL-2, thereby effectively competing with IL-2 and subsequently inhibiting IL-2–driven T-cell proliferation. Basiliximab has a high volume of distribution, almost completely saturating IL-2R on the peripheral lymphocytes within 24 hours of a single dose of 2.5–25 mg in renal transplant recipients. The half-life in adults is approximately 13.4 days, and IL-2R saturation and suppression can last for 4–6 weeks.

Two meta-analyses evaluating the efficacy and safety of basiliximab in renal transplant recipients have been published.[51,52] Both studies showed that basiliximab was more effective than placebo in reducing acute cellular rejection 6 months after transplantation. However, both meta-analyses illustrated no significant differences in patient or graft survival rates between basiliximab and placebo groups at 1 year after transplantation. However, in studies involving high-risk transplant recipients (*i.e.,* high HLA-mismatch) the risk of rejection may be higher with IL-2R mAb in comparison with rATG.[53]

The highly selective and short-term immunosuppressive effects of basiliximab, which are confined to the highly immunogenic period immediately after transplantation, make this class of drug a useful substitute to steroids in early steroid-withdrawal or steroid-free regimens, as illustrated in studies with liver transplantation. Several prospective clinical trials using basiliximab induction to facilitate early steroid withdrawal or complete steroid avoidance after kidney transplantation were tried and proven safe. However, similar to alemtuzumab, the use of complete CNI avoidance protocols should be practiced with extreme caution; in several studies that withheld CNIs for any more than a very short period, despite adequate immunosuppression with IL-2R mAbs, MMF, and steroids, the acute cellular rejection rates after renal transplantation were much higher than with CNI use.[54] Thus as with ATG and alemtuzumab, patients receiving basiliximab are still exposed to chronic administration of maintenance immunosuppression and its associated comorbidities.

Glucocorticoids

Corticosteroids have complex immunosuppressive as well as antiinflammatory effects (Chapter 86). They act principally by binding to cytoplasmic glucocorticoid receptors, although at higher doses they can exhibit receptor-independent effects as well. The steroid–receptor complex translocates to the nucleus, where it is able to alter the expression of multiple cytokines through DNA-binding and by targeting transcription factors, such as AP-1 and NF-κB. Corticosteroids reduce the expression of many molecules important in the immune response, including IL-1, -2, -3, and -6; TNF-α; IFN-γ; and a number of chemokines. By inhibiting cyclooxygenase, corticosteroids are also able to reduce the production of inflammatory mediators, such as leukotrienes and prostaglandins.

Corticosteroids have been the mainstay of maintenance immunosuppression regimens for several decades. However, because of the agents' wide-ranging effects, not only involving the immune response but also a multitude of organ systems, and its direct nephrotoxicity and diabetogenic properties, in modern immunosuppressive regimens, an attempt is made to avoid or withdraw corticosteroids at the earliest possible time after transplantation (Table 81.1).

Antiproliferative Agents

Azathioprine and MMF interfere with DNA synthesis and prevent cell cycle progression. In the context of allograft rejection, this impairs the clonal expansion of alloreactive T cells.

Azathioprine is metabolized in the liver to the purine analogue 6-mercaptopurine and incorporated into DNA. By inhibiting purine nucleotide synthesis (and therefore DNA and RNA synthesis), it reduces gene transcription and prevents cell cycle progression. The effects of azathioprine are not lymphocyte specific, and patients must be monitored closely for bone marrow suppression.

MMF is metabolized in the liver to mycophenolic acid, which is a noncompetitive, reversible inhibitor of inosine monophosphate dehydrogenase (IMPDH). Cells are able to generate purines either *de novo* by converting inosine monophosphate to guanosine monophosphate (catalyzed by IMPDH) or from guanine via the salvage pathway. The salvage pathway is less active in lymphocytes and therefore are relatively dependent on the *de novo* pathway of purine synthesis compared with other cell types. As a result, the effects of MMF are more lymphocyte specific than azathioprine, and it is less myelosuppressive. Both agents are commonly used in modern maintenance immunosuppression regimens.

Calcineurin Inhibitors

The introduction of cyclosporine in the early 1980s presented a great step forward in transplant-related immunosuppression because this was the first drug able to selectively block T-cell activation. Subsequently a second CNI, the macrolide antibiotic tacrolimus, replaced cyclosporine in most immunosuppressive regimens. Almost 99% of renal transplant recipients currently receive CNIs.[55] Both cyclosporine and tacrolimus bind cytoplasmic immunophilins (cyclophilin in the case of cyclosporine and FK506-binding protein 12 [FKBP12] in the case of tacrolimus) to form complexes that can inhibit the calcium-dependent phosphatase calcineurin, a rate-limiting enzyme in the TCR signal–transduction pathway. By preventing translocation of the transcription factor NFAT to the nucleus, calcineurin inhibition impairs upregulation of many molecules important for T-cell proliferation and the generation of an effective immune response, including the cytokines IL-2, IL-4, TNF-α, and IFN-γ, and costimulatory molecules, such as CD154. However, CNIs are well known for their nephrotoxic properties and are thought to be major contributors to chronic allograft dysfunction in renal transplantation. Thus in modern immunosuppressive regimens, an attempt is made to minimize or even replace this drug class with less harmful yet equally effective agents.

Mechanistic Target of Rapamycin Inhibitors

Sirolimus (rapamycin) and everolimus bind to the same immunophilin as tacrolimus (FKBP12), although the complexes they form are unable to interact with calcineurin. Instead they bind to the regulatory kinase mTOR, which has a critical role in cytokine receptor signal transduction. The usual actions of mTOR are to activate the ribosomal enzyme p70 S6 kinase and block an inhibitory protein 4E-BP1, both of which are required for translation of proteins necessary for progression from the G_1 (growth) phase to the S (DNA synthesis) phase of the cell cycle. Therefore inhibition of this pathway in T cells blocks the action of cytokines, such as IL-2, -4, and -15, preventing cell cycle progression and clonal expansion.

In addition to inhibiting clonal expansion of effector T cells, there is growing evidence supporting the role of rapamycin in promoting the generation and survival of Tregs, which have been shown in several animal models to be capable of preventing allograft rejection.[56] For instance, rapamycin promotes the *de novo* conversion of alloantigen-specific Tregs.[44] This is in contrast to CNIs, which inhibit such conversion. This property makes rapamycin an attractive agent to potentially promote tolerance, as discussed further in the following section.

TOLERANCE

Billingham et al. first introduced the term *transplantation tolerance* in 1953, with the report that inoculation of fetal mice with lymphoid cells from an allogeneic adult donor mouse of a different strain led to later acceptance of skin grafts from the original skin graft donors,[57] and for their seminal work Medawar received the Nobel Prize for Medicine in 1960.

The condition in which an SOT recipient exhibits a well-functioning graft and lacks histological signs of rejection after receiving no immunosuppression for at least 1 year is referred to as *clinical operational tolerance* (COT).[58] Importantly, patients demonstrating COT must also be capable of responding to other non–transplantation-related immune challenges, including infections. Over recent years, experimental models have shown that it is possible to exploit the processes by which immune homeostasis and tolerance to self antigens are maintained to induce tolerance to alloantigen (Chapter 12). The optimal outcome for patients after transplantation would be harnessing of these mechanisms to induce specific immunological unresponsiveness or tolerance to the graft, thus avoiding the adverse side effects associated with current immunosuppressive regimens (see Table 81.1).

Unfortunately, more than 5 decades of clinical experience in SOT have demonstrated that COT in humans is extremely difficult to induce intentionally and may only be possible in a subset of transplant recipients. Many successful experimental techniques can produce durable hyporesponsiveness to mismatched allografts in rodent models, but scarcely any have been successfully translated into large animal models or clinical trials. Progress in the field of transplantation tolerance has also been hampered by the

THERAPEUTIC PRINCIPLES

Mechanisms of Tolerance

Deletion

Deletion of alloreactive or autoreactive T cells can be achieved centrally in the thymus or in the periphery. Infusion of donor bone marrow into a recipient who has been conditioned with nonmyeloablative irradiation or immunotherapy enables antigen-presenting cells (APCs) to access the thymus and trigger the deletion of maturing thymocytes. In the periphery, deletion can be triggered by alloantigen recognition under suboptimal conditions, including costimulation blockade, as well as immunodepletion by activation of apoptotic cell death and cytolysis.

Anergy

This is the functional inactivation of the T-cell response to restimulation by alloantigen or self antigen and has been described both *in vivo* and *in vitro*. Some forms of T-cell anergy are also reported to result in the development of regulatory activity. Costimulation blockade, as well as inhibition of downstream proliferative pathways, can trigger anergic states in T cells.

Immunoregulation

This active process results in the regulation of one cell population by the activity of another cell population. Various populations of leukocytes have been described as having the ability to control immune responsiveness to alloantigen stimulation in both innate and adaptive immune responses. This mechanism, although well described in experimental models, has yet to be introduced therapeutically in a sustained, clinical setting.

Clonal Exhaustion

This can occur as a result of chronic antigen stimulation or antigen recognition under suboptimal conditions. The consequence is either deletion or functional inactivation of the cells responding to the recognized antigen. An example of such exhaustion can be seen in liver transplantation, where the large number of donor-derived APCs migrating from the liver to draining lymphoid tissues after transplantation can trigger this type of response.

Ignorance

This is an uncommon mechanism in the induction phase of unresponsiveness to alloantigen as it is difficult to introduce donor cells or tissue without alerting the immune system to their presence, in transplantation. This mechanism, however, does describe the natural state of some self-reactive CD4 T-cell populations found in healthy individuals with no autoimmune pathology.

lack of definitive laboratory parameters able to give a clear indication of whether a particular recipient is tolerant of his or her graft. Furthermore, COT appears to be somewhat organ dependent. Recipients of a liver graft are more capable of developing COT because of the immune-privileged status of the liver (see Fig. 81.7). Clinical experience in liver transplantation now demonstrates clear evidence that a permanent and stable immunosuppressive-free state can be safely attempted and sometimes achieved in patients who have received a liver transplant for nonimmunological liver diseases.[59] However, COT has never been reported after intestinal, islet, or whole-organ pancreas transplantation, whereas two exceptional cases of COT have been described after lung and heart transplantations.[60,61] Since the first renal transplant over 55 years ago, only sporadic cases of COT have been documented after renal transplantation in the absence of genetic identity between donor and recipient.[59]

KEY CONCEPTS

Barriers to Transplantation Tolerance

- T-cell memory
  - Presensitization—direct exposure to alloantigen (*i.e.*, pregnancy or blood transfusion).
  - Heterologous immunity—cross-reactivity in the T-cell repertoire among antiviral, antibacterial, environmental, and transplantation antigens.
  - Homeostatic proliferation—induced by immunodepletory antibodies (*i.e.*, alemtuzumab).
  - Memory T cells generated by the above mechanisms—can result in rapidly formed effector immune responses upon rechallenge. These T cells are less sensitive to T-cell depleting antibodies and costimulatory blockade and thus may be more resistant to some tolerance induction strategies.
- B-cell response
  - Recipients treated with immunodepletory antibodies display a general increase in the naïve B-cell population.
  - There is a prevalent development of alloantibody in recipients treated with depleting antibody therapy.
  - Much focus of tolerogenic strategies is on the T-cell response. However, recent data suggest that the humoral immune system may play a more significant role than previously thought, possibly contributing to more long-term outcomes. Further work in this area continues.
- Lack of tolerance signature
  - An episode of acute rejection can severely affect graft survival in most transplanted organs.
  - There is absence of validated biomarkers of tolerance or predictors of rejection.
  - It is clinically difficult to justify often high-risk tolerizing strategies in patients who would otherwise do moderately well with standard immunosuppression.

In the next section, we discuss the current clinical strategies for tolerance induction as well as much-needed future approaches to produce more short-acting, antigen-specific agents that can optimize outcomes in the clinic.

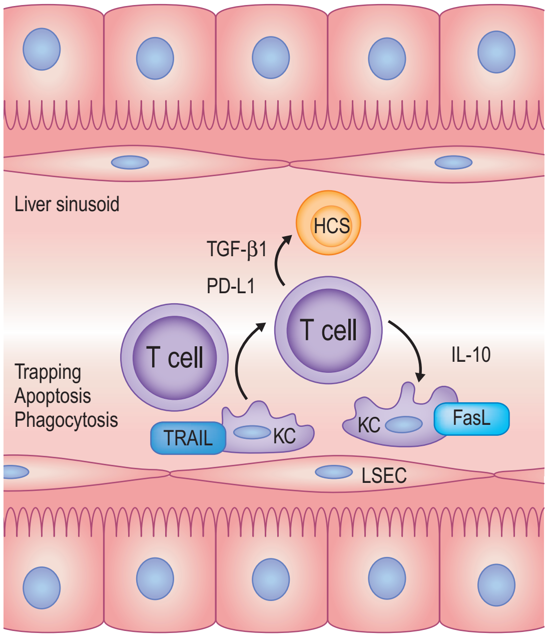
Molecule-Based Tolerogenic Protocols

The induction of COT through the administration of presumed tolerogenic drugs and subsequent immunosuppression minimization is one such strategy currently under investigation. Promising results of the first use of alemtuzumab in human subjects by Calne et al.[31,62] with low-dose cyclosporine alone as maintenance therapy introduced the concept of "*prope*" tolerance (*i.e.*, graft acceptance with reduced immunosuppression). Several groups have subsequently studied alemtuzumab induction and maintenance therapy without nephrotoxic CNI agents. Unfortunately, results of pilot studies have been poor, with high acute rejection rates (as much as 36% of recipients in one study). As discussed above, an attempt to induce donor allograft tolerance using alemtuzumab alone as an induction therapy with no maintenance immunosuppression at all resulted in all of the patients developing acute rejection within the first month after transplantation.[31] Encouragingly, physicians from the University of Wisconsin have demonstrated excellent long-term graft survival in 90% of patients maintained with only low-dose rapamycin monotherapy; these recipients had undergone alemtuzumab induction and an initial 60-day-only treatment with a CNI after transplantation to prevent acute rejection. It should be noted, however, that 50% of the patients developed alloantibody to either HLA class I or II

Tolerogenesis in the Liver

The liver is an important site for primary T-cell activation, but this takes place in an environment biased toward tolerance:

- Liver endothelium expresses adhesion molecules that facilitate the sequestration of circulating activated T cells, particularly CD8 T cells. This gives the liver a role in systemic immunoregulation.
- These activated T cells may undergo apoptosis owing to FasL and TRAIL expressed on Kupffer cells (KC) and may also be phagocytosed.
- Constitutive exposure of liver cells to traces of endotoxin and other microbial products via blood from the systemic circulation and intestine results in down-modulation of costimulatory molecules and the synthesis of IL-10 by Kupffer cells (KC) and liver sinusoidal endothelial cells (LSECS).
- Non haematopoietic liver cells, including LSECs and hepatic stellate cells (HSC) act as APCs, presenting antigen to T cells in the presence of immunosuppressive cytokines (IL-10 and TGF-β1) and inhibitory cell surface igands (PD-L1). Anergy, rather than activation is promoted. Thus immune responses to liver antigens can often result in tolerance.



Clinical Observations

- Positive cross-match or blood type incompatibility has little affect on graft survival
- HLA matching is not a prerequisite to liver transplantation
- Reduced incidence of hyperacute rejection ompared to other organs in SOT
- Frequent spontaneous recovery after severe rejection
- Acute rejection does not impact severely on long-term graft and patient survival
- Liver allograft protects other extra-hepatic allograft from rejection if derived from the same donor
- Lower overall incidence of chronic rejection that is reversible in up to 30% cases
- Clinical operational tolerance (COT) has been achieved most reliably and frequently in liver transplantation than with any other solid organ

FIG 81.7 Tolerogenesis in the Liver. The liver is an important site for primary T-cell activation, but this takes place in an environment biased toward tolerance. Liver endothelium expresses adhesion molecules that facilitate the sequestration of circulating activated T cells, particularly CD8 T cells. This gives the liver a role in systemic immunoregulation. These activated T cells may undergo apoptosis as a result of FasL and TRAIL expressed on Kupffer cells (KCs) and may also be phagocytosed. Constitutive exposure of liver cells to traces of endotoxin and other microbial products via blood from the systemic circulation and intestine results in downmodulation of costimulatory molecules and the synthesis of interleukin-10 (IL-10) by KCs and liver sinusoidal endothelial cells (LSECS). Nonhematopoietic liver cells, including LSECs and hepatic stellate cells, act as APCs, presenting antigen to T cells in the presence of immunosuppressive cytokines (IL-10 and transforming growth factor-β_1 [TGF-β_1]) and inhibitory cell surface ligands (PD-L1). Anergy, rather than activation, is promoted. Thus immune responses to liver antigens can often result in tolerance.

molecules and 40% demonstrated C4d$^+$ graft biopsies indicative of subacute humoral responses to the allograft.[39]

In 2003, investigators at the University of Pittsburgh published the results of a single center trial in which they administered an immunosuppressive regimen deemed to be tolerogenic to 82 adult kidney, liver, pancreas, and intestinal transplant recipients.[63] Their working hypothesis was that the need for continuous high-dose immunosuppression could be avoided in most cases with the use of a strong lymphocyte-depleting regimen before engraftment, followed by the administration of low-dose tacrolimus monotherapy. The goal of the induction treatment was the nonspecific removal of clones of immune cells responsible for rejection before contact with foreign donor antigens occurred while minimizing exposure to maintenance immunosuppression. At the time of publication of the study, 64% renal, 70% liver, 42% pancreas, and 54% intestinal transplant recipients were on spaced doses, but no patient could be weaned completely off immunosuppression. Nevertheless, the striking reduction in daily dosing was a step forward because it led to a significant reduction in overall immunosuppression-related morbidity.

It is important to note that other protocols based on a similar strategy have since been attempted, including post–liver transplantation protocols (the liver being an organ more capable of developing COT compared with any other). However, none of these protocols has achieved COT, nor have these protocols shown any impact on overall outcomes convincingly. In most situations, it appears that leukocyte depletion is not accompanied by a permanent and complete deletion of alloaggressive donor-reactive cells, and the establishment of a regulatory network is required to maintain tolerance.

Full Chimerism

The more robust experimental strategies for the induction of tolerance to foreign antigen utilize the mechanisms of central deletion to eliminate T-cell clones with specificity for the foreign antigens in question, thereby preventing them from entering the periphery. This can be reliably achieved by the establishment of hematopoietic chimerism through bone marrow transplantation (BMT). Stable engraftment of donor hematopoietic stem cells (HSCs) results in repopulation of the recipient thymus with donor-type thymic DCs, with the result that developing T cells with antidonor specificity are deleted by negative selection.

Full chimerism can be obtained rapidly through the ablation of the recipient's marrow and immune system with high-dose radiation and/or chemotherapy; in addition it can be induced more slowly by nonablative conditioning regimens, followed by the infusion of donor's marrow to colonize the recipient completely. This phenomenon paves the way for the onset of tolerance in the case of a subsequent SOT from the same donor.[64]

As proof of concept, numerous patients have undergone successful BMT for hematological indications and have subsequently been successfully transplanted with a kidney from the same donor, without the requirement for increased immunosuppression.

It is important to highlight that in all these cases, the use of BMT was justified on the basis of the need for treatment of hematological malignancies.

An alternative approach to achieving full chimerism and hence tolerance to donor alloantigens while avoiding GvHD or engraftment syndrome, in HLA-*mis*matched unrelated stem cell/renal transplant human recipients has been reported.[65] Results were achieved with the introduction of a mobilized cellular product enriched for tolerogenic graft facilitating cells (FCs) as well as HSCs and T cells (the treatment with HSCs in combination with FCs is termed *FCRx*) rather than just bone marrow graft alone. These bone marrow–derived FCs, which are $CD8^+$ but do not express a TCR, potently enhance engraftment of allogeneic HSCs in conditioned recipients. FCs are composed predominantly of a plasmocytoid precursor DC subpopulation, induce the generation of antigen-specific Tregs *in vitro* and *in vivo,* and have been found to effectively prevent GvHD in mice. In the first phase of this single-center clinical study, eight HLA-mismatched kidney transplant recipients were treated with FCRx and nonmyeloablative conditioning during the peritransplantation period. One year later, five recipients exhibited durable macrochimerism with no incidence of GvHD or engraftment syndrome, and were immunosuppression-free—that is, these patients were clinically operationally tolerant. *In vitro* studies indicated that chimeric donor lymphocytes were tolerized to the recipient, and a significant increase in the CD4 Treg/T effector cell population ratio was observed in these patients when compared with those who had lost chimerism. The ability to establish high levels of donor multilineage chimerism in haploidentical and highly mismatched unrelated donor–recipient pairs without the development of GvHD through the use of novel cellular therapies could have exciting therapeutic implications for disorders for which HSC transplantation can provide a "functional cure," including inherited metabolic disorders, hemoglobinopathies, and autoimmune disease, as well as in solid organ transplantation. Longer-term follow-up data on these initial patients as well as data from the enrollment of additional patients into this protocol continue to be encouraging. A phase III multicenter clinical trial is planned.

Mixed Chimerism

A promising approach is the induction of mixed hematopoietic chimerism, which can be achieved in experimental models with far less toxic induction therapy. Nonmyeloablative conditioning regimens have been applied in the clinical setting in patients who develop renal failure as a consequence of multiple myeloma. Further refinements of such protocols are being explored in experimental models. Examples of these include either a combination of depleting anti-CD4 and anti-CD8 antibodies together with mild, nonmyeloablative total body irradiation or costimulatory blockade with anti-CD154 and/or CTLA-4 Ig. When these induction protocols are followed by BMT, the result is mixed chimerism (the continued survival of both donor and recipient hematopoietic progenitor cells). Animals that have undergone these therapies have demonstrated durable tolerance to donor-type allografts and have a much lower incidence of GvHD compared with full chimeras.

In particular, the Massachusetts General Hospital (MGH) transplantation group showed in mouse and NHP models that nonmyeloablative regimens could result in transient mixed chimerism with accompanying long-term transplantation tolerance. Indeed, the same group has reported the long-term follow-up of human recipients, who underwent combined HLA-matched bone marrow and kidney transplantation for end-stage renal disease with or without multiple myeloma.[66,67]

Regulatory T Cells

There is now abundant evidence for the existence of populations of regulatory lymphocytes with the ability to suppress immune responses by other leukocytes (Chapter 18). Tregs can be divided into two populations: thymic-derived naturally occurring $CD25^+CD4^+$ cells (tTregs) and induced Tregs or iTregs that are either differentiated from $CD25^-CD4^+$ nonregulatory cells or expanded from $CD25^+CD4^+$ cells in response to antigen. Expression of the transcription factor FOXP3 is essential for the development and function of Tregs; however, tTregs and iTregs differ in origin, antigen experience, methylation patterns of FOXP3, and suppressive mechanisms. Both tTregs and iTregs have been demonstrated to play important roles in transplant tolerance.

tTregs develop within the thymus under the direction of FOXP3 and have a critical role in limiting immune responses to self antigens, as demonstrated by experimental models where mice depleted of $CD25^+CD4^+$ cells subsequently develop inflammatory bowel disease (IBD) and widespread autoimmune phenomena. Mutation of FOXP3 in humans is responsible for immune dysfunction/polyendocrinopathy/enteropathy/X-linked (IPEX) syndrome (Chapter 35). There is evidence for the involvement of Tregs in the downregulation of immune responses to tumors and chronic infections, as well as allogeneic transplants.[70]

Many strategies exist for the *in vivo* or *ex vivo* generation and/or expansion of Tregs. The most common *in vivo* approaches are based on the fact that stimulation with anti-CD3 and anti-CD28 in the presence of high concentrations of IL-2 stimulates proliferation and that exposure to antigen increases Treg frequency and/or potency by either expanding tTregs or inducing the generation of iTregs from cells that do not originally possess regulatory activity. There is a large body of data demonstrating that *ex vivo* expanded Tregs can protect allografts from rejection.[56,71]

Transplant recipients are treated with diverse immunosuppressive drug combinations, which may have a different impact on Tregs. As mentioned in the previous section, it has been demonstrated that CNIs, in particular cyclosporine, are detrimental to Tregs, whereas the mTOR inhibitor rapamycin was shown to be beneficial to Tregs both in terms of *in vivo* generation and function in mouse models and *in vitro* cultures of human Tregs. Adoptive transfer of a low number of alloantigen-specific Tregs under a cover of low-dose rapamycin was found to be capable of inducing long-term survival of heart transplants in an unmanipulated host, an outcome otherwise difficult to obtain.[72]

There are significant differences between human and mouse Tregs, in particular the differences in FOXP3 expression: In humans, FOXP3 is also expressed transiently by activated nonregulatory T cells that also upregulate CD25 expression. Thus not all $CD25^{+}FOXP3^{+}CD4^{+}$ cells will be genuine Tregs, and therefore isolation strategies based on $CD25^{hi}CD4^{+}$ are likely to be imperfect. Stable expression of FOXP3 in Tregs depends on DNA demethylation at the Treg-specific demethylated region (TSDR), a conserved, CpG-rich region within the FOXP3 locus. The TSDR is selectively demethylated in tTregs (which are more stable than pTregs), and therefore the methylation status of the TSDR region can act as a marker for the quality and stability of a Treg population. Cell-surface markers that can enrich for Tregs have also been described. In humans, $CD127^{lo}CD25^{+}CD4^{+}$ T cells are characterized by a higher level of FOXP3 expression and a more pronounced suppressive capacity.[70,71] Importantly, *ex vivo* expanded $CD25^{hi}CD4^{+}$ and $CD127^{lo}CD25^{+}CD4^{+}$ Tregs are effective at inhibiting vasculopathy in a humanized mouse model, whereas $CD127^{lo}CD25^{+}CD4^{+}$ T cells are five times more efficient than those not expressing a low level of CD127.[56]

A number of strategies for the isolation or enrichment of human Tregs have been described, but there is no consensus as to which strategy produces the optimal population for use in cell therapy applications. The most commonly used expansion protocol at present is based on stimulation by anti-CD3/anti-CD28 beads in the presence of high doses of recombinant IL-2, supplemented in some protocols with rapamycin. Despite the generation of sufficient numbers of Tregs for cellular therapy, this mode of expansion is antigen nonspecific without any enrichment steps for the cells of interest. The concept of expanding or generating donor alloantigen-reactive Tregs is much more appealing for clinical application in the setting of transplantation and a number of studies have shown that donor alloantigen-reactive Tregs are more potent on a cell-per-cell basis at controlling allograft rejection. Further work to ascertain the long-term stability and plasticity of Tregs *in vivo,* as well as the potential effects of Tregs in antitumor and antiviral immunities, as theoretically the infusion of large numbers of potent suppressive cells may compromise the immune response toward infectious pathogens and tumor cells.[73]

Several other immunomodulatory cell groups are currently under investigation as potential tolerogenic therapies, such as mesenchymal stromal cells (MSCs) and regulatory macrophages. MSCs were originally isolated from the bone marrow but can now be isolated from almost any human tissue, and they possess fascinating tissue repair and immunoregulatory properties. Preclinical models indicate that MSCs can promote engraftment of allogeneic cells, tissues, and organs as well as prevent and/or treat rejection. Although there are no published reports, to date, on their potential to induce tolerance in the setting of clinical SOT, their use as an induction therapy instead of anti-IL2R has demonstrated a lower incidence of acute rejection, decreased risk of opportunistic infection, and better estimated renal function at 1 year.[68] Regulatory macrophages have been used in two safety trials in renal transplantation. Overall, these trials demonstrated that infusion of regulatory macrophages is feasible and that there is the potential for immunosuppression minimization.[69] There are several ongoing clinical trials examining the application, safety, and efficacy of cellular therapies, including those for the prevention and treatment of GvHD in hematopoietic stem cell transplantation (HSCT) and the prevention of rejection in solid organ transplantation. (*e.g.,* the ONE Study; www.onestudy.org).[74]

Biomarkers of Rejection or Tolerance

Concerns have been raised regarding the evaluation of tolerance induction protocols, such as those using cellular therapies in recipients of kidney, heart, or pancreas, as an episode of acute rejection could severely affect graft function and survival. In these settings, most physicians are reluctant to withdraw immunosuppressive drugs in the absence of validated biomarkers of transplantation tolerance and rejection prediction. Biomarkers are thus needed to better evaluate the immune status of transplant recipients, determine the "tolerance signature" of recipients, predict and diagnose graft rejection noninvasively, and individualize immunosuppressive therapy.

Technical advances in multiparameter flow cytometry, antigen-specific lymphocyte assays, and genome-wide analyses have led to the development of powerful and more standardized immunomonitoring techniques to characterize alloimmune responses. Histological analysis remains the gold standard method by which to detect graft rejection, but graft biopsy is an invasive procedure with inherent risks and so cannot be routinely performed for surveillance purposes. It is also associated with sampling error and interpretation biases.

Molecular or cell-based biomarkers are being developed as alternative surveillance methods.[20,75] At present, it is notable that for the many biomarkers identified, there are discrepancies among data from different laboratories investigating the same parameter. This is likely caused by differences in recipient and donor populations, clinical management, and the laboratory techniques used to perform the assays. Therefore it is of great importance that the assays used to detect biomarkers are optimized, validated, and standardized and that candidate biomarkers are investigated in large patient cohorts from different transplantation centers using the same shared and validated assays.[76]

ON THE HORIZON

Initiatives in Clinical Utilization of Tolerance

- Clinical introduction of antigen-specific regulatory T-cell therapy
- Identification of tolerance signatures and predictors of rejection
- Use of immunomodulatory stem cell populations (*i.e.,* mesenchymal stromal cells)

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Merrill JP, Murray JE, Harrison JH, et al. Successful homotransplantation of the human kidney between identical twins. J Am Med Assoc 1956;160(4):277–82.
2. Salahudeen AK, Haider N, May W. Cold ischemia and the reduced long-term survival of cadaveric renal allografts. Kidney Int 2004;65(2):713–18.
3. Andrade CF, Waddell TK, Keshavjee S, et al. Innate immunity and organ transplantation: the potential role of toll-like receptors. Am J Transplant 2005;5(5):969–75.
4. Hart DN, Fabre JW. Kidney-specific alloantigen system in the rat. Characterization and role in transplantation. J Exp Med 1980;151(3):651–66.
5. Qureshi F, Rabb H, Kasiske BL. Silent acute rejection during prolonged delayed graft function reduces kidney allograft survival. Transplantation 2002;74(10):1400–4.

6. Whitelegg A, Barber LD. The structural basis of T-cell allorecognition. Tissue Antigens 2004;63(2):101–8.
7. Lakkis FG, Arakelov A, Konieczny BT, et al. Immunologic “ignorance” of vascularized organ transplants in the absence of secondary lymphoid tissue. Nat Med 2000;6(6):686–8.
8. Afzali B, Lombardi G, Lechler RI. Pathways of major histocompatibility complex allorecognition. Curr Opin Organ Transplant 2008;13(4):438–44.
9. Auchincloss H Jr, Sultan H. Antigen processing and presentation in transplantation. Curr Opin Immunol 1996;8(5):681–7.
10. Smyth LA, Herrera OB, Golshayan D, et al. A novel pathway of antigen presentation by dendritic and endothelial cells: implications for allorecognition and infectious diseases. Transplantation 2006;82(1 Suppl.):S15–18.
11. Clarkson MR, Sayegh MH. T-cell costimulatory pathways in allograft rejection and tolerance. Transplantation 2005;80(5):555–63.
12. Heidt S, Segundo DS, Chadha R, et al. The impact of Th17 cells on transplant rejection and the induction of tolerance. Curr Opin Organ Transplant 2010;15(4):456–61.
13. Feng G, Chan T, Wood KJ, et al. Donor reactive regulatory T cells. Curr Opin Organ Transplant 2009;14(4):432–8.
14. Gill RG. NK cells: elusive participants in transplantation immunity and tolerance. Curr Opin Immunol 2010;22(5):649–54.
15. Nath DS, Basha HI, Mohanakumar T. Antihuman leukocyte antigen antibody-induced autoimmunity: role in chronic rejection. Curr Opin Organ Transplant 2010;15(1):16–20.
16. Rosenberg AS. The T cell populations mediating rejection of MHC class I disparate skin grafts in mice. Transpl Immunol 1993;1(2):93–9.
17. Cramer DV, Wu GD, Chapman FA, et al. Lymphocytic subsets and histopathologic changes associated with the development of heart transplant arteriosclerosis. J Heart Lung Transplant 1992;11(3 Pt 1):458–66.
18. Gebel HM, Bray RA. The evolution and clinical impact of human leukocyte antigen technology. Curr Opin Nephrol Hypertens 2010;19(6):598–602.
19. Leffell MS, Zachary AA. Antiallograft antibodies: relevance, detection, and monitoring. Curr Opin Organ Transplant 2010;15(1):2–7.
20. Sanchez-Fueyo A, Strom TB. Immunologic basis of graft rejection and tolerance following transplantation of liver or other solid organs. Gastroenterology 2011;140(1):51–64.
21. de Fijter JW. Rejection and function and chronic allograft dysfunction. Kidney Int Suppl 2010;119:S38–41.
22. Veronese FV, Noronha IL, Manfro RC, et al. Prevalence and immunohistochemical findings of subclinical kidney allograft rejection and its association with graft outcome. Clin Transplant 2004;18(4):357–64.
23. Shapiro R, Basu A, Tan H, et al. Kidney transplantation under minimal immunosuppression after pretransplant lymphoid depletion with Thymoglobulin or Campath. J Am Coll Surg 2005;200(4):505–15, quiz A59–61.
24. Deeks ED, Keating GM. Rabbit antithymocyte globulin (thymoglobulin): a review of its use in the prevention and treatment of acute renal allograft rejection. Drugs 2009;69(11):1483–512.
25. Charpentier B, Rostaing L, Berthoux F, et al. A three-arm study comparing immediate tacrolimus therapy with antithymocyte globulin induction therapy followed by tacrolimus or cyclosporine A in adult renal transplant recipients. Transplantation 2003;75(6):844–51.
26. Mourad G, Garrigue V, Squifflet JP, et al. Induction versus noninduction in renal transplant recipients with tacrolimus-based immunosuppression. Transplantation 2001;72(6):1050–5.
27. Morris PJ, Russell NK. Alemtuzumab (Campath-1H): a systematic review in organ transplantation. Transplantation 2006;81(10):1361–7.
28. Trzonkowski P, Zilvetti M, Chapman S, et al. Homeostatic repopulation by CD28-CD8+ T cells in alemtuzumab-depleted kidney transplant recipients treated with reduced immunosuppression. Am J Transplant 2008;8(2):338–47.
29. Pearl JP, Parris J, Hale DA, et al. Immunocompetent T-cells with a memory-like phenotype are the dominant cell type following antibody-mediated T-cell depletion. Am J Transplant 2005;5(3):465–74.
30. Lopez M, Clarkson MR, Albin M, et al. A novel mechanism of action for anti-thymocyte globulin: induction of $CD4^+CD25^+Foxp3^+$ regulatory T cells. J Am Soc Nephrol 2006;17(10):2844–53.
31. Calne R, Friend P, Moffatt S, et al. Prope tolerance, perioperative campath 1H, and low-dose cyclosporin monotherapy in renal allograft recipients. Lancet 1998;351(9117):1701–2.
32. Watson CJ, Bradley JA, Friend PJ, et al. Alemtuzumab (CAMPATH 1H) induction therapy in cadaveric kidney transplantation—efficacy and safety at five years. Am J Transplant 2005;5(6):1347–53.
33. Wu Z, Bensinger SJ, Zhang J, et al. Homeostatic proliferation is a barrier to transplantation tolerance. Nat Med 2004;10(1):87–92.
34. Trzonkowski P, Zilvetti M, Friend P, et al. Recipient memory-like lymphocytes remain unresponsive to graft antigens after CAMPATH- 1H induction with reduced maintenance immunosuppression. Transplantation 2006;82(10):1342–51.
35. Hanaway MJ, Woodle SE, Mulgaonkar S, et al. Alemtuzumab Induction in Renal Transplantation. NEJM 2011;364(20):1909–19.
36. The 3C Study Collaborative Group, Haynes R, Harden P, et al. Alemtuzumab-based induction treatment versus basiliximab-based induction treatment in kidney transplantation (the 3C Study): a randomised trial. Lancet 2014;384(9955):1684–90.
37. Shyu S, Dew MA, Pilewski JM, et al. Five-year outcomes with alemtuzumab induction after lung transplantation. J Heart Lung Transplant 2011;30(7):743–54.
38. Willicombe M, Brookes P, Santos-Nunez E, et al. Outcome of patients with preformed donor-specific antibodies following alemtuzumab induction and tacrolimus monotherapy. Am J Transplant 2011;11(3):470–7.
39. Knechtle SJ, Pascual J, Bloom DD, et al. Early and limited use of tacrolimus to avoid rejection in an alemtuzumab and sirolimus regimen for kidney transplantation: clinical results and immune monitoring. Am J Transplant 2009;9(5):1087–98.
40. Bloom D, Chang Z, Pauly K, et al. BAFF is increased in renal transplant patients following treatment with alemtuzumab. Am J Transplant 2009;9(8):1835–45.
41. Chatenoud L, Bluestone JA. CD3-specific antibodies: a portal to the treatment of autoimmunity. Nat Rev Immunol 2007;7(8):622–32.
42. Ablamunits V, Bisikirska B, Herold KC. Acquisition of regulatory function by human CD8(+) T cells treated with anti-CD3 antibody requires TNF. Eur J Immunol 2010;40(10):2891–901.
43. Golshayan D, Pascual M. Tolerance-inducing immunosuppressive strategies in clinical transplantation: an overview. Drugs 2008;68(15):2113–30.
44. Knechtle SJ. Immunoregulation and tolerance. Transplant Proc 2010;42(9 Suppl.):S13–15.
45. Tang Q, Henriksen KJ, Boden EK, et al. Cutting edge: CD28 controls peripheral homeostasis of CD4+CD25+ regulatory T cells. J Immunol 2003;171(7):3348–52.
46. Bluestone JA, Liu W, Yabu JM, et al. The effect of costimulatory and interleukin 2 receptor blockade on regulatory T cells in renal transplantation. Am J Transplant 2008;8(10):2086–96.
47. Vincenti F, Charpentier B, Vanrenterghem Y, et al. A phase III study of belatacept-based immunosuppression regimens versus cyclosporine in renal transplant recipients (BENEFIT study). Am J Transplant 2010;10(3):535–46.
48. Larsen CP, Grinyó J, Medina-Pestana J, et al. Belatacept-based regimens versus a cyclosporine A-based regimen in kidney transplant recipients: 2-year results from the BENEFIT and BENEFIT-EXT studies. Transplantation 2010;90(12):1528–35.
49. Vincenti F, Rostaing L, Grinyo J, et al. Belatacept and Long-Term Outcomes in Kidney Transplantation. NEJM 2016;374(4):333–43.
50. Vincenti F, de Andrés A, Becker T, et al. Interleukin-2 receptor antagonist induction in modern immunosuppression regimens for renal transplant recipients. Transpl Int 2006;19(6):446–57.
51. Webster AC, Playford EG, Higgins G, et al. Interleukin 2 receptor antagonists for renal transplant recipients: a meta-analysis of randomized trials. Transplantation 2004;77(2):166–76.

52. Keown P, Balshaw R, Khorasheh S, et al. Meta-analysis of basiliximab for immunoprophylaxis in renal transplantation. Biodrugs 2003;17(4):271–9.
53. Brennan DC, Schnitzler MA. Long-term results of rabbit antithymocyte globulin and basiliximab induction. N Engl J Med 2008;359(16):1736–8.
54. McKeage K, McCormack PL. Basiliximab: a review of its use as induction therapy in renal transplantation. Biodrugs 2010;24(1):55–76.
55. Carvalho-Gaspar M, Jones ND, Luo S, et al. Location and time-dependent control of rejection by regulatory T cells culminates in a failure to generate memory T cells. J Immunol 2008;180(10):6640–8.
56. Nadig SN, Wieckiewicz J, Wu DC, et al. In vivo prevention of transplant arteriosclerosis by ex vivo-expanded human regulatory T cells. Nat Med 2010;16(7):809–13.
57. Billingham RE, Brent L, Medawar PB. Actively acquired tolerance of foreign cells. Nature 1953;172(4379):603–6.
58. Ashton-Chess J, Giral M, Brouard S, et al. Spontaneous operational tolerance after immunosuppressive drug withdrawal in clinical renal allotransplantation. Transplantation 2007;84(10):1215–19.
59. Orlando G, Soker S, Wood K. Operational tolerance after liver transplantation. J Hepatol 2009;50(6):1247–57.
60. Tzakis AG, Reyes J, Zeevi A, et al. Early tolerance in pediatric liver allograft recipients. J Pediatr Surg 1994;29(6):754–6.
61. Mazariegos GV, Sindhi R, Thomson AW, et al. Clinical tolerance following liver transplantation: long term results and future prospects. Transpl Immunol 2007;17(2):114–19.
62. Calne R, Moffatt SD, Friend PJ, et al. Campath IH allows low-dose cyclosporine monotherapy in 31 cadaveric renal allograft recipients. Transplantation 1999;68(10):1613–16.
63. Starzl TE, Murase N, Abu-Elmagd K, et al. Tolerogenic immunosuppression for organ transplantation. Lancet 2003;361(9368):1502–10.
64. Orlando G, Hematti P, Stratta RJ, et al. Clinical operational tolerance after renal transplantation: current status and future challenges. Ann Surg 2010;252(6):915–28.
65. Leventhal J, Abecassis M, Miller J, et al. Chimerism and tolerance without GVHD or engraftment syndrome in HLA-mismatched combined kidney and hematopoietic stem cell transplantation. Sci Transl Med 2012;4(124):124ra128.
66. Spitzer TR, Sykes M, Tolkoff-Rubin N, et al. Long-term follow-up of recipients of combined human leukocyte antigen-matched bone marrow and kidney transplantation for multiple myeloma with end-stage renal disease. Transplantation 2011;91(6):672–6.
67. Kawai T, Sachs DH, Sprangers B, et al. Long-term results in recipients of combined HLA-mismatched kidney and bone marrow transplantation without maintenance immunosuppression. Am J Transplant 2014;14(7):1599–611.
68. Tan J, Wu W, Xu X, et al. Induction therapy with autologous mesenchymal stem cells in living-related kidney transplants: a randomized controlled trial. JAMA 2012;307(11):1169–77.
69. Hutchinson JA, Riquelme P, Sawitzki B, et al. Cutting Edge: immunological consequences and trafficking of human regulatory macrophages administered to renal transplant recipients. J Immunol 2011;187(5):2072–8.
70. Juvet SC, Whatcott A, Bushell AR, et al. Harnessing regulatory T cells for clinical use in transplantation: the end of the beginning. Am J Transplant 2014;14(4):750–63.
71. Putnam AL, Brusko TM, Lee MR, et al. Expansion of human regulatory T-cells from patients with type 1 diabetes. Diabetes 2009;58(3):652–62.
72. Raimondi G, Sumpter TL, Matta BM, et al. Mammalian target of rapamycin inhibition and alloantigen-specific regulatory T cells synergize to promote long-term graft survival in immunocompetent recipients. J Immunol 2010;184(2):624–36.
73. DuPage M, Bluestone JA. Harnessing the plasticity of CD4+ T cells to treat immune-mediated disease. Nat Rev Immunol 2016;16(3): 149–63.
74. Trzonkowski P, Bachetta R, Battaglia M, et al. Hurdles in therapy with regulatory T cells. Sci Transl Med 2016;7(304):1–18.
75. Banham GD, Clatworthy MR. B-cell biomarkers in transplantation – from genes to therapy. Tissue Antigens 2016;85(2):82–92.
76. Poste G. Bring on the biomarkers. Nature 2011;469(7329):156–7.

MULTIPLE-CHOICE QUESTIONS

1. The transfer of an organ between genetically identical individuals is termed:
   A. An autograft
   B. An allograft
   C. A xenograft
   D. An isograft
   E. A haplograft

2. The proportion of T cells that are capable of directly recognizing allogeneic major histocompatibility complex–peptide complexes is:
   A. Low before transplantation
   B. Only the memory T-cell subpopulation
   C. About the same as the frequency directed against a nominal antigen
   D. Zero given that T cells are self-restricted
   E. Approaching 10% of T cells

3. What mechanism is responsible for the presence of memory T cells reactive against a graft in a patient receiving his or her first transplant?
   A. Previous blood transfusion
   B. Previous pregnancy
   C. Heterologous immunity
   D. Direct allorecognition
   E. All of the above

82

Immune Reconstitution Therapy for Immunodeficiency

Luigi D. Notarangelo, Sung-Yun Pai

Following the discovery of the major human leukocyte antigens (HLAs) in 1958, hematopoietic stem cell transplantation (HSCT) came into practice to provide treatment for a variety of congenital and acquired disorders. Cure of an infant with severe combined immunodeficiency (SCID) in 1968 was the first successful experience in HSCT, marking the beginning of a new era in medicine. Shortly thereafter, partial success was achieved also with HSCT in a child with Wiskott-Aldrich syndrome (WAS).

For many years, the successful use of HSCT in severe primary immune deficiency (PID) was largely restricted to transplantation from HLA-matched related donors (MRDs), as HSCT from mismatched related family donors (MMRDs), usually represented by haploidentical parents, was followed by severe complications, graft-versus-host disease (GvHD) in particular. In the late 1970s, it was demonstrated in animal models that removal of mature T lymphocytes from the graft obtained from mismatched marrow allowed successful reconstitution upon injection into lethally irradiated recipient animals. This important achievement opened the way to a broader use of HSCT in severe forms of PID. More recently, transplantation of hematopoietic stem cells (HSCs) from volunteer matched unrelated donors (MUDs) and of umbilical cord blood (UCB) has been increasingly used in individuals with PID. Overall, since 1968, over 2000 transplantations have been performed in patients with PID, most of them in children with SCID.[1,2] The increasing number of transplantations over the years reflects an increased awareness of PID, the broader availability of diagnostic tools (including newborn screening for SCID), improved outcomes as a result of advances in supportive and critical care before and after HSCT, increasingly improved strategies to prevent GvHD, and the greater availability of MUD and UCB for transplantation.

HEMATOPOIETIC STEM CELL TRANSPLANTATION: GENERAL CONSIDERATIONS

Sources of Hematopoietic Stem Cells for Transplantation

Several sources of HSCs are available for transplantation (Table 82.1). HSCs can be retrieved from bone marrow, peripheral blood, or UCB (Chapter 2). Bone marrow stem cells are most commonly obtained by multiple aspirations along the iliac crests, usually while the donor is under general anesthesia. The volume of bone marrow that is obtained may vary between 500 mL and 1 L or even more, depending on the type of transplant (HLA-identical or T cell–depleted) and the weight of the recipient. Blood group matching for ABO antigens is not required for HSCT, as mature red blood cells (RBCs) or anti-ABO antibodies can be removed by RBC depletion and plasma depletion, respectively. In the case of HLA-identical transplantation, marrow stem cells are then injected intravenously without further manipulation into a central line in the recipient. In the case of mismatched transplantation, bone marrow cells are usually T cell depleted *in vitro* (see below); stem cells are then enumerated and injected intravenously. More recently, injection of unmanipulated bone marrow cells from MMRD, followed by *in vivo* administration of cyclophosphamide to prevent GvHD, has been also used.[3]

HSCs can also be retrieved from peripheral blood, following *in vivo* administration of granulocyte–colony-stimulating factor (G-CSF) to the donor, usually at the dosage of 10 μg/kg/day for 5 days, which allows mobilization of stem cells. In this case, stem cells can also be purified by positive selection (see below), enumerated, and injected.

Finally, UCB is another rich source of HSCs. At birth, cord blood is collected in a heparinized medium and stored in liquid nitrogen, with small aliquots preserved for HLA typing. Whenever sufficient compatibility is identified between a patient and stored cord blood, the latter is thawed and injected into the recipient without further manipulation. With this procedure, the number of HSCs that can be transplanted is dictated by their concentration in the cord blood sample and by the volume of the sample itself. More recently, *in vitro* expansion of cord blood stem cells, and transplantations with multiple cord blood units, have been attempted to overcome this limitation.

Donor Selection and Manipulation of the Graft

HSCT From a Related HLA-Identical Donor

The use of unfractionated stem cells from an HLA-identical sibling offers the best chance of rapid engraftment and immune reconstitution. In such cases, the HLA-identity between recipient and donor minimizes the risk of GvHD. Furthermore, the mature T cells contained in the graft provide a first line of immune reconstitution after transplantation, as they may expand and lead to a rapid increase in the number of circulating T lymphocytes as early as 2 weeks after HSCT. Engraftment of these mature T cells transplanted with the bone marrow of an MRD is particularly important in infants with SCID, as it is early evidence of immune reconstitution in a severely immunocompromised host.

TABLE 82.1 Sources of Hematopoietic Stem Cells for Transplantation

Hematopoietic Stem Cell Transplantation (HSCT) From a Related Donor

- Bone marrow from a human leukocyte antigen (HLA)–genotypically identical sibling
- Bone marrow from an HLA-phenotypically identical family member
- Bone marrow from a haploidentical parent
  - T cell depleted by negative selection with soybean lectin agglutination and rosetting with sheep red blood cells (E-rosetting)
  - T cell depleted by negative selection with monoclonal antibodies
  - Positive selection of CD34$^+$ cells

HSCT From a Matched Unrelated Donor

- Unmanipulated bone marrow
- T cell–depleted bone marrow by means of positive selection of CD34$^+$ cells
- Positively selected peripheral blood CD34$^+$ cells

HSCT From an Unmanipulated Related or Unrelated Cord Blood

KEY CONCEPTS

Sources of Stem Cells and Selection of Donors for Hematopoietic Stem Cell Transplantation in Primary Immunodeficiencies

- Sources of hematopoietic stem cells (HSCs) for transplantation include bone marrow, peripheral blood, and cord blood.
- If the donor is a genotypically human leukocyte antigen (HLA)–identical sibling, unmanipulated bone marrow is used as source of stem cells.
- Whenever the donor is HLA-mismatched to the recipient, T-cell depletion must be performed to eliminate mature T lymphocytes from the graft. Methods for T-cell depletion of the bone marrow include use of soybean lectin agglutination and E-rosetting, depletion with monoclonal antibodies (mAbs), and positive selection of stem cells.
- Cord blood is a rich source of stem cells. However, the volume of cord blood is limited, so its use is mainly restricted to young patients.
- The number of volunteers included in Bone Marrow Donor Registries is expanding. Consequently, there is a continuous increase in the number of matched unrelated donor (MUD) transplantations performed for patients with primary immunodeficiencies.
- Whenever cord blood or MUD stem cells are used, conditioning regimen must be given to the recipient before transplantation to facilitate engraftment of donor stem cells.
- Therapeutic options for patients with severe combined immunodeficiency (SCID) and other severe forms of primary immunodeficiency (PID) include transplantation from HLA-genotypically identical donor, mismatched related donors (MMRD), and MUDs.

HSCT From a Haploidentical Donor

Unfortunately, the option of related HLA-identical HSCT is limited only to a minority of patients. When no such donor is available, stem cell transplantation from a haploidentical parent should be considered, particularly in infants with SCID.

The rationale for haploidentical HSCT is based on the ability of donor-derived stem cells to repopulate the recipient's vestigial thymus and give rise to fully mature T lymphocytes. Indeed, this is a life-saving procedure that has been successfully applied to several hundreds of infants with SCID.[1,2]

However, this procedure requires careful removal of T lymphocytes from the graft, as these would otherwise cause severe GvHD. Several methods are available to attain T-cell depletion.

In the past, the method of soybean lectin agglutination and E-rosetting was frequently used. With this method, soybean lectin allowed agglutination of the majority of mature marrow cells, which were removed by sedimentation. Further depletion of T lymphocytes was achieved by rosetting with sheep erythrocytes (E-rosetting technique) and density gradient centrifugation. Importantly, T-cell depletion by soybean lectin agglutination and E-rosetting maintains all immature marrow cells in the final preparation.

T-cell depletion can also be achieved by incubation of bone marrow with monoclonal antibodies (mAbs) to T lymphocytes plus complement. Campath-1 G, Leu 1, and other mAbs have been used for this purpose, but the degree of T-cell depletion that is achieved with these agents is less effective than with the soybean lectin and E-rosetting, and therefore a higher incidence of GvHD has been reported. More recently, use of mAbs directed against the αβ form of the T-cell receptor (TCR) and against CD19 has entered clinical practice, with excellent results.[4]

However, the most common method to obtain transplantable HSCs is represented by positive selection of CD34$^+$ cells using mAbs. This method allows very robust depletion of mature T cells; however, it also removes immature CD34$^-$ cells and other cells (especially stromal marrow cells) that can facilitate stem cell engraftment.

Selection of the best donor is another important aspect of T cell–depleted haploidentical HSCT for SCID. In general, the donor is represented by one of the parents, since the volume of bone marrow that can be collected is much higher than it would be if a haploidentical sibling were to serve as donor. Maternal T-cell engraftment *in utero* is a common finding in infants with SCID and has been observed in almost 40% of patients with SCID. In such cases, T cell–depleted HSCT should be performed using the mother as donor, if possible, since transplantation from the father might cause a graft-versus-graft reaction.

In Utero Haploidentical HSCT

The identification of a growing number of immunodeficiency-causing genes has resulted in continuous improvement in prenatal diagnosis, which, in most cases of severe immunodeficiency, can now be accomplished on chorionic villi DNA at 10–11 weeks of gestation. This has prompted prenatal transplantation of parental positively selected CD34$^+$ stem cells into the peritoneum of fetuses prenatally diagnosed with SCID, under ultrasound guidance.

The rationale underlying *in utero* HSCT is based on a lower risk of graft rejection resulting from decreased fetal immunocompetence (although this consideration is not relevant in the case of fetuses with SCID), a presumed induction of tolerance to paternal antigens (which might favor successful engraftment after postnatal transplantation from the same donor), the predicted competition between donor and autologous stem cells at a time when several empty niches should be available for stem cell engraftment, the potential ability to provide preemptive treatment (thus reducing the risk of postnatal infection), and the lower cost of the procedure that does not require prolonged hospitalization. However, *in utero* HSCT is associated with the potential risks of fetal loss and of GvHD. Finally, if maternal T cells had engrafted into the fetus with SCID, transplantation of paternal CD34$^+$ cells might cause graft-versus-graft reaction.

In a series of five *in utero* HSCT performed at the University of Brescia, Italy, all three cases with B^+ SCID survived, with evidence of T-cell reconstitution. However, transient or no immune reconstitution was observed in two fetuses with B^- SCID. The use of *in utero* HSCT for SCID is now largely unjustified, also in consideration of the fact that newborn screening allows rapid identification of infants with SCID and permits rapid referral to HSCT, with long-term survival that exceeds 90%.[5]

HSCT From Matched Unrelated Donors

Since the first successful HSCT performed in 1977 in an infant with SCID, HSCT from MUD has been increasingly used to treat severe PID. MUD HSCT has been shown to be more effective than T cell–depleted HSCT in patients with immunodeficiencies other than SCID,[1] and it has been successfully used also in infants with SCID.[6]

Transplantation from MUD has been facilitated by the increasing number of volunteer donors included in registries worldwide. In addition, advances in the quality of the techniques used for HLA typing permits this identification of an optimal MUD and reduction in the risk of GvHD. As of April 2016, more than 28 million donor volunteers and UCB units were included in the Bone Marrow Donors Worldwide (BMDW) registry. At present, it takes only a few weeks to identify a MUD. However, the probability of finding a suitable donor is lower for selected ethnic or racial groups that are poorly represented among volunteer donors.

Importantly, MUD HSCT requires use of a preparative chemotherapy regimen in the recipient (even in the case of SCID) and GvHD prophylaxis (because of likely disparity between donor and recipient at minor histocompatibility loci), whereas neither one is necessary for related HLA-identical HSCT in SCID infants.

HSCT Using Unmanipulated Cord Blood

As opposed to MUD HSCT, which requires identification, willingness, and medical clearance of an adult volunteer, stored cord blood is readily available as a source of stem cells for transplantation. In addition, the risk of GvHD at any given degree of HLA matching is lower when using cord blood compared with MUD HSCT so that greater HLA disparity with the recipient can be tolerated. However, the number of cells contained in any defined unit is still a major limitation of cord blood. Low cell dose is not usually a problem for transplants performed in infants with SCID or other severe forms of immune deficiency because of the low weight of the recipient. Indeed, unrelated umbilical cord stem cell transplantation has been successfully used in dozens of patients with severe PID.[7] In practice, an unrelated HSC donor should be simultaneously searched for in cord blood banks and in bone marrow donor registries for patients lacking an HLA-identical sibling HSC donor. The option of performing cord blood transplants should be based on urgency of the transplantation, the cell dose required, and the number of HLA disparities.

As for HSCT from MUD, transplantation using cord blood usually requires pretransplantation conditioning and GvHD prophylaxis, irrespective of the underlying disease.

Complications of Hematopoietic Stem Cell Transplantation

A variety of complications can compromise the success of HSCT. Among these, incompatibility between donor and recipient can lead to graft rejection by the host immune system or to GvHD caused by alloreactivity of donor-derived lymphocytes to the recipient's cells. Furthermore, conditioning regimens can cause toxicity of several organs. Myeloablative regimens cause anemia, thrombocytopenia, and leukopenia. Consequently, supportive treatment with RBC and platelet transfusions is necessary during the aplastic phase. Finally, the leukopenia predisposes the patient to an increased risk of life-threatening bacterial or fungal infections.

The frequency and severity of these complications depend on the type of transplant, the possible use of a conditioning regimen, and specific considerations related to the underlying disorder and to the clinical status of the recipient before transplantation.

KEY CONCEPTS

Rejection of Donor Stem Cells by the Host Immune System and Reaction of Donor's T Lymphocytes to the Host

The host T cells are responsible for the elimination of donor stem cells, regardless of the degree of human leukocyte antigen (HLA) compatibility. Under these circumstances, other than in patients with severe combined immunodeficiency (SCID), a chemotherapy-based conditioning regimen must be used before transplantation to allow for donor stem cell engraftment.

If T cells are present in the graft, and especially if there is HLA-incompatibility between the donor and the recipient, the donor's T lymphocytes may react to host alloantigens and cause graft-versus-host disease (GvHD).

Risk factors for GvHD include HLA-mismatch between donor and recipient, older age at transplantation, gender mismatch, and previous viral infections.

GvHD may develop early after transplantation (acute GvHD), or at 100 days or more after transplant (chronic GvHD). GvHD is one of the major causes of death and long-term disability after stem cell transplantation.

Prevention is the best approach to the management of GvHD. Prevention of GvHD is based on selection of optimal donor and vigorous T-cell depletion in the case of HLA mismatch between the donor and the recipient. Immunosuppressive drugs, such as cyclosporine, are another potent form of GvHD prevention.

Graft Rejection

Graft rejection reflects the presence of immunocompetent cells in the host that specifically recognize and react to donor-derived stem cells. Several factors influence the likelihood of graft rejection, in particular *(i)* the degree of immunocompetence of the host; *(ii)* the degree of HLA disparity between donor and recipient; *(iii)* the number and source of stem cells infused; *(iv)* the type of conditioning regimen used; and *(v)* the possible presensitization of the host to donor histocompatibility antigens.

In the case of infants with SCID, graft rejection is an unlikely event because of the virtual lack of T cells that characterizes these conditions. However, natural killer (NK) cells are present in several forms of SCID and may contribute to graft rejection. Indeed, among infants with SCID who received HSCT without conditioning chemotherapy, reduced overall survival and increased requirement of additional procedures have been observed in those with NK^+ SCID compared with NK^- SCID (Fig. 82.1).[8] In other forms of PID, there is sufficient immune function in the host to allow for rejection of donor-derived stem cells, unless an appropriate conditioning regimen is used before HSCT.

In children with nonmalignant disease, the most commonly used myeloablative conditioning regimen consists of busulfan

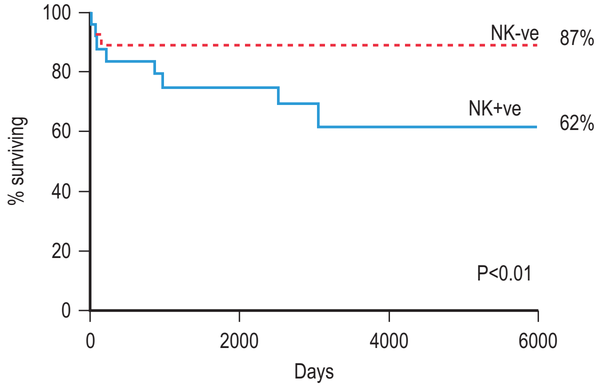


FIG 82.1 Kaplan-Meyer survival curve after nonconditioned transplantation for infants with severe combined immunodeficiency (SCID) with normal natural killer (NK) cell count (NK$^+$) or with lack or markedly reduced number of NK cells (NK$^-$). (With permission from Hassan A, Lee P, Maggina P, et al. Host natural killer immunity is a key indicator of permissiveness for donor cell engraftment in patients with severe combined immunodeficiency. J Allergy Clin Immunol 2014;133:1660–6.)

TABLE 82.2 Staging and Grading of Acute Graft-Versus-Host Disease (aGvHD)

Stage	Skin	Liver (Bilirubin)	Gastrointestinal system (GI) (Stool output)
0	None	<2 mg/dL	Adult: <500 mL/day Child: <10 mL/kg/day
1	Rash <25% BSA	2–3 mg/dL	Adult: 500–999 mL/day Child: 10–19.9 mL/kg/day or persistent nausea, vomiting, or anorexia, with a positive upper GI biopsy
2	Rash 25–50% BSA	3–6 mg/dL	Adult: 1000–1500 mL/day Child: 20–30 mL/kg/day
3	Rash >50% BSA (generalized erythroderma)	6.1–15 mg/dL	Adult: >1500 mL/day Child: >30 mL/kg/day
4	Generalized erythroderma with bullous formation	>15 mg/dL	Severe abdominal pain with or without ileus

Grade

Grade 0	No stage 1–4 of any organ
Grade 1:	Stage 1 or 2 skin involvement; no liver or gut involvement
Grade 2:	Stage 1–3 skin involvement; Grade 1 liver or gut involvement
Grade 3:	Stage 2 or 3 skin involvement
Grade 4:	Stage 1–4 skin involvement; stage 2–4 liver or gut involvement

and fludarabine, with or without the addition of antithymocyte globulin (ATG) or anti-CD52 (alemtuzumab) mAb. However, infants and children with preexisting organ damage are highly sensitive to the toxic effects of drugs. In these cases, reduced intensity regimens are preferred. To this purpose, lower doses of busulfan are used. Treosulfan is an alternative to busulfan and has reduced toxicity.[9] Finally, the combination of fludarabine and melphalan is another common form of reduced-intensity conditioning.

Acute Graft-Versus-Host Disease

Acute GvHD (aGvHD) is the result of alloreactivity of donor-derived T lymphocytes versus the recipient's antigens and is one of the most severe complications of HSCT. It may occur as early as 1 week after HSCT and is potentially fatal. Clinical manifestations of aGvHD include maculopapular skin rash (that tends to be confluent), diarrhea, and liver abnormalities (hepatomegaly, elevated liver enzymes, increased levels of conjugated bilirubin).[10] The disease may progress to severe skin manifestations, with exfoliative dermatitis, and significant liver and gut damage (with intractable watery or bloody diarrhea, protein-losing enteropathy, and abdominal pain). In the most severe cases, leakage of intravascular fluids into the interstitium (so-called third space filling) leads to generalized edema. Bone marrow aplasia and a high susceptibility to infections (including reactivation of herpesvirus infections) are also often observed in severe aGvHD.

The severity of aGvHD is evaluated according to grading (Table 82.2). Major risk factors for aGvHD include HLA-mismatch between donor and recipient, older age of the recipient or donor, gender mismatch, and stem cell source.[10] However, aGvHD may also be observed following related HLA-identical HSCT, particularly when a conditioning regimen is used.

Finally, transfusion-associated aGvHD is a very severe complication after HSCT, which can be effectively prevented by using irradiated (1500–3000 rad) and filtered blood derivatives.

Chronic Graft-Versus-Host Disease

Chronic GvHD (cGvHD) has traditionally been defined as symptoms that persist or appear after 100 days since the time of transplantation. As the types of donor, stem cell sources, and conditioning regimens have changed, it is now recognized that symptoms of aGvHD can present later than 100 days and that cGvHD is better defined by its distinct clinical manifestations rather than the time of onset alone. These clinical manifestations include skin changes (scleroderma-like lesions, hyperpigmentation, hyperkeratosis, skin atrophy, ulcerations), tissue fibrosis and limitation of joint motility, fibrosis of exocrine glands ("sicca syndrome"), fibrosis of lungs and liver, increased susceptibility to infections, immune dysregulation, and autoimmunity.[10] Consequently, cGvHD poses a major burden on the patients' quality of life and can be fatal.

Although the incidence of cGvHD is lower in children than in adults treated by allogeneic HSCT, the risk factors and the spectrum of clinical manifestations are similar.

Acute GvHD represents a major risk factor for cGvHD, and yet cGvHD can be observed even without preceding aGvHD, and when present, it does not represent merely the continuation of aGvHD. Older age of the recipient at HSCT, transplantation from a multiparous female donor into a male recipient (with reactivity to Y chromosome-associated antigens), and incompatibility at minor histocompatibility loci represent additional risk factors for cGvHD.[10] Furthermore, use of peripheral blood stem cells carries an increased risk of cGvHD compared with use of bone marrow stem cells.

Prevention of GvHD

Prevention is the most effective approach to GvHD, and use of a fully matched donor remains the best method of prevention. Alternatively, if a related HLA-mismatched donor is used for transplantation, it is essential that the graft is vigorously T-cell depleted.

Whenever a conditioning regimen is used in the transplantation protocol, pharmacological prophylaxis of GvHD must also be included, even in the case of HSCT from a related HLA-identical donor.

Long-standing approaches to prevention of GvHD include cyclosporine daily for 6 months or methotrexate (15 mg/m^2 on the first day after HSCT, and then 10 mg/m^2 at days +3, +6, and +11 after transplantation), or a combination of the two. Newer regimens substitute another calcineurin inhibitor, tacrolimus, for cyclosporine. Serotherapy with ATG or alemtuzumab is often included in the preparatory regimen, with the aim of reducing the risk of GvHD. The timing of administration of alemtuzumab as part of the preparatory regimen has opposed effects on the prevention of aGvHD and the speed of T-cell reconstitution. In particular, use of alemtuzumab in the days that immediately precede transplantation has more potent effects on aGvHD prevention but is associated with delayed T-cell reconstitution, whereas the opposite is observed when the drug is administered a few weeks before HSCT. Finally, mycophenolate mofetil and methotrexate can also be used to prevent GvHD.

Attempts to prevent cGvHD have been less satisfactory. In particular, use of a prolonged immune suppression after transplant does not decrease the incidence of cGvHD.

Treatment of GvHD

Once GvHD has developed, treatment is mainly based upon the use of immunosuppressive drugs. Corticosteroids remain the first-line therapy and are usually effective, especially for mild and moderate forms of aGvHD. Second-line therapy includes ATG, mycophenolate mofetil, cyclosporine, or tacrolimus. mAbs, such as anti-CD25 (daclizumab), are also used. However, the efficacy of second-line agents in severe or steroid unresponsive aGvHD is limited.

Treatment of cGvHD is also based on immunosuppression, but with limited efficacy.[11,12] Topical steroids and calcineurin inhibitors may alleviate mucosal and skin symptoms. Systemic steroids have been shown to improve survival, but at the risk of significant adverse effects. Ursodeoxycholic acid may be useful in cGvHD with significant liver involvement. Extracorporeal photopheresis can be used with the goal to induce tolerance; typically, its benefits, if present, are delayed until 2–3 months after initiation of treatment. Use of hydroxychloroquine, mycophenolate mofetil, anti-TNF-α mAb, etanercept (a recombinant form of soluble tumor necrosis factor [TNF]-receptor), and anti-CD20 antibody (rituximab) remains investigational.

Infections

Infections represent one of the major complications following HSCT. Patients with severe PID are intrinsically highly susceptible to infections. In infants with SCID and with other forms of combined immune deficiency, viral and opportunistic infections can develop before transplantation and are one of the factors that adversely affect the outcome of HSCT itself.[2] Regardless of the type of underlying PID, T cell–depleted HSCT carries a high risk of infections because of the longer time required to achieve immune reconstitution. Furthermore, use of a pretransplant conditioning regimen with myeloablative and immunosuppressive drugs, and GvHD prophylaxis, contribute to the increased susceptibility to infections after HSCT.

Strict isolation of the patients during and after HSCT, and prophylactic administration of antibiotics, have been associated with a better survival rate, particularly after related HLA-mismatched transplantation for SCID.[3] In spite of this, infections remain the major cause of death. Challenging viruses in infants with SCID include adenovirus, cytomegalovirus (CMV), parainfluenza type III virus, and Epstein-Barr virus (EBV). In particular, CMV infection after HSCT can cause interstitial pneumonia, enteritis, hepatitis, and encephalitis. EBV can also cause lymphoproliferative disease. Filtering of blood derivatives removes leukocytes and thus reduces the risk of transfusion-associated infections. Several antiviral drugs (acyclovir, ganciclovir, foscarnet, cidofovir) are now available and have good results, especially against CMV,[13] and preemptive administration of anti-CD20 mAb is efficacious in the prevention of EBV-driven lymphoproliferative disease. Furthermore, transfusion of virus-specific cytotoxic T lymphocytes (CTLs) represents another important resource to fight severe viral infections.[14]

Pneumocystis jiroveci is a common cause of pneumonia in severely immunocompromised patients. Treatment is based on intravenous cotrimoxazole (20 mg/kg/day).

Aspergillus infection is a severe complication in patients with chronic granulomatous disease (CGD) and in patients with profound neutropenia. Voriconazole offers some advantage compared with liposomal amphotericin B for treatment of invasive aspergillosis, whereas prophylactic itraconazole reduces the incidence of fungal infections in patients with CGD before transplantation.

Bacterial infections are usually amenable to successful treatment, if the pathogen is identified, and appropriate and aggressive use of antibiotics is initiated. Prophylactic administration of immunoglobulins (Igs) following HSCT also reduces the frequency and severity of infections. Administration of Bacille Calmette-Guérin (BCG) vaccine at birth to prevent tuberculosis is still practiced in many countries worldwide. In severely immunocompromised infants who undergo HSCT, this live vaccine may cause disease, which often manifests at the time of engraftment with local and systemic immune reconstitution inflammatory syndrome (IRIS).

Toxicity Related to Conditioning Regimen

Chemotherapeutic agents that are used in the conditioning regimen of HSCT often cause significant short-term and long-term toxicity. Chemotherapeutic drugs that damage the liver vascular endothelium, particularly busulfan and cyclophosphamide, can cause veno-occlusive disease (VOD), which is clinically marked by painful hepatomegaly, jaundice, ascites, fluid retention, and weight gain and can ultimately result in fatal multiorgan failure (MOF). Defibrotide is the most effective agent studied to date in the treatment of VOD. Busulfan can also cause seizures and lung damage. Cyclophosphamide can cause hemorrhagic cystitis, a syndrome of inappropriate antidiuretic hormone secretion, or more rarely, cardiac disturbances.

Long-term hormonal complications are more common when total body irradiation (TBI) is used. However, the busulfan and cyclophosphamide regimen can cause delayed puberty or sterility, and thyroid dysfunction is frequently observed, even in patients who have not received TBI. Delayed or incomplete tooth eruption is also a possible consequence of conditioning regimens. Effects on final height and growth, as well as long-term neurocognitive effects, are emerging as more children are treated and followed up. Patients with defects of DNA repair (*e.g.*, some forms of SCID) are particularly at risk for drug-related toxicities.

CLINICAL PEARLS

Considerations of Stem Cell Transplantation Unique to Severe Combined Immunodeficiency (SCID)

Patients with SCID are strongly impaired in their ability to reject allogeneic cells, including stem cells. Therefore no chemotherapy is required in these patients in order to achieve T-cell reconstitution following stem cell transplantation.

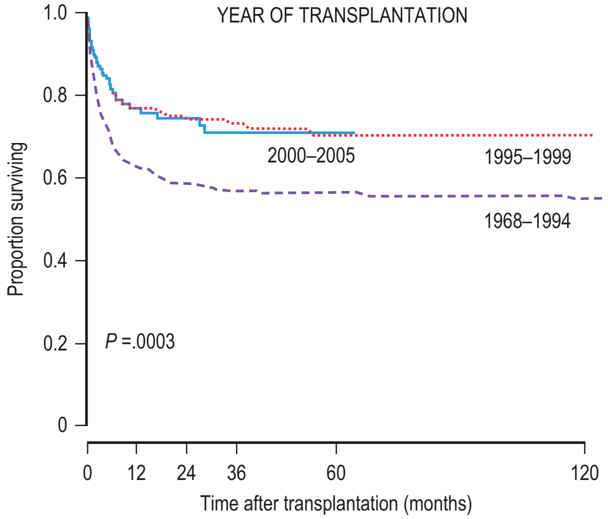
The quality and the kinetics of T-cell reconstitution following stem cell transplantation depend on the type of transplant.

If unmanipulated bone marrow from genotypically human leukocyte antigen (HLA)-identical related donor is used, mature T cells contained in the graft expand as early as 2 weeks after transplantation and provide a rapid source of immune competence.

A similar phenomenon is also observed following unmanipulated matched unrelated donor (MUD) transplantation. In this case, however, drugs used in conditioning regimen and for prevention of graft-versus-host disease (GvHD) decrease the degree of early expansion of donor-derived T cells.

In contrast, appearance of naïve T cells occurs only at 3 months or more after transplantation, regardless of degree of HLA matching between donor and recipient. Consequently, following haploidentical transplantation, there is a prolonged period during which the recipient remains lymphopenic and at high risk of infections.

In the absence of pretransplant conditioning, following haploidentical transplantation for SCID, engraftment is usually restricted to T lymphocytes. This may cause persistent B-cell dysfunction. Furthermore, patients with natural killer (NK)$^+$ SCID may show some ability to reject stem cells from HLA-mismatched donors. For these reasons, several centers include use of conditioning regimen for the hematopoietic stem cell transplantation (HSCT) in babies with SCID who do not have a matched related donor available.



Months	0	6	12	24	36	60	120
Number at risk							
1968–1994	361	245	187	172	166	151	88
1995–1999	157	113	95	78	66	40	3
2000–2005	181	111	74	49	26	4	0

FIG 82.2 Cumulative probability of survival in patients with severe combined immunodeficiency (SCID) after hematopoietic stem cell transplantation (HSCT) according to the period in which transplanted. (With permission from Gennery AR, Slatter MA, Grandin L, et al. Transplantation of hematopoietic stem cells and long-term survival for primary immunodeficiencies in Europe: entering a new century, do we do better? J Allergy Clin Immunol 2010;126:602-10.e1–11.)

HSCT for SCID

General Considerations

SCID is a medical emergency and is uniformly fatal unless promptly diagnosed and successfully treated. With a few exceptions in which alternative strategies (gene therapy, enzyme replacement therapy) can be used, allogeneic HSCT represents the most effective form of treatment for these disorders.

SCID is also a unique situation in which the virtual lack of T lymphocytes strongly impairs the ability of the recipient to reject the graft. Furthermore, donor-derived lymphoid progenitor cells have a striking advantage for *in vivo* T-cell differentiation. Consequently, use of pretransplantation chemotherapy and immune suppression is not required to attain T-cell reconstitution. However, although there is a general consensus that no conditioning is needed for HSCT from an HLA-identical related donor, there is controversy as to whether chemotherapy should be used for T cell–depleted mismatched HSCT. Although use of conditioning chemotherapy carries the risk of drug-related toxicity, it favors the engraftment of donor-derived stem cells and may lead to a better recovery of humoral immunity and of improved thymopoiesis. In approximately 30% of the cases, unconditioned, T cell–depleted MMRD transplantation may fail to induce T-cell reconstitution in infants with SCID. In this case, booster transplants without chemotherapy may be helpful. A recent report analyzed the outcome of such a procedure.[15] Among 49 patients with SCID who received such booster transplants at Duke University Medical Center, 31 (63%) survived for up to 28 years. Older age at HSCT and persistence of preexisting infections represented significant risk factors for death. In several patients, the use of two parents as sequential donors was associated with immune reconstitution.

In contrast, chemotherapy is typically used for HSCT from unrelated donors and—irrespective of donor type—for “leaky” forms of SCID with presence a significant number of autologous, partially functioning T lymphocytes.

Survival Following HSCT for SCID

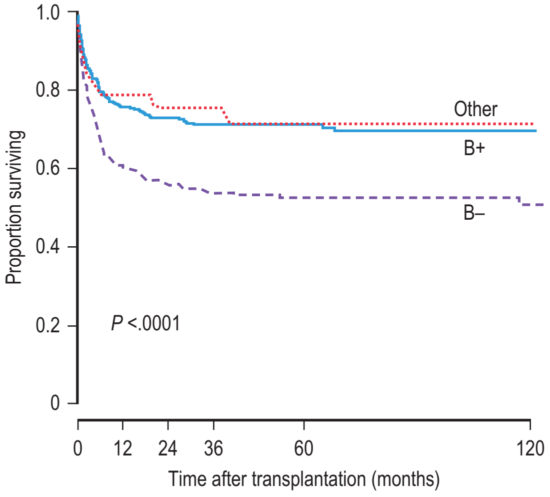
The two largest series of patients with SCID treated by HSCT in Europe and in North America included 699 and 240 cases, respectively.[1,2]

In the European study, which included transplantations performed between 1968 and 2005, 203 patients received HSCT from genotypically (n = 135) or phenotypically (n = 68) identical related donors, with 10-year overall survival rates of 84% and 64%, respectively. This figure is clearly superior to the 54% survival rate observed among 415 patients who had received related HLA-mismatched HSCT. Finally, in this study, 81 patients received HSCT from MUDs, and 66% survived.[1] In the North American study, which included transplantations performed between 2000 and 2009, the overall 5-year survival rate was 74%.[2] In particular, the survival rate was highest among recipients of HSCT from HLA-matched sibling donors (MSDs) (97%), followed by unconditioned HSCT from MMRDs (79%), HSCT from URDs (74%), HSCT from MMRDs with conditioning (66%), and HSCT from UCB (58%).

Survival after HSCT for SCID has improved over the years (Fig. 82.2),[1] reflecting more effective treatment and prevention of disease-related and transplantation-associated complications, such as infections and GvHD. In particular, according to the European study,[1] the current survival rate after HSCT from HLA-identical related donors is 90%, whereas 3-year survival after HSCT from related HLA-mismatched donors has improved from 49% in transplantations performed up to 1994 to 69% in those performed between 1995 and 1999. However, no further improvement in survival has been observed since that time.

Several factors influence survival after HSCT for SCID. In particular, younger age at transplantation leads to superior survival. In the North American study, among 68 infants who were treated before 3.5 months of life, 64 (94%) survived.[2] Similarly high survival rates were observed also among infants older than 3.5 months with no history of infections or whose infections had resolved before HSCT (90% and 84% 5-year survival rates, respectively).[2] Collectively, these data emphasize the importance of early diagnosis to prevent infections and provide strong support to newborn screening for SCID. For patients of any age who do not have MSDs and have an active infection at the time of HSCT, a higher survival rate (65%) is observed with unconditioned transplantation from MMRDs.[2]

The SCID genotype and phenotype may also affect the outcome of HSCT. In the European series, infants with B^+ SCID had superior 10-year survival after HSCT compared with infants with B^- SCID (70% vs 51%, respectively; Fig. 82.3),[1] although no significant difference in survival between these two groups of infants with SCID was observed in the North American study.[2] Furthermore, haploidentical transplantation without conditioning or with immunosuppression only in patients with recombinase activating gene (RAG) or Artemis deficiency is associated with a high risk of graft failure or of incomplete T-cell reconstitution.[16]



Months	0	6	12	24	36	60	120
Number at risk							
B+	345	247	195	168	147	111	55
B–	300	184	135	109	91	69	33
Other	54	38	26	22	20	15	3

FIG 82.3 Cumulative probability of survival in patients with $T^- B^-$ or $T^- B^+$ severe combined immunodeficiency (SCID) after hematopoietic stem cell transplantation (HSCT). (With permission from Gennery AR, Slatter MA, Grandin L, et al. Transplantation of hematopoietic stem cells and long-term survival for primary immunodeficiencies in Europe: entering a new century, do we do better? J Allergy Clin Immunol 2010;126:602-10.e1–11.)

Factors that may contribute to such poor outcome may include the presence of autologous NK cells that may mediate graft rejection. Consistent with this, in a UK series of 77 infants with SCID who received HSCT, survival rate was higher among patients with NK^- SCID than among recipients with NK^+ SCID (see Fig. 82.1).[17] Furthermore, RAG and Artemis deficiencies cause a later block in T-cell development compared with the majority of genetic defects associated with B^+ SCID. This may lead to a stronger competition between donor-derived and autologous T-cell progenitors in patients with B^- SCID who receive unconditioned HSCT.

Finally, in a series of 106 patients with SCID caused by adenosine deaminase deficiency (ADA^- SCID), unconditioned transplantation from MMRDs was associated with a high risk of graft failure.[18]

CLINICAL PEARLS

Outcome of Hematopoietic Stem Cell Transplantation (HSCT) for Primary Immunodeficiencies

For severe combined immunodeficiency (SCID), the 5-year survival rate of patients transplanted between 2000 and 2009 is over 74% and is higher (97%) for patients who receive transplant from matched sibling donors.

Improvements in clinical care (both in critical care and in prevention and treatment of infections) have also contributed to the improved outcome of HSCT for SCID, including transplantation from matched unrelated donors (MUDs). However, the decision to attempt a MUD transplantation for SCID must be weighed against the risks associated with the time interval required to identify such a donor.

Factors influencing survival after transplantation for SCID include younger age at transplant and lack of active infections at the time of transplantation.

The decline of T-cell function that is observed at 10 years or more after transplantation remains a concern and may cause clinical problems. Therefore there is a need for improvements in the procedures used to facilitate and sustain stem cell engraftment and/or to boost donor-derived immunity.

For primary immunodeficiencies other than SCID, there has been a progressive improvement in the outcome following stem cell transplantation. In particular, results are good both for human leukocyte antigen–identical transplants (with a survival rate of 70% or more, depending on the disease) and for MUD or cord blood transplantations. Reduced-intensity conditioning regimens have been often used in these patients, in the attempt to reduce the risks of drug-related toxicity.

Complications Following HSCT for SCID

Despite advances in prophylaxis and treatment, infections (especially those caused by viruses) remain a significant cause of death after HSCT for SCID. In a report of 166 transplantations performed at their center, Buckley et al.[18a] indicated that viral infections accounted for 30 of the 40 deaths observed. CMV and adenovirus were responsible for nine deaths each, and 18 additional deaths resulted from infections caused by EBV, enteroviruses, parainfluenza virus type 3, varicella virus, herpes simplex virus, and respiratory syncytial virus. In a series of 240 patients with SCID who received HSCT in North America between 2000 and 2009, infections accounted for 24 of 62 deaths.[2]

Viral and opportunistic infections are more common early after HSCT, especially in recipients of T cell–depleted haploidentical HSCT, because of the delay in achieving immune reconstitution. Incomplete recovery of immune function at 1 year after

HSCT is associated with a higher risk of late infections. In a single-center study of 90 patients with SCID treated with HSCT, 11 (12%) developed significant infectious complications 2–17 years after transplantation.[19] Among late infections, chronic skin warts caused by papilloma virus have been observed in a significant fraction of infants with γc or JAK3 deficiency after HSCT.[19] This complication may occur also in patients who attain robust immune function and probably results from signaling defects that involve extrahematopoietic cells, such as keratinocytes.

GvHD is another major complication of HSCT for SCID. In a series of 240 patients with SCID who received HSCT in North America between 2000 and 2009, the cumulative incidence of aGvHD of grade 2–4 at 100 days after transplant was 20%, with no significant difference based on donor type.[2] By contrast, in their article describing the experiences of HSCT for SCID at two centers (Brescia, Italy; and Toronto, Canada), Grunebaum et al. reported that aGvHD developed in four (31%) of 13 patients who received related HLA-identical HSCT, 18 (45%) of 40 patients treated with T cell–depleted haploidentical transplantation, and 30 (73%) of 41 patients receiving MUD HSCT.[6] Buckley et al. reported that GvHD occurred in 45 (30.2%) of 149 patients given T cell–depleted mismatched parental bone marrow, eight (47%) of 17 patients given unfractionated HLA-identical marrow, and four (80%) of five patients given placental blood.[18a] In most cases, GvHD occurred when there was presence of transplacentally acquired T lymphocytes. In most cases in this study, the GvHD observed was mild (grade I or II) and required no treatment; however, 11 patients developed GvHD grade III or IV and required treatment with steroids, cyclosporine, and/or tacrolimus. None of these patients has died since then, but one has developed cGvHD. Although continuous improvement in HLA typing has resulted in a progressively lower incidence of aGvHD in the last decades, these data illustrate the need for careful monitoring of infants treated with other than related HLA-identical HSCT and call for adherence to guidelines on the use of immunosuppression for GvHD prophylaxis after MUD HSCT or after conditioned T cell–depleted haploidentical transplantation.

Chronic GvHD disease has been reported in 10 (11%) of 90 patients who have survived for at least 2 years after receiving HSCT for SCID in Paris. Six of them developed disseminated cGvHD, and three died of cGvHD and related infectious complications.[19] A similar rate of cGvHD (15% at 2 years post-HSCT) has been observed in the North American series of 240 infants with SCID.[2]

Immune dysregulation and autoimmunity represent additional complications of HSCT for SCID. In a joint series of 94 infants with SCID who received transplantations in Brescia (Italy) and Toronto (Canada), we have reported that six of 41 patients who received MUD HSCT, and five of 40 children treated with T cell–depleted haploidentical transplantation developed autoimmune cytopenias.[6] These complications may develop a few months after HSCT (when skewing of the T-cell repertoire may be observed) or may persist, particularly in infants with delayed and incomplete immune reconstitution. In particular, Neven et al. have reported that among 90 long-term survivors after HSCT for SCID, 12 patients suffered from autoimmune and inflammatory complications at more than 2 years after HSCT for SCID, and in six of them, the onset of such complications was within the first 2 years after transplantation.[19] These late manifestations of immune dysregulation are often associated with incomplete immune reconstitution and may lead to a poor outcome.

Finally, prolonged nutritional support has been reported after HSCT for SCID. This complication is more frequent among patients treated by mismatched related or unrelated donor HSCT, especially if cGvHD, immune dysregulation, and poor immune reconstitution are also present. Infants with Artemis deficiency (leading to impaired DNA repair) are at particularly high risk for late complications, including growth retardation, requirement for nutritional support, and dental abnormalities.[16,19]

Quality and Kinetics of T-Cell Immune Reconstitution

The effectiveness of HSCT in SCID is well illustrated by the normalization of the number and function of T lymphocytes that is achieved after transplantation (Figs. 82.4 and 82.5). The efficacy of the procedure has been demonstrated in all forms of SCID, although the T-lymphocyte count after HSCT tends to be lower in patients with adenosine deaminase (ADA) deficiency, possibly reflecting irreversible thymic damage. Moreover, normalization of T-lymphocyte count after HSCT demonstrates the ability of stem cells to seed and differentiate in a vestigial thymus.

The kinetics of T-lymphocyte reconstitution differs substantially, depending on the type of transplant. The unmanipulated graft from a related HLA-identical donor contains mature T lymphocytes. Homeostatic as well as antigen-driven expansion of these mature T cells occurs as early as 2 weeks after transplantation (Fig. 82.6).[1] These T cells have a memory (CD45RO) phenotype, are fully competent, and, in fact, provide the recipient with functional immunity.

Mature T cells are present also in bone marrow grafts collected from MUDs. However, the use of conditioning in MUD HSCT impairs, at least in part, immune development in such transplants, compared with unconditioned HSCT from related HLA-identical donors (see Fig. 82.6).

In contrast, newly generated, naïve ($CD45RA^+$ $CD31^+$) T lymphocytes do not appear in circulation until 3–6 months after HSCT, irrespective of the type of transplant (HLA-identical or mismatched), and their number tends to peak at approximately 1 year after HSCT, when a fully polyclonal T-cell repertoire is usually observed. These naïve T lymphocytes are the product of ongoing active thymopoiesis, as shown by the fact that they contain T-cell receptor excision circles (TRECs). TRECs are extrachromosomal DNA episomes, which are generated during V(D)J recombination (Chapter 4) and are not duplicated during mitosis. Therefore TRECs identify newly generated naïve T lymphocytes, and their enumeration in peripheral blood is used as a method to identify babies with SCID at birth.[20]

The kinetics of T-cell reconstitution is influenced by the recipient's age. Transplantations performed early in life (at <3.5 months of age) lead to superior thymic output.[2] This may reflect lack of thymic damage (which is often observed in older infants after infections); alternatively, it is possible that a younger thymus has intrinsic superior ability to support active thymopoiesis.

Quantitation of TRECs sequentially after HSCT is an accepted approach to assess engraftment of *bona fide* stem cells and to monitor persistence of immunity. Although an earlier study had shown that levels of TRECs tended to decline at 10 years after HSCT in recipients of unconditioned mismatched-related transplants, more recent observations from the same group indicate that robust thymopoiesis and generation of a diversified repertoire of T lymphocytes were maintained over the long term after HSCT.[21]

Vigor of immune reconstitution at >2 years after HSCT is influenced by the donor type, use of conditioning, and SCID

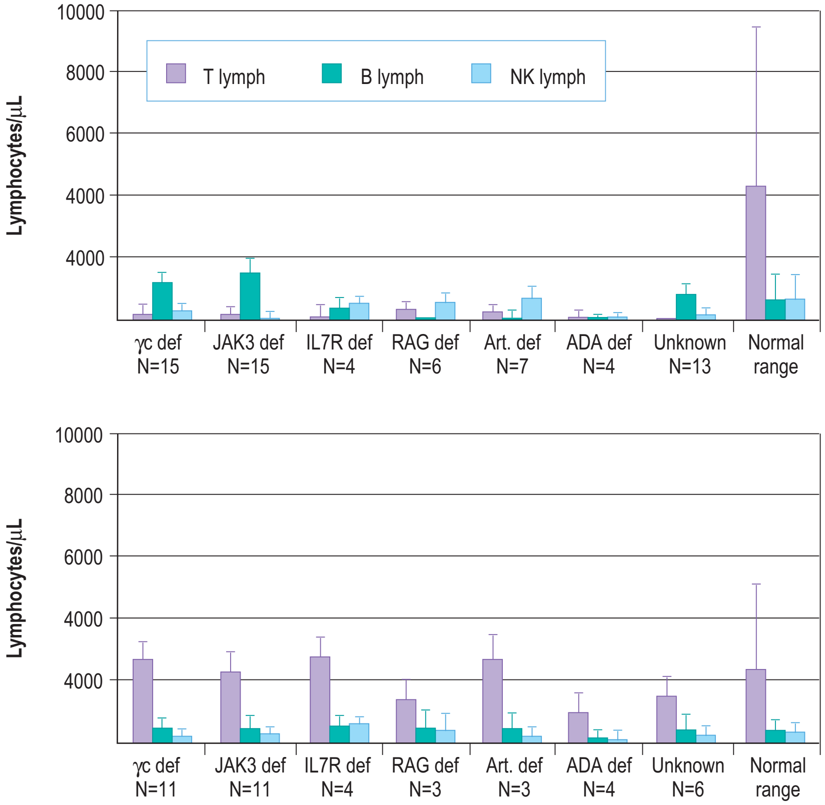


FIG 82.4 Mean (±SE) numbers of $CD3^+$ T cells, $CD19^+$ B cells, and $CD16^+$ natural killer cells before transplantation (A) and at the most recent evaluation after transplantation (B) performed at the University of Brescia, Italy, according to the type of severe combined immunodeficiency (SCID).

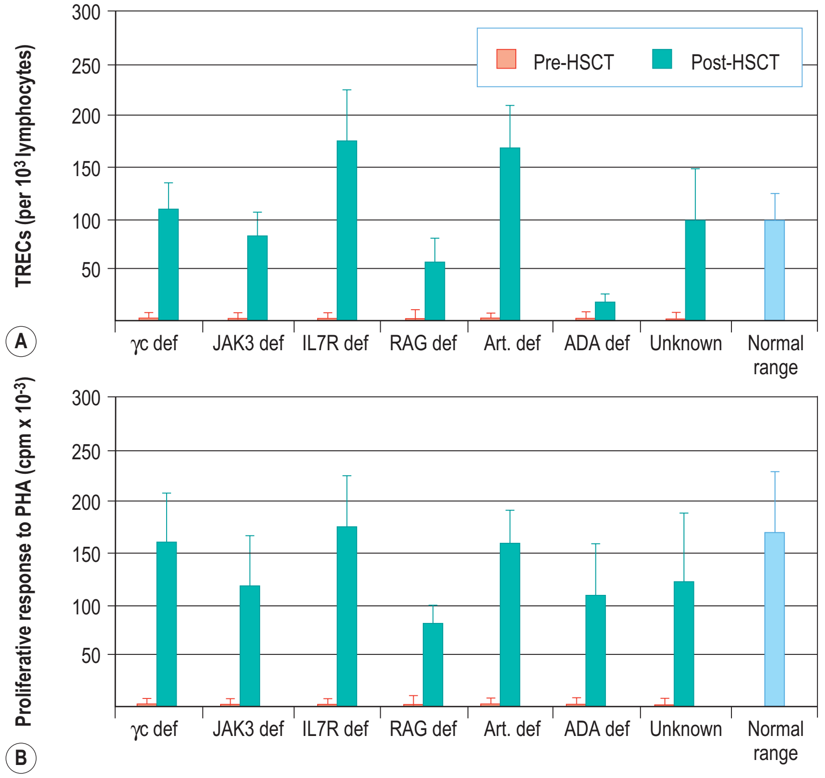


FIG 82.5 Number of T-cell receptor excision circles (TRECs) (A) and *in vitro* proliferative response to phytohemagglutinin (PHA) (B) before transplantation and at the last follow-up after transplantation in a series of 42 infants with severe combined immunodeficiency (SCID) treated at the University of Brescia, Italy, according to the type of SCID.

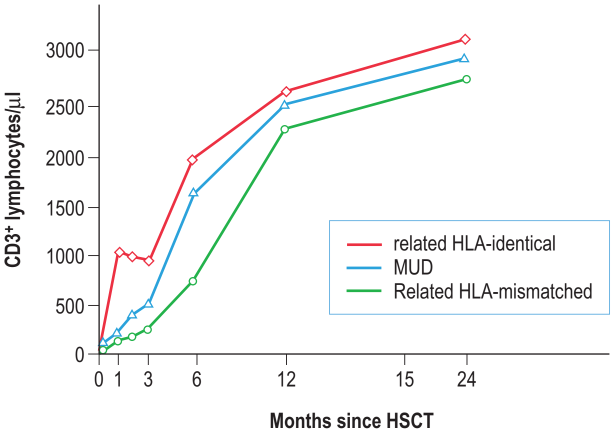


FIG 82.6 Kinetics of CD3$^+$ T lymphocytes reconstitution in 48 infants with severe combined immunodeficiency (SCID) following human leukocyte antigen (HLA)-identical (n = 12), matched unrelated donor (MUD) (n = 15) or haploidentical (n = 21) hematopoietic stem cell transplantation (HSCT), performed at the University of Brescia, Italy. Geometrical mean CD3$^+$ T-cell counts are shown.

phenotype. In particular, in a large series of patients with SCID who received HSCT in North America between 2000 and 2009, higher counts of CD3 cells at 2–5 years after transplantation were associated with HSCT from MSDs, use of conditioning, and B$^+$ or NK$^-$ SCID phenotype.[2] Use of conditioning resulted also in a higher count of naïve CD4 cells.[2]

Reconstitution of B- and NK-Cell Immunity

In contrast to what is observed for T lymphocytes, the engraftment of B cells after HSCT for SCID is often problematic and delayed.

In their series, Buckley et al. reported that only five of 17 survivors of HLA-identical HSCT and 33 of 109 survivors of unconditioned haploidentical HSCT had evidence of donor-derived B lymphocytes; overall, 63 of 126 survivors required intravenous immunoglobulins.[18a] Booster transplants have been used to overcome these problems; 33 of 49 patients who received such booster transplants at Buckley et al.'s institution were reported to be alive with improved immune function.[18a]

In Europe, use of pretransplantation conditioning has been advocated in haploidentical HSCT with the goal of facilitating engraftment of stem cells and thus also of B lymphocytes. Mazzolari et al. have reported that donor B-cell engraftment was achieved in only one of 11 survivors after unconditioned HSCT compared with 26 of 29 patients who received a myeloablative conditioning regimen.[22]

Recently, higher B-cell counts and a better chance of attaining independence from Ig replacement therapy in recipients of conditioned HSCT for SCID have been confirmed also in a large North American study.[2] Finally, attainment of normal B-cell function may also depend on the nature of the genetic defect, as shown by the fact that among infants with B$^+$ SCID, those who have an *IL-7RA* gene defect usually develop normal B-cell immunity after HSCT, even if no donor-derived B cells are present, whereas patients with γc or JAK3 deficiency (both of which compromise B-cell function) often remain dependent on Ig substitution therapy if engraftment of donor-derived B cells is not achieved. This reflects the importance of interleukin-21 (IL-21)–mediated, γc/JAK3-dependent signaling for germinal-center reaction, plasma-cell differentiation, and antibody production.

More limited data are available about reconstitution of NK cell function. In patients with NK$^-$ SCID, NK cells are often the first cells to appear after haploidentical HSCT. Lower NK-cell counts are observed at long-term follow-up after HSCT in patients with γc or JAK3 defects.

HSCT for Combined Immunodeficiencies Other Than SCID

The Primary Immune Deficiency Treatment Consortium (PIDTC) has established criteria to distinguish typical SCID from other forms of combined immunodeficiency characterized by residual (although lower than normal) T cell–mediated immunity.[23] Omenn syndrome is a fatal disorder, unless treated with HSCT. Satisfactory results have been obtained with transplantation from related HLA-identical donors, but less so from haploidentical donors. However, an improved outcome has been observed in the last 15 years. Mazzolari et al. have reported that 9 of 11 patients with Omenn syndrome were alive after HSCT; importantly, of these, only one had a matched sibling, whereas two infants had a phenotypically identical related donor, three were treated with MUD HSCT, and five received a haploidentical HSCT.[22]

HSCT has been attempted also in other predominant T-cell immunodeficiencies, such as purine nucleoside phosphorylase deficiency, cartilage hair hypoplasia, and other forms of T-cell activation deficiency. Overall, survival after HSCT for these disorders is worse than in typical SCID (approximately 50%).[1]

Major histocompatibility complex (MHC) class II deficiency remains a very difficult disease to treat with transplantation. In the European series, cumulative survival after HSCT performed in 1995–2005 was around 40%.[1] Many patients with this deficiency fail to reconstitute the number of circulating CD4 T lymphocytes, probably because the lack of expression of HLA class II molecules on thymic epithelial cells prevents positive selection of CD4 lymphocytes.

The mainstay of treatment for patients with complete DiGeorge syndrome is represented by thymic transplantation.[24] However, HSCT or even transplantation of unmobilized peripheral blood mononuclear cells may be attempted if an HLA-identical donor is available; in such cases, immune reconstitution is provided by mature T lymphocytes contained in the graft.

Patients with CD40 ligand (CD40L) deficiency suffer from recurrent bacterial and opportunistic (*Pneumocystis jiroveci, Cryptosporidium parvum*) infections, resulting in increased mortality in childhood and young adulthood. This has prompted use of HSCT in the treatment of this disease. In a series of 38 patients transplanted in Europe, 26 (68.4%) survived.[25] Early age at transplantation and lack of preexisting pulmonary disease were associated with a more favorable outcome. Successful outcome has been also reported after HSCT for CD40 deficiency.

X-linked immunodeficiency with ectodermal dystrophy, caused by mutations of the *IKBKG* (*NEMO*) gene that impair nuclear factor (NF)-κB signaling, may also present with features of combined immunodeficiency. Colitis is also frequently seen and may be caused by immune dysregulation as well as abnormalities of NF-κB signaling in the gut epithelial cells. Initial observations supported the notion that use of myeloablative regimens for HSCT is poorly tolerated in this condition, also because of increased mucosal toxicity. However, graft failure has been observed when using reduced-intensity conditioning, and in contrast, several patients have tolerated myeloablative regimens well and have attained immune reconstitution.[26] Yet, persistence of the colitis after HSCT has been reported in a proportion of the patients.[27] Conflicting results have also been reported after

HSCT in patients with hypermorphic mutations of the *IKBA* gene, another cause of combined immune deficiency with ectodermal dystrophy.[28]

Wiskott-Aldrich Syndrome

Bone marrow transplantation for correction of WAS was attempted as early as 1968, with partial success. Full correction following HSCT was first reported in 1978 when a more robust conditioning regimen was used. Since then, results of related HLA-identical HSCT in WAS have been consistently good, with continuous improvement in recent years, and excellent results have been also achieved with HSCT from MUDs. A multiinstitutional study of 194 patients with WAS treated with HSCT showed an overall survival of 84%, and a 5-year survival as high as 89.1% for transplantations performed since year 2000.[29] For patients treated with MUD HSCT, younger age (<5 years) at transplantation was associated with an improved outcome. Patients with preexisting autoimmunity or recurrent/severe infections had a higher rate of complications after transplantation. Mixed chimerism was associated with the presence of immunological abnormalities and increased risk of autoimmunity, and myeloid chimerism <50% was associated with persistent thrombocytopenia.[29] Use of a conditioning regimen other than fully myeloablative conditioning was associated with reduced disease-free survival in another study of 14 patients.[30]

Cytotoxicity Defects

Familial hemophagocytic lymphohistiocytosis (FHL) comprises a genetically heterogeneous group of disorders of T cell– and NK cell–mediated cytotoxicity. Although chemotherapy may induce remission, patients with FHL tend to relapse and ultimately die, mostly as a result of multiorgan failure observed in the accelerated phase of the disease. At present, HSCT is the only curative approach to FHL. However, patients with FHL often are critically ill, with extensive organ involvement and/or active infections, and may suffer from refractory disease. For these reasons, patients are unusually prone to developing transplantation-related toxicities and complications. Recently, significantly improved outcomes with increased survival have been reported after HSCT with the use of a reduced-intensity conditioning regimen.[31] However, use of reduced-intensity conditioning is associated with a higher risk of mixed and recipient chimerism (65%), even for HSCT from matched donors.[31] Selection of optimal family donors is an important issue for HSCT in FHL. Functional and genetic studies should be used to screen potential family donors to avoid obtaining transplants from genetically affected but as-yet asymptomatic subjects.

Excellent results have been reported also after HSCT for X-linked lymphoproliferative disease type 1 (XLP1). In an international series of 91 patients with XLP caused by a *SH2D1A* gene defect, survival was 81.4% in 43 patients treated with HSCT compared with 62.5% in 48 untransplanted patients.[32] Furthermore, the majority of untransplanted survivors required Ig replacement therapy, whereas good immune reconstitution was achieved in most transplanted patients. However, survival was poorer (50%) among patients with a previous history of hemophagocytic lymphohistiocytosis (HLH), suggesting that ideally the transplantation should be performed before onset of EBV infection.[32] More recently, a single-center experience involving 16 patients with XLP1 showed that reduced-intensity conditioning with alemtuzumab, fludarabine, and melphalan is effective even in patients with XLP1 with a history of HLH (Fig. 82.7).[33] Survival was 80% at 1 year and 71% in the long term; mixed chimerism was common but, in most cases, was sufficient to control the disease. Infectious complications (especially viral infections) were recorded in the majority of the patients.[33]

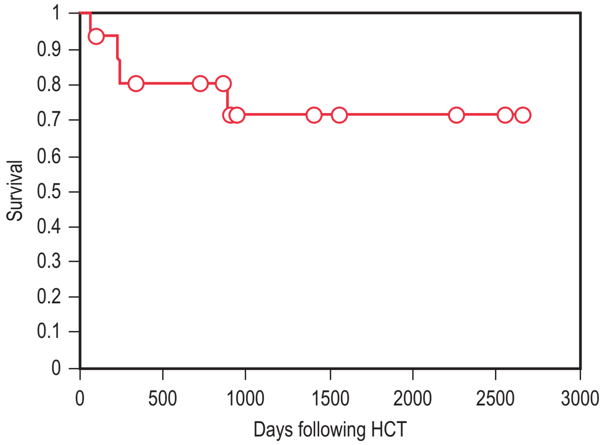


FIG 82.7 Long-term survival of patients after reduced intensity conditioning hematopoietic stem cell transplantation (HSCT) for X-linked lymphoproliferative disease type 1 (XLP1). (With permission from Moratto D, Giliani S, Bonfim C, et al. Long-term outcome and lineage-specific chimerism in 194 Wiskott–Aldrich Syndrome patients treated by hematopoietic cell transplantation between 1980–2009: an international collaborative study. Blood 2011;118:1675–84.)

XLP type 2 (XLP2), caused by mutations in the X-linked inhibitor of apoptosis (*XIAP*) gene, is associated with various phenotypes, including XLP, HLH, and severe colitis. An international survey of transplantation for this condition identified 19 patients who underwent HSCT, in which fully myeloablative (n = 7), reduced-intensity (n = 11), or intermediate-level (n = 1) conditioning was used.[34] Myeloablative conditioning was associated with poor survival (14%) and a high rate of serious, transplantation-related toxicities. Survival was better (55%) in patients who received reduced-intensity conditioning and was especially good (86%) in those in remission from HLH.[34]

The hematological and immune complications of Chediak-Higashi syndrome (CHS) can be cured with bone marrow transplantation. In a series of 35 patients, 5-year survival after HSCT was 62%.[35] Mortality was higher in patients who were in the life-threatening accelerated phase of the disease at the time of transplantation and in those who received HSCT from an alternative related donor.[35] Use of MUDs represents a valid option for HSCT in patients with CHS. However, the long-term outcome of patients with CHS treated with HSCT remains unclear, especially since neurological deterioration has been consistently observed several years after transplantation.

Griscelli syndrome type 2 (GS2) is a genetic disease characterized by hemophagocytic lymphohistiocytosis and a high risk of neurological complications. Correction of the cytotoxicity defect after HSCT has been reported in a small series of patients with GS2, but neurological sequelae remain a challenge and may develop even in patients without preexisting neurological problems.[36]

Phagocytic Cell Disorders

Although regular administration of prophylactic antibiotics and antifungal agents (with the possible addition of interferon [IFN]-γ) has clearly improved the outcome in patients with CGD, this remains a severe disorder with a rather high risk of complications and death, especially in patients with oxidase-null phenotype.[37] Therefore there has been a renewed interest in HSCT for CGD. In a survey of the European experience in 1985–2000, 27 patients with CGD were treated with HSCT; most of them received unmodified bone marrow from an HLA-identical sibling after myeloablative conditioning. Overall, 23 of the 27 patients were reported to have survived, and the disease had been cured in 22 of them.[38] In this study, all 18 patients who were infection-free at the time of transplantation survived. However, a real breakthrough in the treatment of CGD was provided by a prospective multicenter study in 56 patients with CGD, 42 of whom had high-risk features, including treatment-refractory infections and severe inflammatory complications.[39] In this study, reduced-intensity conditioning (with low-dose or targeted busulfan administration, high-dose fludarabine, and serotherapy) and transplantation from matched related (n = 21) or unrelated (n = 35) donors resulted in a 2-year overall survival rate of 96% and an event-free survival rate of 91% (Fig. 82.8). Stable (≥90%) myeloid chimerism was documented in 93% of surviving patients. A low rate of aGvHD of grade III/IV (4%) and of cGvHD (7%) was observed.

HSCT is a successful and lifesaving procedure also in patients with the complete form of leukocyte adhesion defect type 1 (LAD-1). A multicenter study of 36 such patients who underwent HSCT between 1993 and 2007 reported an overall survival of 75%, with similar results when matched related or unrelated donors were used. Mortality was higher (four of eight cases) after haploidentical HSCT. Stable mixed multilineage chimerism is sufficient to cure the disease.[40] More recently, employing blockade of the P40 subunit of IL-23 and IL-12 with ustekinumab has been reported resolve inflammatory symptoms in a patient with LAD-1 after 1 year of therapy.[40a]

Administration of recombinant G-CSF is the treatment of choice for patients with severe congenital neutropenia (SCN). However, a subgroup of these patients fail to respond to G-CSF, and some of them are at high risk for development of myelogenous leukemia. A retrospective multicenter study of 136 patients with SCN who underwent allogeneic HSCT between 1990 and 2012 reported a 3-year overall survival rate of 82%, with 17% transplantation-related mortality.[41] Transplantation at a younger age (<10 years) and use of matched related or unrelated donors were associated with an improved outcome.

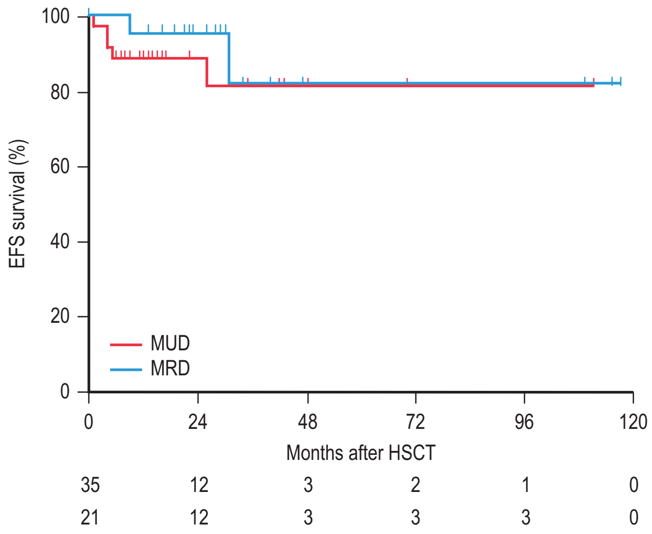


FIG 82.8 Event-free survival (EFS) in patients with chronic granulomatous disease (CGD) after reduced-intensity conditioning hematopoietic stem cell transplantation (HSCT) from matched related donors (MRD) or matched unrelated donors (MUD). (With permission from Eapen M, DeLaat CA, Baker KS, et al. Hematopoietic cell transplantation for Chediak-Higashi syndrome. Bone Marrow Transplant 2007;39:411–5.)

Other Primary Immune Deficiencies

IFN-γ receptor 1 deficiency leads to severe mycobacterial infections, with a high rate of mortality early in life. Although HSCT should theoretically correct the disease, results have been disappointing, with few exceptions. An international survey of eight transplanted patients showed that only two were in full remission 5 years after transplantation. The high level of IFN-γ in these patients inhibits engraftment of donor-derived HSCs, accounting for the high rejection rate. Correction of the disease has been reported with HSCT following control of mycobacterial infection, normalization of IFN-γ levels, and use of myeloablative conditioning.

Immunodysregulation/polyendocrinopathy/enteropathy/X-linked (IPEX) syndrome is a severe disorder with immune dysregulation caused by mutations of the *FOXP3* gene, which plays a critical role in the development and function of regulatory T cells (Tregs). Patients with IPEX often die early in infancy. HSCT is the only curative approach to treat this disease.[42] Selective advantage for donor-derived Tregs has been observed in patients with mixed chimerism after HSCT.[43] However, other studies have reported a higher risk of graft failure and increased occurrence of posttransplantation autoimmune cytopenias following HSCT with reduced-intensity conditioning.[44]

Dedicator of cytokinesis 8 (DOCK8) deficiency is a recently described form of combined immunodeficiency with elevated serum IgE, cutaneous viral infections, and a high risk of malignancy. HSCT represents an effective form of treatment for this disease. In a series of 11 patients, seven of whom received HSCT from related donors and four from unrelated donors, survival was excellent (91%), and selective advantage for donor-derived T cells and switched memory cells was noticed.[45] In other studies, successful outcomes were reported with either myeloablative or reduced-intensity conditioning. HSCT has been also successfully used in patients with combined immune deficiency resulting from DOCK2 deficiency.[46]

FUTURE TRANSLATIONAL RESEARCH FOR HCST IN THE TREATMENT OF PID

Although HSCT has clearly shown its efficacy in patients with severe forms of PID, several goals remain to be met. In particular, the main areas of interest for future development include *(i)* methods to improve and sustain engraftment of stem cells; *(ii)* strategies to facilitate engraftment and to reduce the incidence of GvHD in recipients of mismatched or matched unrelated transplants; *(iii)* attempts to improve thymopoiesis, with the goal to accelerate immune reconstitution and to avoid or postpone long-term decline of immunity; and *(iv)* design strategies to reduce the burden of infections after HSCT.

Generation of a diversified pool of naïve T lymphocytes after HSCT depends on thymic function. Use of cytotoxic drugs and GvHD are significant risk factors for posttransplantation T-cell deficiency as they interfere with normal thymic function.[47] Use of mAbs against CD117 (c-kit) has been proposed to displace autologous stem cells and favor engraftment of donor-derived cells, while avoiding the drug-related toxicity of chemotherapy and radiotherapy.[48] Recognition that B-cell engraftment after HSCT is hampered by competition between host and donor early B-cell precursors[49] may prompt novel forms of conditioning regimens targeted to host pro-B lymphocytes, in an attempt to facilitate development of donor-derived B cells, even in the unconditioned HLA-identical setting.

Strategies aimed at improving thymic function after HSCT may include *(i)* protection of thymic stroma that supports thymopoiesis and *(ii)* direct stimulation of early T-cell progenitors. Keratinocyte growth factor (KGF) is potentially attractive because it protects the thymic stroma. Administration of KGF before HSCT has been shown to enhance thymopoiesis and peripheral T-cell numbers and to reduce the incidence and severity of GvHD in murine models of HSCT.[50] However, more must be learned about the long-term outcome of this treatment before clinical trials can be started.

Attempts to accelerate immune reconstitution might be based on the use of cytokines, such as IL-7, which promote T-cell development and maturation in the thymus. However, experience in severely immunodeficient mice has shown that infusion of IL-7 provides limited benefit in the MHC-compatible or partially compatible setting, whereas it has facilitated T-cell development in a fully mismatched setting.

Finally, besides use of virus-specific CTLs,[14] infusion of donor T cells that have antiinfective activity but no GvHD activity might help reduce infection-related mortality after HSCT. Such strategy could be based on removal of alloreactive cells by negative selection using anti-CD25 or anti-CD69 antibodies.

ON THE HORIZON

The following strategies may help improve even further the outcome of hematopoietic stem cell transplantation (HSCT) for primary immunodeficiency (PID) in the future:

Methods to improve and sustain engraftment of stem cells

Strategies to facilitate engraftment and reduce the occurrence of graft-versus-host disease (GvHD) in patients with mismatched or matched unrelated donor (MUD) transplant

Approaches to increase thymopoiesis and accelerate immune reconstitution

Design of therapeutic strategies to reduce the infectious complications of HSCT

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Gennery AR, Slatter MA, Grandin L, et al. Transplantation of hematopoietic stem cells and long-term survival for primary immunodeficiencies in Europe: entering a new century, do we do better? J Allergy Clin Immunol 2010;126:602–10.e1–11.
2. Pai SY, Logan BR, Griffith LM, et al. Transplantation outcomes for severe combined immunodeficiency, 2000-2009. N Engl J Med 2014;371:434–46.
3. Klein OR, Chen AR, Gamper C, et al. Alternative-donor hematopoietic stem cell transplantation with post-transplantation cyclophosphamide for nonmalignant disorders. Biol Blood Marrow Transplant 2016;22:895–901.
4. Handgretinger R. Negative depletion of CD3$^+$ and TcRαββ$^+$ T cells. Curr Opin Hematol 2012;19:434–9.
5. Kwan A, Abraham RS, Currier R, et al. Newborn screening for severe combined immunodeficiency in 11 screening programs in the United States. JAMA 2014;312:729–38.
6. Grunebaum E, Mazzolari E, Porta F, et al. Bone marrow transplantation for severe combined immunodeficiency. JAMA 2006;295:508–18.
7. Gennery AR, Cant AJ. Cord blood stem cell transplantation in primary immune deficiencies. Curr Opin Allergy Clin Immunol 2007;7:528–34.
8. Hassan A, Lee P, Maggina P, et al. Host natural killer immunity is a key indicator of permissiveness for donor cell engraftment in patients with severe combined immunodeficiency. J Allergy Clin Immunol 2014;133:1660–6.
9. Slatter MA, Rao K, Amrolia P, et al. Treosulfan-based conditioning regimens for hematopoietic stem cell transplantation in children with primary immunodeficiency: United Kingdom experience. Blood 2011;117:4367–75.
10. Ferrara JL, Levine JE, Reddy P, et al. Graft-versus-host disease. Lancet 2009;373:1550–61.
11. Wolff D, Gerbitz A, Ayuk F, et al. Consensus conference on clinical practice in chronic graft-versus-host disease (GVHD): first-line and topical treatment of chronic GVHD. Biol Blood Marrow Transplant 2010;16:1611–28.
12. Wolff D, Schleuning M, von Harsdorf S, et al. Consensus conference on clinical practice in chronic GVHD: second-line treatment of chronic graft-versus-host disease. Biol Blood Marrow Transplant 2011;17:1–17.
13. Tomblyn M, Chiller T, Einsele H, et al. Guidelines for preventing infectious complications among hematopoietic cell transplantation recipients: a global perspective. Biol Blood Marrow Transplant 2009;15:1143–238. Erratum in: Biol Blood Marrow Transplant 2010;16:294.
14. Naik S, Nicholas SK, Martinez CA, et al. Adoptive immunotherapy for primary immunodeficiency disorders with virus-specific T lymphocytes. J Allergy Clin Immunol 2016;137:1498–505.
15. Teigland CL, Parrott RE, Buckley RH. Long-term outcome of non-ablative BMT in patients with SCID. Bone Marrow Transplant 2013;48:1050–5.
16. Schuetz C, Neven B, Dvorak CC, et al. SCID patients with ARTEMIS vs RAG deficiencies following HCT: increased risk of late toxicity in ARTEMIS-deficient SCID. Blood 2014;123:281–9.
17. Hassan A, Lee P, Maggina P, et al. Host natural killer immunity is a key indicator of permissiveness for donor cell engraftment in patients with severe combined immunodeficiency. J Allergy Clin Immunol 2014;133:1660–6.
18. Hassan A, Booth C, Brightwell A, et al. Outcome of hematopoietic stem cell transplantation for adenosine deaminase-deficient severe combined immunodeficiency. Blood 2012;120:3615–24.
18a. Buckley RH. Transplantation of hematopoietic stem cells in human severe combined immunodeficiency: longterm outcomes. Immunol Res 2011;49(1-3):25–43.
19. Neven B, Leroy S, Decaluwe H, et al. Long-term outcome after hematopoietic stem cell transplantation of a single-center cohort of 90 patients with severe combined immunodeficiency. Blood 2009;113:4114–24.
20. Kwan A, Abraham RS, Currier R, et al. Newborn screening for severe combined immunodeficiency in 11 screening programs in the United States. JAMA 2014;312:729–38.
21. Sarzotti-Kelsoe M, Win CM, Parrott RE, et al. Thymic output, T-cell diversity, and T-cell function in long-term human SCID chimeras. Blood 2009;114:1445–53.
22. Mazzolari E, Forino C, Guerci S, et al. Long-term immune reconstitution and clinical outcome after stem cell transplantation for severe T-cell immunodeficiency. J Allergy Clin Immunol 2007;120:892–9.
23. Shearer WT, Dunn E, Notarangelo LD, et al. Establishing diagnostic criteria for severe combined immunodeficiency disease (SCID), leaky SCID, and Omenn syndrome: the Primary Immune Deficiency Treatment Consortium experience. J Allergy Clin Immunol 2014;133:1092–8.
24. Markert ML, Devlin BH, Chinn IK, et al. Thymus transplantation in complete DiGeorge anomaly. Immunol Res 2009;44:61–70.

25. Gennery AR, Khawaja K, Veys P, et al. Treatment of CD40 ligand deficiency by hematopoietic stem cell transplantation: a survey of the European experience, 1993–2002. Blood 2004;103:1152–7.
26. Permaul P, Narla A, Hornick JL, et al. Allogeneic hematopoietic stem cell transplantation for X-linked ectodermal dysplasia and immunodeficiency: case report and review of outcomes. Immunol Res 2009;44:89–98.
27. Pai SY, Levy O, Jabara HH, et al. Allogeneic transplantation successfully corrects immune defects, but not susceptibility to colitis, in a patient with nuclear factor-κB essential modulator deficiency. J Allergy Clin Immunol 2008;122:1113–18.
28. Fish JD, Duerst RE, Gelfand EW, et al. Challenges in the use of allogeneic hematopoietic SCT for ectodermal dysplasia with immune deficiency. Bone Marrow Transplant 2009;43:217–21.
29. Moratto D, Giliani S, Bonfim C, et al. Long-term outcome and lineage-specific chimerism in 194 Wiskott–Aldrich Syndrome patients treated by hematopoietic cell transplantation between 1980–2009: an international collaborative study. Blood 2011;118:1675–84.
30. Stepensky P, Krauss A, Goldstein G, et al. Impact of conditioning on outcome of hematopoietic stem cell transplantation for Wiskott-Aldrich syndrome. J Pediatr Hematol Oncol 2013;35:e234–8.
31. Marsh RA, Jordan MB, Filipovich AH. Reduced-intensity conditioning haematopoietic cell transplantation for haemophagocytic lymphohistiocytosis: an important step forward. Br J Haematol 2011;154:556–63.
32. Booth C, Gilmour KC, Veys P, et al. X-linked lymphoproliferative disease due to SAP/SH2D1A deficiency: a multicenter study on the manifestations, management and outcome of the disease. Blood 2011;117:53–62.
33. Marsh RA, Bleesing JJ, Chandrakasan S, et al. Reduced-intensity conditioning hematopoietic cell transplantation is an effective treatment for patients with SLAM-associated protein deficiency/X-linked lymphoproliferative disease type 1. Biol Blood Marrow Transplant 2014;20:1641–5.
34. Marsh RA, Rao K, Satwani P, et al. Allogeneic hematopoietic cell transplantation for XIAP deficiency: an international survey reveals poor outcomes. Blood 2013;121:877–83.
35. Eapen M, DeLaat CA, Baker KS, et al. Hematopoietic cell transplantation for Chediak-Higashi syndrome. Bone Marrow Transplant 2007;39: 411–15.
36. Pachlopnik Schmid J, Moshous D, Boddaert N, et al. Hematopoietic stem cell transplantation in Griscelli syndrome type 2: a single-center report on 10 patients. Blood 2009;114:211–18.
37. Kuhns DB, Alvord WG, Heller T, et al. Residual NADPH oxidase and survival in chronic granulomatous disease. N Engl J Med 2010;363:2600–1.
38. Seger RA, Gungor T, Belohradsky BH, et al. Treatment of chronic granulomatous disease with myeloablative conditioning and an unmodified allograft: a survey of the European experience, 1985–2000. Blood 2002;100:4344–50.
39. Güngör T, Teira P, Slatter M, et al. Reduced-intensity conditioning and HLA-matched haemopoietic stem-cell transplantation in patients with chronic granulomatous disease: a prospective multicentre study. Lancet 2014;383:436–48.
40. Qasim W, Cavazzana-Calvo M, Davies EG, et al. Allogeneic hematopoietic stem-cell transplantation for leukocyte adhesion deficiency. Pediatrics 2009;123:836–40. Erratum in: Pediatrics 2009;123:1436.

40a. Moutsopoulos NM, Zerbe CS, Wild T, et al. Interleukin-12 and interleukin-23 blockade in leukocyte adhesion deficiency type 1. N Engl J Med 2017;376:1141–6.

41. Fioredda F, Iacobelli S, van Biezen A, et al. Stem cell transplantation in severe congenital neutropenia: an analysis from the European Society for Blood and Marrow Transplantation. Blood 2015;126:1885–92.
42. Rao A, Kamani N, Filipovich A, et al. Successful bone marrow transplantation for IPEX syndrome after reduced-intensity conditioning. Blood 2007;109:383–5.
43. Horino S, Sasahara Y, Sato M, et al. Selective expansion of donor-derived regulatory T cells after allogeneic bone marrow transplantation in a patient with IPEX syndrome. Pediatr Transplant 2014;18:E25–30.
44. Kucuk ZY, Bleesing JJ, Marsh R, et al. A challenging undertaking: Stem cell transplantation for immune dysregulation, polyendocrinopathy, enteropathy, X-linked (IPEX) syndrome. J Allergy Clin Immunol 2016;137:953–955.e4.
45. Al-Herz W, Chu JI, van der Spek J, et al. Hematopoietic stem cell transplantation outcomes for 11 patients with dedicator of cytokinesis 8 deficiency. J Allergy Clin Immunol 2016;138:852–9.
46. Dobbs K, Domínguez Conde C, Zhang SY, et al. Inherited DOCK2 Deficiency in Patients with Early-Onset Invasive Infections. N Engl J Med 2015;372:2409–22.
47. Krenger W, Blazar BR, Holländer GA. Thymic T-cell development in allogeneic stem cell transplantation. Blood 2011;117:6768–76.
48. Czechowicz A, Kraft D, Weissman IL, et al. Efficient transplantation via antibody-based clearance of hematopoietic stem cell niches. Science 2007;318:1296–9.
49. Liu A, Vosshenrich CA, Lagresle-Peyrou C, et al. Competition within the early B cell compartment conditions B cell reconstitution after hematopoietic stem cell transplantation in non-irradiated recipients. Blood 2006;108:1123–8.
50. Wils EJ, Cornelissen JJ. Thymopoiesis following allogenic stem cell transplantation: new possibilities for improvement. Blood Rev 2005;19:89–98.

MULTIPLE-CHOICE QUESTIONS

1. What distinguishes the kinetics of immune reconstitution following hematopoietic stem cell transplantation (HSCT) with unmanipulated bone marrow from a matched sibling donor versus a T cell–depleted transplantation from a haploidentical donor in an infant with severe combined immunodeficiency (SCID)?
   A. Transplantation from a matched sibling donor will provide rapid engraftment of B lymphocytes.
   B. Transplantation from an human leukocyte antigen (HLA)–matched donor allows rapid T-cell engraftment.
   C. Only transplantation from a matched sibling donor permits durable T-cell reconstitution.
   D. Transplantation from an HLA-matched donor allows rapid multilineage hematopoietic engraftment.
   E. There are no differences in the kinetics of immune reconstitution.

2. Which of the following factors negatively affect outcome of HSCT for SCID?
   A. B^+ SCID phenotype
   B. History of previous infections
   C. Presence of active infections at the time of transplantation
   D. Age younger than 2 months of life
   E. Transplantation from a haploidentical donor

3. Use of reduced intensity conditioning regimens may reduce drug-related toxicity but often results in mixed chimerism after HSCT for primary immune deficiency (PID). In many cases, mixed chimerism is sufficient to control the disease phenotype, but in some forms of PID, it is associated with an increased risk of complications. Can you identify in which of the following immunodeficiency this is the case?
   A. Familial hemophagocytic lymphohistiocytosis (FHL)
   B. Chronic granulomatous disease (CGD)
   C. Immune dysfunction/polyendocrinopathy/enteropathy/X-linked (IPEX) syndrome
   D. Wiskott-Aldrich syndrome (WAS)
   E. Leukocyte adhesion defect type 1 (LAD-1)

83

Hematopoietic Stem Cell Transplantation for Malignant Diseases

Pashna N. Munshi, Scott D. Rowley, Robert Korngold

Hematopoietic stem cell transplantation (HSCT) is effective treatment for most hematological malignancies, including leukemia, lymphoma, multiple myeloma (MM), and clonal myelodysplastic and myeloproliferative diseases (MPDs), as well as nonmalignant diseases, such as autoimmune disorders and hemoglobinopathies. Autologous HSCT (auto-HSCT) is commonly used as therapy for patients with malignancies sensitive to chemotherapy or radiotherapy in a dose-responsive manner. These patients receive intensive cytoreductive regimens designed to eliminate all tumor cells but which, in so doing, also destroy the patient's hematopoietic function needed for blood formation. Infusion of previously collected hematopoietic stem cells (HSCs) will rescue the patient from the marrow-ablative effects of this treatment. Allogeneic HSCT (allo-HSCT), in addition to reconstitution of bone marrow function, achieves an immunotherapeutic benefit from the donor natural killer (NK) and T cells infused into the graft attacking residual tumor cells that persist after the conditioning regimen, thereby greatly reducing the risk of later relapse of the disease. Thus allo-HSCT, in contrast to auto-HSCT, does not require administration of dose-intensive regimens to achieve complete tumor cell kill, and nonmyeloablative regimens may be used to "condition" the host for transplantation.

Auto-HSCT (including syngeneic twins) is justified by the dose-sensitivity of most hematological malignancies. Although there is some evidence that a more robust immunological recovery after auto-HSCT predicts for a lower risk of relapse, possibly opening an area of research in graft (or host) modification to enhance such recovery, treatment of the disease is primarily a result of the dose-intensive, myeloablative chemotherapy or radiotherapy administered. Infusion of cells is only required to recover hematopoiesis, and the stem cell infusion is, therefore, intended to treat the deleterious effect of chemotherapy on bone marrow function and not the disease itself. The primary complications of auto-HSCT result from the administration of a dose-intensive regimen and include a period of marrow hypoplasia, possibly requiring blood transfusions and antibiotics. Nonhematological toxicities, including mucositis resulting in inanition and diarrhea, and damage to other organs, such as the lung, liver, and kidney, limit the amount of chemotherapy that can be administered. Currently, the treatment-related mortality risk for most treatment regimens is ≤5%. Relapse of disease, particularly for patients who come to transplantation with chemotherapy-refractory disease, is the primary cause of failure of auto-HSCT. Improvements in the outcome of auto-HSCT will require new conditioning regimens with greater tumor cell kill, effective posttransplantation consolidation therapies, or strategies (*e.g.*, tumor vaccines) to induce an effective immune responsiveness to the residual disease after transplantation.

THERAPEUTIC PRINCIPLES

Autologous and Allogenic Transplantation

Autologous Transplantation

- Based on chemotherapy or radiotherapy dose-sensitivity of disease being treated
- Requires collection and storage of adequate hematopoietic stem cells (HSCs), preferably before extensive alkylating agent or purine analogue therapy
- Lower risk of graft failure (no immunological rejection)
- No routine posttransplantation immunosuppression
- Minimal risk of graft-versus-host disease (GvHD)
- No graft-versus-tumor (GvT) effect
- More rapid posttransplantation immune reconstitution
- Risk of tumor cell contamination in HSC product
- Not useful for diseases in which normal HSCs cannot be collected (*e.g.*, chronic myelogenous leukemia, myelodysplasia)

Allogeneic Transplantation

- Rescues bone marrow function after dose-intensive therapy
- Effective with reduced-intensity conditioning regimens
- Achieves a GvT effect in many malignancies
- Risk of GvHD distinct from the beneficial GvT effect
- Higher risk of transplantation-related complications that may offset the benefit of the GvT effect
- Risk of immunological graft rejection
- Slow posttransplantation immune reconstitution
- No risk of tumor-cell contamination with healthy donor

Allo-HSCT has a much lower relapse risk compared with auto-HSCT as a result of a beneficial immunological graft-versus-tumor (GvT) effect achieved by engraftment of the donor immune system. Allograft recipients, however, face a much higher risk of treatment-related mortality (TRM) from the detrimental immunological GvH response against healthy tissues of the patient. The principal complication of allo-HSCT is graft-versus-host disease (GvHD), which can occur early (acute GvHD, within the first several weeks) (Table 83.1) or late (chronic GvHD, months to several years) after transplantation. The rate of overall incidence of moderate to severe acute GvHD (aGvHD) is 35%–80% for all patients receiving a human leukocyte antigen (HLA)–matched related or unrelated donor stem cell transplant, and aGvHD is a primary cause of death in 10–20% of these patients. Chronic GvHD (cGvHD), a clinicopathologically distinctive form of this

TABLE 83.1 Acute Graft-Versus-Host Disease (GvHD)

Clinical Staging

Stage	Skin	Liver	Gut
1	Maculopapular rash <25% BSA	Bilirubin 2–3 mg/dL	Diarrhea 500–1000 mL/day, or Persistent nausea
2	Maculopapular rash 25–50% BSA	Bilirubin 3–6 mg/dL	Diarrhea 1000–1500 mL/day
3	Generalized erythroderma	Bilirubin 6–15 mg/dL	Diarrhea >1500 mL/day
4	Desquamation and bullae formation	Bilirubin >15 mg/dL	Pain ± ileus

Clinical Grading

Overall Grade	Skin	Liver	Gut	Functional Impairment
0 (none)	0	0	0	0
1 (mild)	1–2	0	0	0
2 (moderate)	1–3	1	1	1
3 (severe)	2–3	2–3	2–3	2
4 (life-threatening)	1–4	1–4	1–4	3

Adapted from: Harris AC, Young R, Devine S, et al. International, Multicenter Standardization of Acute Graft-versus-Host Disease Clinical Data collection: A Report from the Mount Sinai Acute GVHD International Consortium. Biol Blood Marrow Transplant 2016;22:4–10.

alloreaction, occurs in up to 80% of recipients, may involve aspects of regulatory T cell (Treg) dysfunction and autoimmune-like responses, and may require years of therapy before tolerance is achieved allowing withdrawal of immunosuppressant medications. Other significant complications of allo-HSCT relate to problems of inadequate reconstitution of the immune system of the patient and the concomitant risk of opportunistic infections, including viral and fungal infections that are rarely observed in healthy hosts. Despite the apparent risks associated with allo-HSCT, this treatment holds great promise as a curative therapy for several tumor types, particularly with strategies either to enhance the efficacy of the GvT effect or decrease the toxicity of the graft-versus-host (GvH) response. Efforts have been made in recent years to make HSCT less toxic by development of low-dose, nonmyeloablative conditioning regimens that allow the treatment of older patients or those with comorbid health issues that otherwise preclude treatment with high doses of chemotherapy.[1]

IMMUNE MECHANISMS RELATED TO ALLO-HSCT

Histocompatibility

The HLA major histocompatibility complex (MHC; Chapter 5) is the primary consideration in the selection of a donor for allo-HSCT, since its loci contribute significantly to host-versus-graft (HvG), leading to immunological rejection of donor HSC, and to GvH (leading to GvHD and GvT) reactions. HLA antigens are classified as class I (HLA-A, -B, -C) and class II (HLA-DR, -DQ, -DP) molecules, and typing of donors and patients can be performed using low- or high-resolution techniques. The low-resolution serological techniques can determine a phenotypic mismatch (*e.g.,* A02 vs A03), whereas high-resolution molecular techniques can identify allelic genotypic differences (*e.g.,* A*02:01 vs A*02:02). The risks of aGvHD, cGvHD, and transplantation-related mortality increase with the number of HLA mismatches, and ideally, genotypically matched unrelated donors are sought for patients lacking an HLA-identical sibling. Disparity at HLA-A, -B, -C, and DRB1 alleles are definite risk factors for survival after unrelated donor transplantation, whereas single HLA-DQ or -DP mismatches appear to be better tolerated and/or permissive, meaning they do not have a deleterious impact on clinical outcome.[2]

Not all mismatches result in a deleterious clinical outcome, and identification of permissive mismatches will improve the number of donors available to the patient. Algorithms accounting for possible permissive mismatching in the selection of unrelated donors are currently being developed and tested.[3]

GvHD and GvT occurring after HLA-compatible sibling transplantation demonstrate the importance of minor histocompatibility antigens (miHAs) on the outcome of allo-HSCT. MiHAs are derived from polymorphic sites in normal proteins between individuals and are constantly processed by proteasome activity and presented on the cell surface by MHC molecules; they can thus be recognized by T cells from a HLA-matched donor. As a result, hundreds of miHAs may be variably expressed on host tissues and can trigger an alloresponse from donor T cells, thereby causing GvHD. Unfortunately, only a few miHAs have been identified, such as those on the Y chromosome leading to a higher risk of GvHD in male recipients of cells from female donors; miHA matching of a male donor for a male recipient is not a component of current algorithms for donor selection.

Research in the field of additional criteria for donor selection is ongoing. Human NK cells possess clonally distributed, inhibitory receptors termed "killer cell immunoglobulin-like receptors" (KIRs) that recognize epitopes shared by groups of HLA class I alleles (KIR ligands). KIR–ligand mismatching in the GvH direction appears to result in lower risk of relapse- and nonrelapse-related mortality after allo-HSCT. Activating KIRs transduce signals to activate NK cells, and the presence of these cells is associated with a lower risk of leukemia relapse after unrelated and haploidentical transplantation as well as protection against certain viral infections, such as human immunodeficiency/acquired immunodeficiency syndrome (HIV/AIDS) or hepatitis C virus (HCV) infection. This effect appears to be more evident in haploidentical transplantation and in umbilical cord blood (UCB) transplantation. The benefit of KIR–ligand mismatching is less obvious (and more controversial) in unrelated transplantations and after transplantation with T cell–replete grafts in which the antitumor effects of NK cells may be obscured by GvH reactions mediated by T cells.

Graft-Versus-Host Disease

GvHD is caused by mature donor T cells contaminating the HSC inoculum, which recognize HLA or miHA differences expressed by host antigen-presenting cells (APCs) and tissues (Fig. 83.1).[4] Cytokines released from host cells after a patient has received dose-intensive tumor cytoreductive chemotherapy or radiotherapy conditioning create an inflammatory environment that enables the generation of a response of infused donor T cells against host antigens. This initiates a cascade of T-cell activation events, which results in proliferation, release of additional inflammatory cytokines, and the generation of effector T cells that can infiltrate target tissue, particularly the lymphoid system, intestinal tract, skin, and liver and mediate the destruction of host cells in those organs. Both CD4 and CD8 T cells can be

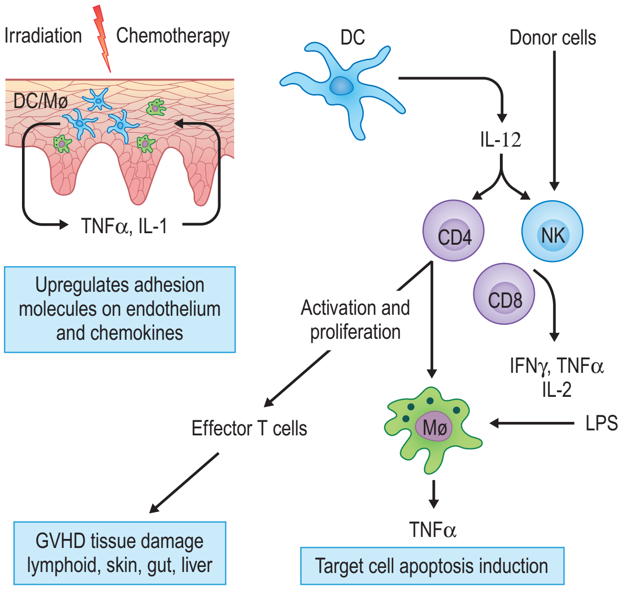

Irradiation
Chemotherapy
DC/Mø
TNFα, IL-1
Upregulates adhesion molecules on endothelium and chemokines
DC
Donor cells
IL-12
CD4
NK
CD8
Activation and proliferation
IFNγ, TNFα
IL-2
Effector T cells
Mø
LPS
TNFα
GVHD tissue damage lymphoid, skin, gut, liver
Target cell apoptosis induction

FIG 83.1 Development of Graft-Versus-Host Disease (GvHD). Cytoreductive preconditioning treatment of patients with hematological malignancies with total body irradiation (TBI) or chemotherapeutic drugs causes damage to epithelium in the skin and gastrointestinal tract and activates the release of inflammatory cytokines by dendritic cells (DCs) and macrophages (Mϕ) in those tissues. These cytokines include tumor necrosis factor-α(TNF-α) and interleukin (IL)-1, which upregulate adhesion molecules and chemokine release in the vascular endothelium of the tissues. Activated DCs also migrate to the lymphoid system, where they can present recipient histocompatibility antigens to infused donor T cells that are in the hematopoietic stem cell (HSC) graft. DCs release IL-12, which helps activate CD4 and CD8 T cells, as well as natural killer (NK) cells. These responding cells proliferate and produce additional inflammatory cytokines, including interferon-γ (IFN-γ), TNF-α, and IL-2. Mϕ are activated by both IFN-γ and lipopolysaccharide (LPS) produced by bacteria found in the intestinal tract, and these cells then produce high levels of more TNF-α. TNF-α has many properties, including direct induction of apoptosis on cells in the tissues of target organs of GvHD, but it also helps effector T cells to home to and enter tissue sites through the vascular endothelium. Effector T cells specific for host histocompatibility antigens then get reactivated and perform their effector functions, including release of inflammatory and cytolytic cytokines and direct killing of recipient-type cells in the lymphoid compartment and in the skin, gut, and liver.

involved in GvHD, depending on the specific class I or class II HLA or miHA disparities involved.

The simplest way to avoid the development of GvHD is to deplete the donor HSC graft of T cells before infusion to a cell dose below 10^5 cells/kg body weight. This approach has succeeded in significantly diminishing the incidence of GvHD, but other complications related to the ensuing delay in immune reconstitution of recipients, an increased risk of relapse from loss of the GvT effect, and a higher rate of engraftment failure (also from loss of the GvH effect) have resulted in inconsistent improvement in long-term survival, compared with T cell–replete products. The conundrum is that the same alloreactive donor T cells that mediate GvHD can also cause a GvT response, although there may also be additional tumor-specific or hematopoietic tissue specific T cells.[5] Thus the overriding goal is to be able to manipulate the donor HSC inoculum in such a way as to avoid GvHD but to still be able to mediate a GvT effect.[6]

KEY CONCEPTS

Graft-Versus-Host Disease (GvHD)

- Caused by donor–recipient differences in:
  - Major histocompatibility complex (MHC) molecules
  - Minor histocompatibility antigens (miHAs)
- Mediated by mature donor CD4 and/or CD8 T cells
- Requires inflammatory cytokines
- Primary target organs include lymphoid system, skin, gastrointestinal tract, and liver
- Acute and chronic forms

Clinical Aspects of aGvHD

Usually developing within the first 3 months after transplantation, aGvHD is a clinical diagnosis with characteristic but nondiagnostic pathological findings, with the most common presenting manifestations including skin rash; nausea, anorexia, and diarrhea; and jaundice, depending on the target organ(s) most affected.[7,8] In addition to the increased risks of developing aGvHD related to the extent of HLA and miHA disparity, additional factors include the advanced age of either the donor or the recipient, gender disparity (female donor–male recipient), donor parity (female donors), and infusion of T cell–replete HSC products. Conditioning with reduced-intensity regimens, with lower regimen-related toxicities to nonhematological tissues, also results in a lower risk of aGvHD and may delay the onset of its initial manifestation. Interactions between microbial-associated molecules and innate immune receptors (*e.g.,* Toll-like receptors [TLRs]) appear to be involved in GvHD pathogenesis, as demonstrated in both murine models and human transplantation. Research has shown an interaction between the host's gut microbiota and immune system.[9] On the basis of this knowledge, experiments to decrease transplantation-related GvHD by altering the gut microbiome are being actively pursued.

Pharmacological agents are the mainstay of aGvHD prophylaxis. Most patients receive a combination of a calcineurin inhibitor (tacrolimus or cyclosporine) along with an antimetabolite, such as methotrexate or mycophenolate mofetil (MMF). Methotrexate is associated with delayed engraftment, mucositis, idiopathic pneumonia syndrome, and other transplantation-related complications, which has prompted the development of other combination regimens, such as a calcineurin inhibitor in combination with sirolimus or MMF, or reduced doses of methotrexate. The addition of antithymocyte globulin (ATG) to the conditioning regimen lowers the incidence of both GvH and HvG reactions because of its persistence for several days after HSC infusion, effectively depleting T cells from the graft as well as the host. However, patients treated with ATG may face higher risks of infectious complications, including Epstein-Barr virus (EBV)–associated posttransplantation lymphoproliferative disorder as a result of the greater immunosuppression achieved. Novel combinations of agents continue to be explored to reduce the cost and morbidity of aGvHD and its treatments.

The administration of high-dose cyclophosphamide, an alkylating agent of the nitrogen mustard family, after transplantation reduces the risks of acute and/or chronic GvHD and is a prime example of drug-induced immunological tolerance, a concept first demonstrated by Schwartz and Dameshek in their experiments using 6-mercaptopurine.[10] Subsequently, in 1963, Berenbaum and Brown demonstrated immune tolerance to skin allografts in adult mice by using cyclophosphamide.[11] Post-transplantation cyclophosphamide given on days 3 and 4 after transplantation is now a standard regimen used in haploidentical allogeneic transplantations in combination with other immunosuppressive drugs and as a single agent in matched unrelated and related transplantations, reducing the complications of graft failure and life-threatening aGvHD associated with transplantation of HLA disparate grafts.[12]

Glucocorticoids with a calcineurin inhibitor remain the standard approach to initial systemic management of clinically significant aGvHD. About 30–50% of patients will respond to initial therapy, and patients who fail to respond have a poor prognosis, as additional agents added for control greatly increase the risk of opportunistic infections and other treatment-related complications. The use of higher doses of corticosteroids, or the addition of ATG, for example, in the initial treatment of aGvHD do not improve patient outcomes and should be reserved for patients who fail initial therapy. A number of drugs, including ATG, pentostatin, switching to tacrolimus from cyclosporine, and newer monoclonal antibodies (mAbs) have shown limited activity in the salvage treatment of patients with steroid-refractory aGvHD. Extracorporeal exposure of peripheral blood mononuclear cells (PBMCs) to the photosensitizing agent 8-methoxypsoralen and ultraviolet A (UVA) radiation (photopheresis) is effective in the treatment of selected diseases mediated by T cells, including both aGvHD and cGvHD, although the mechanism of this effect remains to be elucidated.

Autologous GvHD

A form of aGvHD can also occur after auto-HSCT, probably as a manifestation of immune system dysregulation during reconstitution of the immune system after dose-intensive therapy. Initially studied in murine models of auto-HSCT, it was demonstrated that the abrupt withdrawal of cyclosporine could induce clinical features nondistinguishable from that observed after allo-HSCT. A mechanism that involves the depletion of central memory cells has been proposed, although depletion of Tregs with inhibition of peripheral tolerance may also be involved. Autologous GvHD is also reported in patients not receiving posttransplantation immunomanipulations (spontaneous GvHD). Based on the clinical benefits of GvT observed in allo-HSCT, several clinical trials with administration of cyclosporine with or without interferon (IFN) were performed in patients undergoing auto-HSCT without obvious GvT effect.[13] Both induced and spontaneous autologous aGvHD are usually self-limiting complications, and are usually easily managed with a course of corticosteroids in patients with more extensive involvement, in contrast to the more extensive and difficult-to-manage aGvHD that occurs after allo-HSCT. Furthermore, cGvHD does not occur after auto-HSCT.

A spontaneously occurring severe, and sometimes fatal, reaction involving the gastrointestinal (GI) system with severe diarrhea during the periengraftment period is reported in patients undergoing auto-HSCT for MM and other malignancies ("engraftment syndrome").[14] The presentation is clearly distinct from the spontaneous or induced GvHD previously observed in patients undergoing auto-HSCT. "Engraftment syndrome" may be a consequence of prior treatments with immunomodulatory drugs, such as lenalidomide or thalidomide, the proteasome inhibitor bortezomib, or high doses of potent glucocorticoids before collection and storage of autologous HSCs, resulting in a period of immune dysregulation occurring early after autologous HSCT.[10] Patients with "engraftment syndrome" may require intensive treatment with corticosteroids and a calcineurin inhibitor, with courses more similar to those experienced by patients undergoing allo-HSCT.

Clinical Aspects of cGvHD

cGvHD is the leading cause of late transplant-related mortality (TRM) among those undergoing allo-HSCT and resembles autoimmune disorders, such as scleroderma, Sjögren syndrome, and primary biliary cirrhosis. Diagnosis is, as with aGvHD, based on clinical observations and secondary confirmation with laboratory or pathology tests (Table 83.2). A falling performance status, progressive weight loss, or recurrent infections are usually signs of severe cGvHD. About 50% of long-term survivors will develop cGvHD at a median of 9 months after transplantation, and patients must be monitored closely for this complication for at least 3 years so that appropriate treatment can be initiated before extensive end-organ damage ensues. Permanent ocular or oral sicca syndrome and pulmonary dysfunction can result if cGvHD is not appropriately treated, leading to a marked deterioration in quality of life.

Factors predictive of the development of cGvHD include the degree of HLA and miHA disparity, as well as prior aGvHD, older patient age, the source of HSCs (greater risk after peripheral blood stem cell [PBSC] transplantation than after bone marrow transplantation), gender (female donor–male recipient), and donor lymphocyte infusion (DLI) after transplantation. Inflammatory events, such as sunburn or surgical procedures, can precipitate cGvHD. Patients who develop cGvHD have a higher risk of TRM but a lower risk of relapse as a result of the immunological GvT effect.[15] T-cell depletion of the graft or treatment with ATG may decrease the risk of cGvHD, although this has not been demonstrated in all studies. Most patients require at least two drugs for effective treatment of cGvHD, with the standard initial treatment being glucocorticoids and a calcineurin inhibitor. About half the patients do not achieve a complete remission with first-line therapy, although the manifold signs and symptoms of cGvHD complicate the definition of response to treatment. There are no clear recommendations regarding second-line treatments and various pharmacological and immunological techniques have been used. Photopheresis has an overall response rate of 50–60%, with many patients achieving complete remission. At least transient responses can be achieved with rituximab treatment, an anti-CD20 chimeric antibody, which illustrates the humoral immunity contribution to cGvHD. The demonstration of activation of the Janus kinase (JAK) pathways in activated T lymphocytes has led to studies demonstrating a potential benefit of Jak2 inhibitors in the management of steroid-refractory cGvHD.

GvT Responses

Although recognized in mouse models of transplantation in the 1950s, the first clinical report of a relationship between GvHD and GvT was published in 1979.[16] This relationship between the incidence (but not the severity) of aGvHD or cGvHD and the relapse rate of chronic myeloid leukemia (CML) and, to a lesser extent, acute myeloid leukemia (AML), acute lymphocytic leukemia (ALL), and MM has been observed in patients who received allo-HSCT. Relapses after HSCT are thought to occur

TABLE 83.2 **Staging of Chronic Graft-Versus-Host Disease (GvHD)**

Target Organ	Score 0	Score 1	Score 2	Score 3
Performance score	KPS 100%	KPS 80–90%	KPS 60–70%	KPS <60%
Skin	No symptoms	<18% BSA	19–50% or sclerotic, still able to pinch	>50% or "Hidebound"
Mouth	No symptoms	Mild symptoms, no limitations	Moderate symptoms, decreased oral intake	Severe symptoms with major decrease in intake
Eyes	No symptoms	Mild dry eyes	Moderate dry eyes, drops >3×/day	Severe dry eyes affecting daily activities
Gastrointestinal tract	No symptoms	Symptoms without weight loss	Symptoms with moderate weight loss (5–15%)	Symptoms with weight loss >15%
Liver	Normal LFTs	LFTs elevated <2× upper limits of normal	LFTs elevated 2–5× upper limits of normal	LFTs elevated >5× upper limits of normal
Lungs	No symptoms	Mild symptoms FEV 60–79%	Moderate symptoms FEV 40–59%	Severe symptoms FEV <40%
Joints and fascia	No symptoms	Mild tightness not affecting daily activities	Tightness affecting daily activities	Contractures with significant loss of range of motion
Female genital tract	No symptoms	Symptomatic with middle signs on examination	Symptomatic with dyspareunia	Symptomatic with strictures

At least one diagnostic and one distinctive sign is necessary to make a diagnosis of cGvHD.
BSA, Body Surface Area; FEV, Forced Expiratory Volume; KPS, Karnofsky Performance Status; LFT, Liver Function Test.
Adapted from: Filipovitch AH, Weisdorf D, Pavletic S, et al. National Institutes of Health Consensus Development Project on Criteria for Clinical Trials in Chronic Graft-versus-Host Disease: 1. Diagnosis and Staging Working Group Report. Biol Blood Marrow Transplant 2005; 11: 945–56.

CLINICAL PEARLS

Factors Predicting Chronic Graft-Versus-Host Disease (GvHD)

- Degree of human leukocyte antigen (HLA) incompatibility
  - Major histocompatibility complex (MHC)
  - Minor histocompatibility antigens (miHAs) disparity
- Presence of prior acute GvHD
- Source of stem cells (higher risk in peripheral blood versus bone marrow)
- Donor gender (female donor → male recipient)
- Use of donor lymphocyte infusion (DLI) following hematopoietic stem cell transplantation (HSCT)
- Inflammatory events; surgeries, phototoxicity, alcohol consumption

because of survival of malignant cells harbored in bone marrow after administration of the pretransplantation conditioning regimen and their outgrowth several months later. Relapse remains the major cause of treatment failure after allo-HSCT despite the intended GvT effect of this treatment modality. The ability to mediate an effective GvT response likely depends on several factors, including the presentation of appropriate antigens by MHC class I and/or class II molecules on the tumor cells that can be recognized by effector CD4 or CD8 T cells; lack of strong Treg activity that may be induced by cytokines from the tumor cells; tumor cell susceptibility to lysis by effector T cells (*e.g.*, the level of B-cell lymphoma 2 (BCL-2) expression and the ability to resist apoptosis induction); ability of T cells to home to sites of tumor growth; and the direct effect of immunosuppressive cytokines, such as transforming growth factor-β (TGF-β), produced by the tumor cells.[17] Many types of tumor cells downregulate expression of MHC on their surface, and perhaps CML and AML are most susceptible to GvT responses because the myeloid lineage is adapted for antigen presentation and high MHC expression. A number of novel immunotherapeutic approaches are being developed to overcome these obstacles and enhance GvT responses, keeping in mind that GvHD has to also be avoided or minimized to improve outcomes (Fig. 83.2).

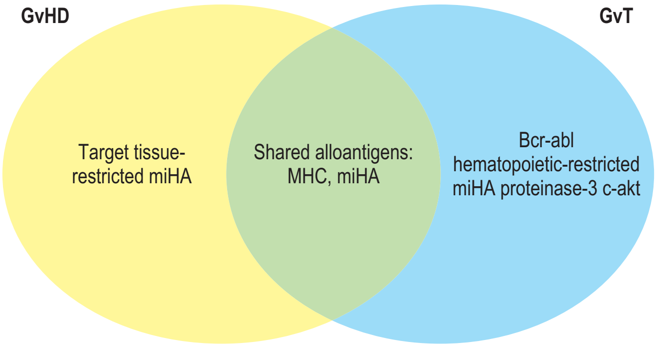

GvHD
GvT
Target tissue-restricted miHA
Shared alloantigens: MHC, miHA
Bcr-abl hematopoietic-restricted miHA proteinase-3 c-akt

FIG 83.2 Graft-Versus-Host Disease (GvHD) and Graft-Versus-Tumor (GvT) Responses. Donor T cell responses to recipient antigens can cause GvHD but can also target residual leukemia cells. T cells causing GvHD may recognize ubiquitous or tissue-restricted antigens (either major histocompatibility complex [MHC] or minor histocompatibility antigens [miHAs]). Many of these recipient antigens may also be expressed by the leukemia cells and allows for a GvT response. Additional leukemia-specific (*e.g.,* bcr-abl, proteinase 3, or c-akt) or tissue-restricted antigens (some miHAs, as those expressed only by certain lineages of hematological cells) may be dominantly expressed by the tumor cells and can be targeted by donor T cells without causing GvHD.

Immunotherapy remains the major strategy to combat relapse occurring after allo-HSCT. This can be achieved by reducing immunosuppression, offering a second allo-HSCT, or infusing additional lymphocytes from the HSC donor (DLI). As one can expect, all these modalities pose the challenge of associated morbidities, particularly GvHD.[6] More defined immunotherapeutic approaches using chimeric antigen receptors (CARs) expressed on T cells (CAR T cells; Chapter 77) also show some promise to target residual tumor cells that remain after allo-HSCT, with a possible low risk of GvHD. However, recent studies in murine models have raised a cautionary note concerning the capacity of CAR T cells to inflame GvHD.[18]

KEY CONCEPTS

Major Issues Related to Success of Allogeneic Hematopoietic Stem Cell Transplantation

- Graft-versus-host disease (GvHD)
- Graft-versus-tumor (GvT) responses
- Kinetics and completeness of immune reconstitution
- Opportunistic infections from delayed immune reconstitution
- Patient comorbid conditions increasing risk of toxicities
- Chemotherapy sensitivity of disease being treated

Immunomodulation is frequently the first treatment option for patients in relapse after allo-HSCT, with rapid withdrawal of immunosuppressive medications and infusion of donor T cells.[6] This treatment approach requires that donor T-cell chimerism be sustained to avoid rejection of effector cells through an HvG mechanism. Patients with low-grade lymphoid malignancies, such as CLL, indolent non-Hodgkin lymphoma (NHL), mantle cell NHL, and CML have the highest likelihood of response. A survey of 25 transplantation programs identified 140 patients who had received DLI and reported a complete response rate of 60% in patients with CML. Response rates were higher in patients with cytogenetic and chronic-phase relapse (75.7%) than in patients with accelerated-phase (33.3%) or blastic-phase (16.7%) relapse, and almost 90% of these patients remained in remission at 2 years after treatment.[19] The activity of DLI in AML and ALL is not as robust in comparison to that of in CML. However, it remains the main treatment strategy to treat relapse following allo-HSCT. DLI can be used in combination with chemotherapy to maintain control of disease control during the time required for the development of the GvT effect.

The number of lymphocytes infused is important in achieving the DLI effect, although it may be possible to induce GvT using doses of lymphocytes that are less likely to result in GvHD. A large retrospective analysis demonstrated that using an initially lower cell dose reduced GvHD and improved survival.[20] One method to reduce GvHD is to genetically insert suicide genes into the T cells being given for DLI, allowing specific ablation of these cells if this complication occurs.[21] Another method to lower GvHD rates complicating DLI is to deplete the DLI product of GvHD-inducing cells. Studies in mice have shown that naïve subset of CD8 T cells lead to more GvHD, whereas the effector memory subsets (T_{EM}) of CD8 and CD4 T cells moderate the graft-versus-leukemia (GvL) effect without causing GvHD.[22,23] The delay in response between DLI and the development of a GvT effect suggests that only a minority of the cells infused recognize the tumor cell antigens and must undergo *in vivo* expansion before the therapeutic effect is achieved. It may be possible to develop leukemia-specific cytotoxic T cells in the laboratory, decreasing the delay in effect and possibly increasing the GvT potential. Donor immunity can be transferred, at least transiently, to the host, as demonstrated by delayed-type transfusion reactions to host red blood cells (RBCs), mediated by donor lymphocytes transfused along with the HSC product.[24] Adoptive transfer of donor immunity against specific targets can be achieved, but its persistence requires immunization of both the donor and the recipient. A clinical utility for such manipulation of the immune system has not yet been demonstrated, but it is clearly of interest as a technique to prevent posttransplantation infections and/or disease relapse.

A rapid and major leap in the field of adoptive transfer of immunity is the development of CAR T–cell therapies for patients with aggressive lymphomas, ALL, MM, and CLL, resulting in long-lasting remissions following heavy pretreatment.[25] Developed in the late 1980s and requiring decades of fine-tuning and extensive clinical trials, CAR T cells have the ability to recognize and target tumor cells via their reprogrammed T-cell receptor (TCR) toward the malignant cell and costimulatory molecules engineered into the patient's own T cells. This form of treatment engaging a combination of gene therapy, cell therapy, and immunotherapy has led to unprecedented results in patients without robust treatment options for their aggressive malignancies or following relapse after allo-HSCT. The primary limitation of CAR T–cell therapy is the need to identify tumor-specific antigens to avoid the serious toxicity of "on target" but "off tumor" response.

Another potential approach to augment GvT responses following HSCT to prevent or treat relapse is the use of checkpoint inhibitors, currently accomplished by mAb blockade of negative regulatory signals from either cytotoxic T lymphocyte antigen-4 (CTLA4) or programmed death (PD)-1 proteins on the surface of activated T cells upon interaction with their respective ligands, B7-2 (CD86) and PD-L1 (B7-H1), on APCs and tumor cells (Chapter 77). Checkpoint inhibitors show strong clinical efficacy in the treatment of relapsed hematological malignancies, such as CLL and Hodgkin lymphoma.[26,27] In a phase I dose-escalation study using the anti-CTLA4 mAb ipilimumab for relapse after allo-HSCT, Bashey et al. concluded that ipilimumab did not induce or exacerbate clinical GvHD but could cause organ-specific immune adverse events and regression of malignancy following allo-HSCT.[28]

Adjuvant Therapy With HSCT

The relatively higher incidence of relapse after high-dose therapy leads to the hypothesis that chemoablation therapy with auto-HSCT, curative for some patients, could be viewed as a platform for other approaches effective in eliminating the minimal residual disease (MRD) in patients destined to experience relapse. Additional or yet higher-dose chemotherapy or radiotherapy, unless directly targeted to the tumor, increases the risk of nonhematopoietic toxicity and TRM from causes other than relapse. The correlation of more rapid lymphocyte recovery with a decreased risk of relapse, although likely a reflection of host factors and not direct evidence of a GvT effect after auto-HSCT, supports attempts to use HSCT as a tumor-debulking platform for posttransplantation immunotherapies. Immunotherapies are of interest in this regard and include administration of post-transplantation cytokines; the addition of tumor-specific antibodies, used before and/or after HSCT as an "*in vivo* purge"; and the development of tumor-specific vaccines, such as with tumor antigen–pulsed dendritic cells (DCs). The immunoglobulin (Ig) idiotype found in most patients with MM and indolent NHL, which is unique to the malignant clone, may serve as a unique tumor-associated antigen for DC pulsing. The treatment appeared to be well tolerated,[29] and after vaccination, idiotype-specific responses have been observed, characterized by T cell–proliferative responses with cytokine release and the production of antiidiotype antibodies, although, again, a clinical benefit remains elusive.

CLINICAL HSCT

Sources of HSCs

Bone marrow, peripheral blood stem cells (PBSCs), or UCB are all appropriate sources of HSCs for transplantation, with the primary differences being the quantities of HSCs in each product and the quantities and functions of other blood cells comprising most of the cellular content of the HSC product that may influence the immunological outcomes of transplantation. PBSCs have virtually replaced bone marrow as the HSC source for auto-HSCT and are widely used for allo-HSCT. Transplantation of PBSCs achieves more rapid hematological recovery in both the auto- and allo-HSCT settings, resulting in lower complication rates and lower costs of treatment. PBSCs appear to improve overall survival and disease-free survival of patients with advanced hematological malignancies who undergo allo-HSCT, although at the cost of an increased risk of extensive cGvHD.[30] Use of PBSCs results in a decreased CD4 T-helper cell-1 (Th1)/Th2 ratio, which may adversely affect the ability to counteract infections and may favor cGvHD development. It is also interesting to note that PBSC products tend to contain about 10 times more T cells than bone marrow products, yet those T cells are less functional, possibly because of inhibition by granulocytes activated by the granulocyte–colony-stimulating factor (G-CSF) used for HSC mobilization.[31]

UCB is a rich source of HSCs, with the major limitation for clinical use being the small quantity of cells collected, resulting in slower recovery of hematological function and increased risk of failure of sustained engraftment. A very significant advantage is the lower risk of GvHD because of the relative immaturity of the donor immune system, allowing for the use of HLA-mismatched units without a prohibitory increase in the risk of GvHD. In general, as with other sources of HSCs, outcomes of UCB transplantation reflect patient characteristics, with lower survival probabilities for patients with advanced diseases or poorer performance status at time of transplantation. A number of investigators proposed combining units of UCB to increase the cell dose infused into the patient, possibly increasing the speed of hematological recovery, but also (in theory at least, but not shown yet in clinical studies) potentiating the GvT effect and reducing the risk of relapse. *Ex vivo* expansion of one or more UCB units is also being studied as a mechanism to achieve more rapid hematological recovery.

Cell dose is an important predictor of outcome for both auto-HSCT and allo-HSCT, and HSCs comprise a very small portion (generally <1%) of the marrow, PBSC, or UCB product. It is now recognized that successful establishment of donor cell chimerism after allo-HSCT is a complex interplay of pre- and posttransplantation suppression of the host immune system, dose of HSCs and accessory cells (including donor lymphocytes) contained in the graft, and donor HLA compatibility (Fig. 83.3). HLA-mismatching, T-cell depletion, and less intensive pretransplantation conditioning regimens all raise the risk of graft failure. Importantly, the duration of aplasia predicts the incidence of TRM after auto-HSCT or allo-HSCT.[32] Auto-HSCT has a negligible risk of engraftment failure if the viability of HSCs is maintained during processing and storage. The speed of hematological recovery is related to the quantity of HSCs reinfused in an exponential relationship.[33] Increasingly higher CD34$^+$ cell (a surrogate marker of immature HSCs) doses result in greater likelihood of rapid recovery of PBSC counts. At lower doses, there is considerable heterogeneity in engraftment speed, especially for platelet recovery, with some patients experiencing quick engraftment despite low doses of PBSCs. Products containing ≥2-3 × 10^6 CD34$^+$ cells/kg recipient weight have more consistent rapid granulocyte and platelet engraftment.

Similarly, cell dose is a predictor of outcome for patients undergoing allo-HSCT. The actuarial 5-year TRM, overall survival, and disease-free survival significantly favored patients receiving marrow or PBSC grafts with higher cell doses,[34] but exceeding a CD34$^+$ cell dose of about 10 × 10^6/kg results in lower survival, likely as a result of higher risk of chronic GvHD.[35]

HSC products are not uniform in their characteristics, and the relative contributions of CD34$^+$ cell dose and the doses of

Influence of graft, donor, and host factors on allogeneic HSC engraftment

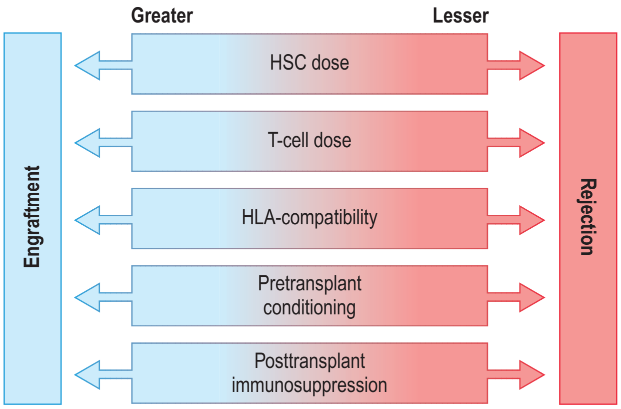


FIG 83.3 **Effect of Donor and Graft Characteristics on Allogeneic Hematopoietic Stem Cell Transplantation (Allo-HSCT).** Hematological engraftment requires that the host-versus-graft (HvG) reaction be overwhelmed, either by administration of a more intensive conditioning regimen before transplantation, use of more immunosuppressive medications after transplantation, closer matching of donor with host, and infusion of donor lymphocytes that can affect development of GvHD.

KEY CONCEPTS

Selection of Hematopoietic Stem Cell (HSC) Products for Transplantation

Bone Marrow

- Adequate quantities of cells can be obtained from most patients and donors
- Cytokine administration before harvesting may increase the quantities of HSCs collected and the number and function of accessory cells affecting transplant outcomes
- Lower risk of chronic graft-versus-host disease (GvHD) than peripheral blood stem cells (PBSCs)

PBSCs

- Faster hematological recovery than with bone marrow or umbilical cord blood (UCB)
- Better survival after related-donor hematopoietic stem cell transplantation (HSCT) for patients with advanced malignancy
- Lower tumor cell contamination than with bone marrow collected from autologous patients
- Cytokine administration before harvesting may increase the quantities of hematopoietic stem cells (HSCs) collected and the number and function of accessory cells affecting transplantation outcomes

UCB

- Relative immaturity of donor immune system permits multiple-antigen mismatched transplantation
- Transplantation outcomes similar to mismatched unrelated bone marrow transplantation
- Slower hematological recovery than with either PBSC or bone marrow
- Availability of stored units facilitates transplantation for patients with immediate need of treatment

accessory cells, including CD4, CD8, and NK-cell populations, on the outcome of transplantation remain unconfirmed. Whether or not this is clinically significant, the ratio and quantity of cell populations collected will be affected by changes in the cytokines and chemokines, used singly or in combination, to mobilize HSCs into peripheral blood for collection or to alter bone marrow products collected from patients and donors.

Purging of Cell Populations

For auto-HSCT, postcollection attempts to purge tumor cells potentially contaminating the graft have not demonstrated a survival advantage for the transplant recipient. T-cell depletion of donor HSCs by various methods is an effective means of reducing the incidence of aGvHD or cGvHD; but it increases the risk of graft failure, opportunistic infections, and relapse, effectively nullifying any advantage. Higher doses of HSCs, which can be achieved by the use of large quantities of PBSCs or by combining marrow and PBSC components, will offset this risk of graft failure. Partial T-cell depletion, as opposed to maximal T-cell depletion, may also reduce the risk of graft failure. Some centers are exploring posttransplantation T-cell add-back after the acute inflammatory effects of the conditioning regimen have resolved in an effort to maintain the GvT effect while still reducing GvHD morbidity and mortality.

Expansion of HSC Products

Ex vivo expansion of UCB HSCs is one potential mechanism to offset the low cell dose and delayed hematological recovery after transplantation. However, no expansion technique has yet achieved this goal, and such endeavors are complicated by the difficulty in identifying the pluripotent HSCs in comparison with lineage-committed progenitor cells. Furthermore, it is likely that expansion techniques will result in T-cell depletion, resulting in the other complications described above.

Hematological Recovery

HSC engraftment encompasses two concepts: *(i)* recovery of hematopoietic and immunological function and *(ii)* the rate at which this recovery occurs. Delay in or failure of sustained engraftment after myeloablative-conditioning regimen administration greatly increases treatment morbidity and cost. Engraftment failure can occur as a result of inadequate HSC quantity from poor collection or loss in postcollection processing, inadequate host support of the infused cells, posttransplantation events or medications, or HvG rejection (see Fig. 83.3). Engraftment failure is a very rare complication of auto-HSCT and is most likely a consequence of poor preservation of HSCs after collection. In allo-HSCT, the risk of engraftment failure is proportional to the donor HLA–miHA disparity, occurring more commonly in unrelated donor transplants than in sibling donor transplants, and with HLA-mismatched transplants. Engraftment failure is also increased by T-cell depletion of marrow inoculum because of the loss of the GvH effect against residual host immune cells.

Chimerism assessment is important in evaluating graft function after allo-HSCT. A fall in PBSC counts could indicate HvG rejection of the graft or early relapse after transplantation or could result from GvHD or viral infection. Documentation of stable persistence of donor T cells in the recipient's blood will help discriminate between these possibilities. It is also important that sustained lymphoid chimerism be demonstrated if DLI is to be used in the treatment of disease relapse after transplantation. The level of donor–host chimerism after allo-HSCT is best demonstrated through evaluation of single nucleotide tandem repeats (short tandem repeats [STRs]) by using molecular analysis. Obviously, such studies are of no value in assessing engraftment after auto-(syngeneic)HSCT.

Much of the emphasis in transplantation has been on myeloid engraftment because initial patient survival depends on recovery of phagocytes and, to a lesser extent, platelets. Immune reconstitution, and in particular donor T-cell reconstitution, in patients receiving HSCT is often hampered by older recipient age, diminished functional status of the thymus, cytokine milieu at the time of transplantation, and posttransplantation immunosuppressive treatments. The thymus involutes rapidly after childhood, and in the older adult, it is only able to contribute a very small portion to the mature T-cell compartment. The thymic tissue may be damaged as a result of a myeloablative conditioning regimen, or it can also be a target of alloreactive donor T cells mediating GvHD. As a result, restoration of the T-cell compartment in patients is often slow, particularly for CD4 T cells, and may be at suboptimal levels for many months to over a year. This situation, of course, endangers the ability of the patient to stave off opportunistic infections, such as from the herpes family of viruses and fungal pathogens. If donor T cells are provided in the HSC inoculum, some reconstitution of the T-cell repertoire, mostly CD8 T cells, is provided by the mechanism of nonthymic homeostatic expansion, although the level of diversity may be limited. Experimentally, administration of cytokines, such as interleukin-7 (IL-7), after HSCT, can enhance thymic function and help donor T-cell reconstitution. B-cell reconstitution, in contrast, is not that problematic in terms of the regeneration of the immune repertoire, although the ability to actually respond effectively to an infection with antibody production may still depend on the availability of antigen-specific CD4 T cells. Administration of Ig to patients with low IgG levels can prevent some of the infectious complications. Patients receiving HSCT who are conditioned with myeloablative regimens usually attain high levels of donor chimerism in their lymphoid compartment within a few months of transplantation. This often correlates with the ability of alloreactive donor T cells, capable of mediating GvHD, to target residual recipient HSC elements so that the primary source of *de novo* lymphoid reconstitution will be from the donor origin. By the same token, high donor chimerism is also associated with a lower incidence of relapse of malignancy.

CONDITIONING REGIMENS

Dose-Intensive and Reduced Intensity Chemotherapy

The pretransplantation regimen is intended to accomplish two goals: ablate the tumor and achieve adequate immunosuppression to allow donor engraftment. For auto-HSCT, only the dose-sensitivity of the tumor being treated need be considered. Lower dose, nonmyeloablative regimens are not used in auto-HSCT because with such regimens, infusion of HSCs to reconstitute marrow function would not be needed. Total body irradiation (TBI) was initially used for conditioning of transplant recipients. This modality achieves tumor cytotoxicity; treatment of sanctuary sites of disease, such as the central nervous system (CNS) and testes; and profound immunosuppression. TBI is usually combined in sequence with chemotherapy agents, such as cyclophosphamide or etoposide. Busulfan-based regimens were developed as alternatives to TBI for patients who had received prior dose-limiting radiotherapy and to avoid the effects of TBI on growth and development in children. A review of several studies that compared the use of busulfan and TBI found no statistically significant difference in survival or disease-free survival for patients with CML or AML.[36] Specific regimens are commonly used in the treatment of certain malignancies, such as dose-intensive melphalan in the treatment of MM and carmustine (BCNU)–containing regimens used in the treatment of lymphoma.

CLINICAL PEARLS

Conditioning Regimens for Hematopoietic Stem Cell Transplantation (HSCT) Transplantation

Autologous HSCT

- Allows dose-intensive therapy
- Allows use of marrow-toxic agents
- Regimens designed for optimal tumor cytotoxicity

Allogeneic HSCT

- Must achieve adequate patient immunosuppression (reduce host-versus-graft [HvG] reaction) to achieve engraftment
- Allows but does not require dose-intensive therapy
- Allows but does not require marrow-toxic agents
- Need not be tumor specific

The myeloablative conditioning regimens currently used have been tested in dose-escalation studies to achieve the maximal tolerated doses in otherwise healthy patients. Nonmarrow toxicities, such as pneumonitis, mucositis, and hepatic venoocclusive disease, limit further dose escalation of standard TBI- or chemotherapy-based regimens. New approaches include the addition of targeted therapies to the conditioning regimen, such as tumor-directed mAbs or radioimmunoconjugates that will not increase the toxicity to other organs. Tandem transplantation with the combination of a dose-intensive regimen with auto-HSCT followed, after recovery from the immediate regimen-related toxicities, by allo-HSCT using a reduced-intensity regimen is a novel approach to combine the benefits of each transplantation modality.

Hematological malignancies are predominantly diseases of older adults. The potent GvT effect that is observed after allo-HSCT allows allograft recipients to be treated with lower-dose nonmyeloablative regimens with the immunosuppressive properties of the regimen to reduce HvG reactions and facilitate engraftment becoming more important than direct cancer cytotoxicity. The primary requirement in developing a reduced-intensity regimen is the need to achieve adequate immunosuppression to permit the development of hematopoietic chimerism, which became feasible with the development of the purine analogue family of drugs. A variety of regimens are available, including combinations of fludarabine with melphalan and fludarabine with busulfan. Among the least toxic are regimens that involve a single fraction of TBI, based on the work by Storb et al.,[37] who proposed that the HvG reaction leading to HSC rejection and the GvH reaction could both be modified by an appropriate immunosuppressive regimen administered after transplantation, allowing a reduction in the intensity of the pretransplantation conditioning regimen.[37,38]

HSCT FOR INDIVIDUAL DISEASES

The treatment of diseases by auto- or allo-HSCT continues to evolve as the understanding of the biology of these diseases becomes more clearly understood. Improvements in nontransplantation treatments available to patients, such as the development of targeted tyrosine kinase inhibitors, which are very effective in the initial treatment of CML, will reduce the numbers of patients requiring HSCT in the management of their disease. Organizations, such as American Society of Blood and Marrow Transplantation, have published guidelines for treatment and reviews of the efficacy of treatment, identifying areas requiring additional research.[39,40]

Acute Myelogenous Leukemia

The primary clinical questions in HSCT in the treatment of AML concern patient selection and timing of treatment. Most patients with AML will achieve remission with initial chemotherapy, but even with appropriate postremission consolidation, the majority of patients (≈65%) will relapse within 1–2 years. Older age, the presence of defined cytogenetic abnormalities, inability to achieve a complete remission (CR) with the initial course of therapy, and history of a preceding marrow disorder or receipt of prior chemotherapy ("secondary AML") are predictors for failure of nontransplantation therapy. Patients with these adverse risk factors may be offered HSCT in first remission in place of nontransplantation consolidation chemotherapy. Numerous studies of patients entering their first remission compared standard consolidation therapy with dose intensification with auto- or allo-HSCT. In general, auto-HSCT was not shown to be more effective than nontransplantation consolidation chemotherapy, whereas allo-HSCT had the lowest risk of relapse. Despite the higher transplantation-related complications, in at least two major studies, patients assigned to allo-HSCT achieved a significantly better disease-free survival (DFS) compared with patients assigned to either chemotherapy or auto-HSCT. Allo-HSCT in first remission is particularly beneficial for patients with adverse risk features, achieving about a 50–70% DFS, and is the treatment of choice for patients with adverse risk cytogenetics or leukemia that arises from prior chemotherapy or other marrow diseases.

Myelodysplastic Syndromes

Myelodysplastic syndrome (MDS) comprises a heterogeneous group of clonal hematological disorders characterized by expansion of abnormal HSCs engendering variable degrees of cytopenia and frequent evolution to AML. Currently, allo-HSCT is the only modality of treatment that can achieve long-term control of disease; auto-HSCT is not feasible because of the inability to collect normal HSCs from these patients. The best results are seen in patients with earlier-stage disease, although for patients with earlier stage MDS, a "watchful waiting" approach, with transplantation performed at the time of disease progression, may be appropriate.

Chronic Myelogenous Leukemia

Allo-HSCT is an appropriate treatment for CML with long-term survival rates >80% for younger patients undergoing related donor transplantation within the first year after diagnosis. However, inhibitors of the tyrosine kinase (TK) encoded by the Philadelphia chromosome have relegated transplantation to the treatment of patients with advanced disease or the rare patient with CML not responsive to that targeted therapy. The probability of long-term survival is lower for patients who undergo allo-HSCT more than 1 year after diagnosis and for patients with more advanced diseases. The effect of prior therapy with a TK inhibitor on the outcome of transplantation is not known. CML is highly responsive to the immunological GvT effect, and many patients in relapse after transplantation can be salvaged by the administration of

DLI. CML is a stem cell disease, and the collection of normal HSCs is unlikely for most patients, precluding auto-HSCT. Of speculative interest is the collection of autologous cells after cytogenetic remission is achieved with the use of a TK inhibitor, for subsequent use at the time of disease relapse.

Myeloproliferative Diseases

MPDs are clonal stem cell diseases and include CML (discussed above), polycythemia vera, primary myelofibrosis, essential thrombocytosis, chronic myelomonocytic leukemia, and MPD not otherwise characterized. These disorders display overlapping clinical features and may also show features more characteristic of MDS, but they exhibit different natural histories and different therapeutic requirements. Allo-HSCT is effective in ablating the abnormal clone in these disorders. The timing of treatment is important because of the frequently long natural history of these diseases. Anemia, older age, and cytogenetic abnormalities all predict poor survival in primary myelofibrosis, suggesting that patients with anemia and an abnormal karyotype should proceed to allo-HSCT.

Acute Lymphoblastic Leukemia

In contrast to treatment of the pediatric patient with ALL (Chapter 78), only very few adults with ALL are cured with nontransplantation induction and consolidation regimens. As with AML, adverse risk features include certain cytogenetic changes and the inability to achieve remission with initial induction chemotherapy. Allo-HSCT in first remission is clearly effective in the management of patients with defined adverse risk cytogenetics, such as translocation involving chromosome 4 and 11, or 9 and 22 (Philadelphia chromosome). Recent randomized studies also demonstrate a survival advantage for patients who have intermediate risk of disease and are undergoing allo-HSCT, especially since current salvage regimens for patients in relapse have a very low likelihood of achieving a second remission. There is little evidence to support the effectiveness of auto-HSCT despite its theoretical potential.

Chronic Lymphocytic Leukemia

Many patients and physicians are reluctant to support aggressive treatment of CLL with allo-HSCT because of the frequently indolent nature and long natural history of this disease, as well as the advanced age of most patients at time of diagnosis. Yet disease progression is inevitable, and the aggressiveness of this disease can be predicted by various features, including chromosome 17p and 11q deletions and detection of ZAP-70 mutation and CD38 expression. Therefore HSCT should be considered a treatment choice, especially for younger patients or those with adverse risk features.

The exquisite sensitivity of the CLL cells to the GvT effect allows for the use of reduced-intensity regimens with lower risks of regimen-related mortality in these, generally, older patients. Durable control of disease can be achieved for about 50% of patients. The extensive infiltration of bone marrow by malignant lymphocytes and the lack of the GvT effect preclude auto-HSCT. In the few clinical studies that have been reported, high CR responses have been obtained with auto-HSCT, but relapses were the frequent cause of treatment failure.

Multiple Myeloma

MM is a malignancy of plasma cells (Chapter 80), with a median patient age of 65 years at time of diagnosis. In MM, dose-intensive melphalan with auto-HSCT can achieve about a 40% complete response rate. Tandem auto-HSCT with two cycles of dose-intensive therapy at 2- to 6-month intervals can improve progressive free survival and in some studies, overall survival. This approach to treatment is still debated. Ultimately, most patients experience relapse and succumb to the disease. Although auto-HSCT is not considered curative for this disease, the median duration of time without treatment is 2–3 years, and ≈15% of patients may not need treatment for progressive disease for 5 years or longer, which greatly improves their quality of life. Allo-HSCT with use of a reduced-intensity conditioning regimen decreases the risk of TRM to <10%, but patients with advanced disease are unlikely to respond to this treatment. The lower risk of complications with nonmyeloablative conditioning regimens has led to studies of tandem auto-HSCT, using an intensive chemotherapy regimen to achieve tumor debulking, followed 2–6 months later by allo-HSCT using a reduced-intensity regimen to achieve the GvT effect. This approach to the treatment of MM is being studied in large multicenter trials, with some studies showing a survival advantage for recipients of allo-HSCT.

Non-Hodgkin Lymphoma

The lymphomas (Chapter 79) represent a diverse group of malignant diseases of B and T lymphocytes, comprising some of the slowest- to the fastest-growing human malignancies, with a range of curability achieved by nontransplantation therapies. As a group, the lymphomas exhibit a strong dose–response relationship to chemotherapy or radiotherapy, and the benefit of dose-intensive treatment with auto-HSCT has been well established. Auto-HSCT avoids the morbidity of allo-HSCT and is the preferred approach for the majority of patients, with some prominent exceptions discussed below. Popular chemotherapy regimens include cyclophosphamide, BCNU, and etoposide (CBV) or BEAM (BCNU, etoposide, cytarabine, and melphalan). However, no single chemotherapeutic or radiation-based regimen has emerged as a superior treatment.

Low-Grade NHL

Low-grade NHL, in general, exhibits a variable and prolonged natural course, with many patients not requiring treatment until symptoms or organ toxicity appear. Therefore most of the experience with HSCT has been in patients after initial relapse rather than at the time of initial diagnosis. Few randomized studies comparing auto-HSCT with nontransplantation therapies for patients have been reported. A number of phase II and registry data have been published for auto-HSCT, and although response rates are high, a continuing pattern of relapse has been observed. In contrast, patients who undergo allo-HSCT experience a higher probability of TRM but a lower risk of relapse after transplantation. Indolent B-cell NHL, similar to other low-grade diseases, such as CLL, appears to be very sensitive to the GvT effect of allo-HSCT, and using reduced-intensity regimens will result in lower transplantation-related complications. The difference in relapse rates between auto-HSCT and allo-HSCT could result from the possible reintroduction of lymphoma cells in the HSC product but is more likely from the lack of a GvT effect of auto-HSCT.

Aggressive NHL

Auto-HSCT is the standard of care for patients with B-cell NHL in first “chemotherapy-sensitive” relapse. The success of this therapy reflects the extent and the responsiveness of the disease to chemotherapy at the time of transplantation, with relapse the major cause of treatment failure. The Parma randomized trial for high-dose therapy followed by auto-HSCT reported a significantly higher event-free survival for patients treated with auto-HSCT than for the group receiving standard-dose treatment (46% vs 12%). Furthermore, it is notable that no patients assigned to the conventional-dose salvage therapy could be rescued at the time of the second relapse with delayed transplantation. The efficacy of auto-HSCT is diminished for patients previously treated with rituximab, an anti–B-cell chimeric antibody, and who relapse within 12 months of initial therapy.[41] Other circumstances, in which HSCT may be indicated, include HSCT for patients who respond slowly to initial therapy, have high-risk disease, are resistant to initial therapy, or are in relapse with chemotherapy-insensitive disease.[40] Results for patients with truly chemotherapy-refractory disease are, however, poor, and such patients should be considered for allo-HSCT.

Mantle cell lymphoma is known for its unremitting clinical course when treated conventionally and has proven relatively resistant to dose-intensive treatment, especially when used for management of relapsed disease. Mantle cell NHL appears to be very sensitive to the GvT effect of allo-HSCT allowing treatment with reduced-intensity regimens.

Burkitt lymphoma, Burkitt-like lymphomas, and lymphoblastic lymphomas are high-grade NHLs associated with relatively poor long-term survival rates. The role of both auto-HSCT and allo-HSCT in these disorders remains unclear. There are no convincing risk models that identify clear indications for transplantation as part of the initial therapy. The lack of an obvious GvT effect in these disorders suggests that auto-HSCT can provide a reasonable treatment. Transplantation registry data suggest that disease status at the time of transplantation is the most important predictor of outcome in patients with high-grade disease.[42] T-cell lymphomas are less common than B-cell lymphomas. There is no well-defined management strategy for these disorders, and treatment is based on the disease stage and immunopathological grade. Disease control with T cell–type treatment regimens, compared with the more common B cell–type treatment regimens, is lower, although the comparison may not be appropriate, as T-cell diseases present at more advanced stages compared with B-cell diseases. The GvT effect may be more evident in T-cell NHL, particularly for virus-associated NHL, and allo-HSCT is the preferred treatment for some patients.

Hodgkin Lymphoma

Many patients with Hodgkin lymphoma will achieve durable remissions with nontransplantation chemotherapy and/or radiation therapy, and algorithms for staging and treatment of this disease are well defined. Dose-intensive therapy with auto-HSCT is available to those patients who do not achieve a remission or who relapse after initial therapy. The outcome for patients whose remission lasted <1 year is dismal with standard dose second-line treatments, and approximately 40–50% of patients with Hodgkin lymphoma who suffer a relapse within 1 year will achieve durable remissions after auto-HSCT. This approach can also overcome drug resistance and lead to an overall survival rate of 34–50% for patients with chemotherapy-refractory disease.

Allo-HSCT is not the first choice for the treatment of relapsed Hodgkin lymphoma because of the higher transplantation-related complications despite the evidence of an effective GvT effect. Allo-HSCT is offered to patients who suffer relapse after auto-HSCT or who are not candidates for auto-HSCT because of aggressive, chemotherapy-refractory disease.

Solid Tumors

Skin and colonic mucosa are primary targets of both acute and cGvHD, and this would suggest that allo-HSCT would be effective therapy in the treatment of cancers of these organs. Yet allo-HSCT is not effective in the control of these cancers, illustrating the discrimination between target antigens of normal tissues of GvH, such as the colonic crypt cells, compared with antigens expressed by tumors derived from these tissues targeted by GvT. With the exception of allo-HSCT in the treatment of renal cell cancer, in which a GvT effect was seen to be clinically evident in some but not all studies, transplantation in the treatment of solid tumors, such as neuroblastoma, Wilm tumor, and germ cell tumors, is limited to auto-HSCT with one or more cycles of dose-intensive chemotherapy.

FUTURE DIRECTIONS

Advances in HLA typing, chemotherapy conditioning regimens, supportive care, and our understanding of the biology of the HvG and GvH reactions have greatly reduced the toxicity of both auto-HSCT and allo-HSCT. Although the diseases treatable with HSCT are those characterized by dose sensitivity, currently available dose-intensive conditioning regimens have been pushed to maximal tolerable doses. Newer approaches to the treatment of these malignancies will likely include the addition of therapies that target the malignancy, such as radioimmunoconjugates that will add minimal toxicity to other organs. The availability of reduced-intensity regimens permits tandem transplantation, using dose-intensive therapy with auto-HSCT to achieve maximal tumor-cell debulking, followed, after recovery from transplantation-related toxicities, by Allo-HSCT, using a reduced-intensity regimen to achieve an immunological GvT effect. Advancements in the field of haploidentical HSCT have facilitated immediate availability of donors for patients with high-risk malignancies when selection of unrelated donors is limited by ethnicity and time constraints. Finally, the ability to use HSCT as a platform for protein- or cellular-based vaccination strategies is the subject of considerable interest.

ON THE HORIZON

Allogeneic Hematopoietic Stem Cell Transplantation (HSCT)

Development of donor selection algorithms
Strategies to maximize graft-versus-tumor (GvT) effect
Advent of haploidentical HSCT allows for easy donor availability
Refining haploidentical HSCT techniques for different diseases and use of KIR testing for donor selection
Combination conditioning regimens
Developing regimens with lower toxicities
Posttransplantation maintenance/treatment
Posttransplantation maintenance/treatment for high risk of relapse after HSCT
Immune effector cell therapies
Platform for other treatments; chimeric antigen receptor (CAR) T–cell therapy, checkpoint inhibitor blockade

Autologous HSCT

Tumor-specific conditioning regimens
Improve ratio of toxicity to tumor-cell cytotoxicity
Posttransplantation vaccination strategies
Checkpoint inhibitor blockade in clinical trials for high-risk malignancies

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Sorror ML, Maris MB, Storer B, et al. Comparing morbidity and mortality of HLA-matched unrelated donor hematopoietic cell transplantation after nonmyeloablative and myeloablative conditioning: influence of pretransplantation comorbidities. Blood 2004;104:961–8.
2. Pidala J, Lee SJ, Ahn KW, et al. Nonpermissive HLA-DPB1 mismatch increases mortality after myeloablative unrelated allogeneic hematopoietic cell transplantation. Blood 2014;124:2596–606.
3. Confer DL, Abress LK, Navarro W, et al. Selection of adult unrelated hematopoietic stem cell donors: beyond HLA. Biol Blood Marrow Transplant 2010;16:S8–S11.
4. Socie G, Blazar BR. Acute graft-versus-host disease: from the bench to the bedside. Blood 2009;114:4327–36.
5. Bleakley M, Riddel SR. Molecules and mechanisms of the graft-versus-leukaemia effect. Nat Rev Cancer 2004;4:371–80.
6. Kolb H-J. Graft-versus-leukemia effects of transplantation and donor lymphocytes. Blood 2008;112:4371–83.
7. Glucksberg H, Storb R, Fefer A, et al. Clinical manifestations of graft-versus-host disease in human recipients of marrow from HLA-matched sibling donors. Transplantation 1974;18:295–304.
8. Przepiorka D, Weisdorf D, Martin P, et al. 1994 Consensus conference on acute GvHD grading. Bone Marrow Transplant 1995;15:825–8.
9. Mathewson ND, Jenq R, Mathew AV, et al. Gut microbiome-derived metabolites modulate intestinal epithelial cell damage and mitigate graft-versus-host disease. Nat Immunol 2016;17:505–13.
10. Schwartz R, Dameshek W. Drug-induced immunological tolerance. Nature 1959;183:1682–3.
11. Berenbaum MC, Brown IN. Prolongation of homograft survival in mice with single doses of cyclophosphamide. Nature 1963;200:84.
12. Robinson TM, O'Donnell PV, Fuchs EJ, et al. Haploidentical bone marrow and stem cell transplantation: experience with post-transplantation cyclophosphamide. Semin Hematol 2016;53:90–7.
13. Yeager AM, Vogelsang GB, Jones RJ, et al. Induction of cutaneous graft-versus-host disease by administration of cyclosporine to patients undergoing autologous bone marrow transplantation for acute myeloid leukemia. Blood 1992;79:3031–5.
14. Keung YK, Beaty MW, Pettenati M, et al. Possible role of engraftment syndrome and autologous graft-versus-host disease in myelodysplastic syndrome after autologous stem cell transplantations: retrospective analysis and review of the literature. Clin Lymphoma Myeloma Leuk 2010;10:129–33.
15. Pidala J, Kim J, Anasetti C, et al. The global severity of chronic graft-versus-host disease, determined by National Institutes of Health consensus criteria, is associated with overall survival and non-relapse mortality. Haematologica 2011;96:1678–84.
16. Weiden PL, Flournoy N, Thomas ED, et al. Antileukemia effect of graft-versus-host disease in human recipients of allogeneic-marrow grafts. N Engl J Med 1979;300:1068–73.
17. van den Brink MRM, Porter DL, Giralt S, et al. Relapse after allogeneic hematopoietic cell therapy. Biol Blood Marrow Transplant 2010;16:S138–S145.
18. Jacoby E, Yang Y, Qin H, et al. Murine allogeneic CD19 CAR T cells harbor potent antileukemic activity but have the potential to mediate lethal GVHD. Blood 2016;127:1361–70.
19. Collins RH, Shpilberg O, Drobyski WR, et al. Donor leukocyte infusions in 140 patients with relapsed malignancy after allogeneic bone marrow transplantation. J Clin Oncol 1997;15:433–44.
20. Guglielmi C, Arcese W, Dazzi F, et al. Donor lymphocyte infusion for relapsed chronic myelogenous leukemia: prognostic relevance of the initial cell dose. Blood 2002;100:397–405.
21. Ciceri F, Bonini C, Stanghellini MT, et al. Infusion of suicide-gene-engineered donor lymphocytes after family haploidentical haematopoietic stem-cell transplantation for leukaemia (the TK007 trial): a non-randomised phase I-II study. Lancet Oncol 2009;10:489–500.
22. Anderson BE, McNiff J, Yan J, et al. Memory CD4+ T cells do not induce graft-versus-host disease. J Clin Invest 2003;112:101–8.
23. Chen BJ, Deoliveira D, Cui X, et al. Inability of memory T cells to induce graft-versus-host disease is a result of an abortive alloresponse. Blood 2007;109:3115–23.
24. Rowley SD, Donato ML, Bhattacharyya P. Red blood cell incompatible allogeneic hematopoietic progenitor cell transplantation. Bone Marrow Transplant 2011;46:1167–85.
25. Jackson HJ, Rafiq S, Brentjens RJ. Driving CAR T-cells forward. Nat Rev Clin Oncol 2016;13:370–83.
26. Bachireddy P, Burkhardt UE, Rajasagi M. Haematological malignancies: at the forefront of immunotherapeutic innovation. Nat Rev Cancer 2015;15:201–15.
27. Ansell SM, Lesokhin AM, Borrello I, et al. PD-1 Blockade with nivolumab in relapsed or refractory Hodgkin's lymphoma. N Engl J Med 2015;372:311–19.
28. Bashey A, Medina B, Corringham S, et al. CTLA4 blockade with ipilimumab to treat relapse of malignancy after allogeneic hematopoietic cell transplantation. Blood 2009;113:1581–8.
29. Leitch HA, Connors JM. Vaccine therapy for non-Hodgkin's lymphoma and other B-cell malignancies. Curr Opin Investig Drugs 2005;6:597–604.
30. Anasetti C, Logan BR, Lee SJ, et al. Peripheral-blood stem cells versus bone marrow from unrelated donors. N Engl J Med 2012;367:1487–96.
31. Vasconcelos ZF, Santos BM, Costa ES, et al. T-lymphocyte function from peripheral blood stem-cell donors is inhibited by activated granulocytes. Cytotherapy 2003;5:336–45.
32. Offner F, Schoch G, Fisher LD, et al. Mortality hazard functions as related to neutropenia at different times after marrow transplantation. Blood 1998;88:4058–62.
33. Rowley SD, Zuehlsdorf M, Braine HG, et al. CFU-GM content of bone marrow graft correlates with time to hematologic reconstitution following autologous bone marrow transplantation with 4-hydroperoxycyclophosphamide purged bone marrow. Blood 1987;70:271–5.
34. Ringden O, Barrett AJ, Zhang MJ, et al. Decreased treatment failure in recipients of HLA-identical bone marrow or peripheral blood stem cell transplants with high CD34 cell doses. Br J Haematol 2003;121:874–85.
35. Heimfeld S. HLA identical stem cell transplant: is there an optimal CD34 cell dose? Bone Marrow Transplant 2003;31:839–45.
36. Socie G, Clift RA, Blaise D, et al. Busulfan plus cyclophosphamide compared with total-body irradiation plus cyclophosphamide before marrow transplantation for myeloid leukemia: long-term follow-up of 4 randomized studies. Blood 2001;98:3569–74.
37. Storb R, Yu C, Wagner JL, et al. Stable mixed hematopoietic chimerism in DLA-identical littermate dogs given sublethal total body irradiation before and pharmacological immunosuppression after marrow transplantation. Blood 1997;89:3048–54.
38. McSweeney PA, Niederwieser D, Shizuru JA, et al. Hematopoietic cell transplantation in older patients with hematologic malignancies: replacing high-dose cytotoxic therapy with graft-versus-tumor effects. Blood 2001;97:3390–400.
39. Oliansky DM, Gordon LI, King J, et al. The role of cytotoxic therapy with hematopoietic stem cell transplantation in the treatment of follicular lymphoma: an evidence-based review. Biol Blood Marrow Transplant 2010;16:443–68.
40. Oliansky DM, Antin JH, Bennett JM, et al. The role of cytotoxic therapy with hematopoietic stem cell transplantation in the therapy of myelodysplastic syndromes: an evidence-based review. Biol Blood Marrow Transplant 2009;15:137–72.
41. Gisselbrecht C, Glass B, Mounier N, et al. Salvage regimens with autologous transplantation for relapsed large B-cell lymphoma in the rituximab era. J Clin Oncol 2010;28:4184–90.
42. Sweetenham JW, Pearce R, Taghipour G, et al. Adult Burkitt's and Burkitt-like non-Hodgkin's lymphoma – outcome for patients treated with high-dose therapy and autologous stem-cell transplantation in first remission or at relapse: results from the European Group for Blood and Marrow Transplantation. J Clin Oncol 1996;14:2465–72.

MULTIPLE-CHOICE QUESTIONS

1. What is the objective of using a conditioning regimen for hematopoietic stem cell transplantation (HSCT)?
   A. To reduce the bulk of malignant cells
   B. To decrease the risk of graft rejection
   C. To improve the engraftment of donor hematopoietic stem cells (HSCs)
   D. All the above

2. What is the main mechanism of graft-versus-leukemia (GvL) effect using donor lymphocyte infusion?
   A. Release of cytokine mediators generated by the graft-versus-host disease (GvHD) reaction initiated by donor lymphocyte infusion (DLI)
   B. Provide CD4 and CD8T cells that recognize target antigens on tumor cells
   C. Enhance inhibitory receptors (killer cell immunoglobulin-like receptors [(KIRs]) recognized by donor natural killer (NK) cells
   D. Increase the T-cell receptor (TCR) repertoire diversity

3. Which of the following increases the risk of relapse after an allogeneic stem cell transplantation?
   A. Human leukocyte antigen (HLA) mismatched donor
   B. $CD34^+$ cell dose less than 2×10^6/kg patient weight
   C. Presence of acute or chronic GvHD
   D. Disease burden prior to stem cell transplantation

4. Use of which strategy decreases the risk for acute and/or chronic GvHD?
   A. Unrelated donor
   B. Female donor for male recipient
   C. Total body irradiation as part of the conditioning regimen
   D. T-cell depletion of graft

84

Immunoglobulin Therapy: Replacement and Immunomodulation

Mark C. Ballow

Over 70 years ago, a cold ethanol fraction of plasma that contained an enriched fraction of gammaglobulin was used intramuscularly as passive immunotherapy for the treatment and protection of infectious pathogens and subsequently as antibody replacement therapy for patients with primary immune deficiency diseases (PIDDs). This Cohn ethanol plasma fraction of immunoglobulin G (IgG) remained the principal form of therapy until 1981 when an intravenous (IV) preparation (*i.e.,* intravenous immunoglobulin [IVIG]) became available. Subsequently, Imbach observed that thrombocytopenia resolved when patients with immune deficiency were treated with IVIG. This observation led to the use of IVIG in patients with autoimmune idiopathic thrombocytopenic purpura, and an expansion on the use of IVIG as immunomodulatory therapy in several FDA approved autoimmune disorders (Table 84.1). Most IVIG preparations are derived from plasma by the Cohn ethanol fractionation method or the Cohn-Oncley modification. IVIG products are modified to prevent the formation of IgG aggregates and to make IVIG suitable for the IV route. Excipients, such as sugars (maltose or D-sorbitol) or amino acids (glycine or L-proline), stabilize the IgG molecules from aggregation. Cold ethanol fractionation as the first step in plasma processing inactivates human immunodeficiency virus (HIV). Treatment with low pH, solvent and detergent, pasteurization or nanofiltration/depth filtration in combination, depending on the Ig product, is used as further 5–7 steps for viral inactivation and removal. Commercial lots of IVIG are derived from ≈15 000 donors (not to exceed 60 000 donors according to the U.S. Food and Drug Administration [FDA] and the Plasma Protein Therapeutics Association). Each lot must contain adequate levels of antibody to measles. These products may vary slightly among manufacturers and from lot to lot, but they are generally comparable with regard to clinical efficacy, but perhaps not tolerability. The newer Ig formulations are isoosmolar, low-sodium, and low-IgA liquid products. The characteristics of IVIG preparations available in the United States are reviewed in Table 84.2.

In this chapter, the application of IVIG as replacement therapy in patients with primary immune deficiency (PID) and the potential mechanisms of action of Ig therapy in the treatment of autoimmune and inflammatory diseases are reviewed. For more information on the use of Ig therapy in specific diseases, the reader is referred to an evidence-based review of the topic elsewhere.[1,2]

KEY CONCEPTS 1

Properties of Intravenous Immunoglobulin (IVIG) and Subcutaneous Immunoglobulin (SCIG)

- Plasma fractionation (first step) by cold ethanol / Cohn-Oncley modification (fraction II)
- >98% IgG; >90% monomeric IgG
- Traces of other immunoglobulins (*i.e.,* IgA and IgM) and serum proteins
- Addition of sugar, and amino acids, stabilizes IgG from aggregation
- Multiple viral inactivation/removal steps
  - Cold ethanol fraction
  - Chromatography
  - Solvent/detergent treatment
  - Caprylate fractionation
  - Nanofiltration
  - Depth filtration
  - Pasteurization
- Intact Fc receptor important for biological function
  - Opsonization and phagocytosis
  - Complement activation
  - Antibody dependent cytotoxicity
- Normal half-life comparable to serum IgG
- Normal proportion of IgG subclasses
- Broad spectrum of antibodies to bacterial and viral agents

REPLACEMENT THERAPY WITH IVIG

Several early trials in the 1980s have been conducted to examine the efficacy of IVIG with intramuscular (IM) gammaglobulin, and a comparison of low-dose with high-dose IVIG therapy. Quartier et al.[3] performed a retrospective study of the clinical features and outcomes of 31 patients with X-linked agammaglobulinemia receiving replacement IVIG therapy between 1982 and 1997. Although early treatment with IVIG and achieving a trough serum IgG level of >500 mg/dL was effective in preventing severe acute bacterial infections, these levels did not prevent pulmonary disease, sinusitis, or meningoencephalitis. The authors suggested that more intensive therapy to maintain a higher serum IgG trough level (*e.g.,* >800 mg/dL) may improve the pulmonary outcome.

Several studies have shown that pulmonary abnormalities are the most important factors associated with morbidity and mortality in patients with PIDs. The number of infections, days on antibiotic therapy, days missed from school or at work, and hospitalized days may not be sufficient indicators of adequate treatment. Therefore the improvement or maintenance of pulmonary function is an important measure of the success of therapy. Orange et al.[4] examined the impact of serum IgG trough levels on pneumonia incidence in patients with PIDDs receiving replacement IVIG therapy in a meta-analysis of published clinical

TABLE 84.1 FDA-Approved Indications for IVIG

- Primary humoral immunodeficiency disease—as replacement therapy
- Immune thrombocytopenic purpura—to prevent severe bleeding
- Children with HIV/AIDS and recurrent infections—to prevent serious bacterial infections
- B-cell chronic lymphocytic leukemia with recurrent infections and humoral immune deficiency—to prevent bacterial infection
- Kawasaki disease—to prevent coronary artery aneurysms
- Bone marrow transplantation—to decrease risk of infection, interstitial pneumonia, and graft-versus-host disease in the first 100 days after transplantation
- Chronic inflammatory demyelinating polyneuropathy
- Multifocal motor neuropathy

FDA, U.S. Food and Drug Administration; *IVIG,* intravenous immunoglobulin; *HIV,* human immunodeficiency virus; *AIDS,* acquired immunodeficiency syndrome.

KEY CONCEPTS 2

Treatment of Patients With Primary Immune Deficiency With Intravenous Immunoglobulin

- Starting dose—400–600 mg/kg every 4 weeks.
- Maintain trough level of >500 mg/dL or a serum immunoglobulin G (IgG) level 450 mg/dL over baseline.
  - Recent studies indicate the most patients may require a trough or steady-state level of 750–850 mg/dL.
  - The most important dose of intravenous immunoglobulin (IVIG) is the dose that keeps the patient free of infections (*i.e.*, biological trough level).
- Higher trough levels (>800 mg/dL) may be necessary to prevent chronic pulmonary changes and enteroviral meningoencephalitis.
- Depending on the IgG catabolism and/or control over infections, decrease infusion interval to every 3 weeks.
- Subcutaneous immunoglobulin (SCIG) offers an advantage in achieving higher steady-state plasma IgG levels.
- It may take ≥3 months to achieve a steady state after a change in dose.
- Monitor serum blood urea nitrogen (BUN) and creatinine and perform liver function tests every 6–12 months.
- Keep a log of dose, manufacturer, lot number, and reactions for each infusion.
- For patients with rate-related adverse side effects. Consider pretreatment with:
  - Acetaminophen
  - Diphenhydramine
  - Nonsteroidal antiinflammatory drugs (NSAIDs)
  - Corticosteroids
- Consider alternate administration of SCIG in patients with frequent systemic adverse reactions.

studies. The results of their analysis showed that trough IgG levels increased 121 mg/dL for every 100 mg/kg increase in dose, which resulted in a decrease in pneumonia incidence by 27%. There was a strong correlation between increasing trough IgG levels and a decrease in pneumonia. Bonagura et al.[5] discussed the value of "biological" trough levels for the serum IgG related to the patient's clinical course (*i.e.,* frequency of infections) instead of arbitrarily choosing a trough value of 500 mg/dL. A study of 90 patients with common variable immune deficiency (CVID) with follow-up of 22 years showed that the dose of IVIG needed to keep a patient free of infection varied among patients, emphasizing that the goal of replacement therapy should be improvement of clinical outcome, not a specific IgG trough level.[6]

The overall consensus among clinical immunologists is that an IVIG dose of 400–600 mg/kg/month is a good *starting point*. (see Key Concepts 2) Catabolism of infused Ig varies among individuals. The rate of elimination of IgG may be higher during a period of active infection. For persons with a higher catabolism of infused IgG or more frequent infections, infusions every 3 weeks with smaller doses may be more efficacious. In the final analysis, trough levels are only a general guide, and the clinical well-being of the patient is a more important parameter. Clearly, with higher trough levels, the incidence of pneumonia and comorbid conditions (*e.g.,* bronchiectasis and meningoencephalitis) is reduced.[3-6]

For replacement therapy in patients with PID all brands of IVIG are probably equivalent, although there are differences in viral inactivation processes (see Table 84.2). The choice of brands may be dependent on the hospital, payor, or home care formulary; the distributor availability and cost; and, most importantly, patient tolerability. The dose, manufacturer, and lot number should be recorded for each infusion to facilitate review procedures for adverse events or other consequences. It is crucial to record all side effects that occur during the infusion. It is also recommended that liver and renal functions be monitored periodically, approximately once or twice a year. Changes in weight or growth in children or changes in clinical response or adverse events may necessitate a more frequent laboratory evaluation. More frequent serum IgG levels may be necessary for patients on home therapy, especially for those receiving subcutaneous (SC) products to evaluate compliance.

In patients with active infection, especially patients with CVID, the initial (first) dose can be halved (*i.e.,* 200–300 mg/kg) and the dose repeated 2 weeks later to achieve a full dose. To minimize cost and inconvenience, self-administration and home treatment have been used successfully. For home therapy, patients need to be selected carefully. Infusions should be done only in the presence of a responsible adult who is ready to provide assistance. Home treatment has been reported to be as effective as hospital treatment in terms of the frequency of infections, days missed from school or work, antibiotics used, and Ig level achieved. Patients receiving home treatment should be seen regularly (*i.e.,* approximately every 6 months) to monitor clinical status, liver function, and serum IgG level.

The first description on the use of the SC route for Ig replacement therapy was published in 1980. It was reported to be safe, well tolerated, and effective in achieving adequate serum IgG levels. In a multicenter study of 165 patients with hypogammaglobulinemia receiving subcutaneous infusions (27 030 at home), a significant reduction in adverse systemic reactions with SC administration, compared with the IM or IV administration, was observed. Anaphylactoid reactions did not occur. Several studies have shown enhancements in quality-of-life measurements.[7,8] Local tissue reactions occur with subcutaneous immunoglobulin (SCIG) infusions, including swelling, soreness, redness, induration, local heat, itching, and bruising. There are advantages and disadvantages for each of the routes of administration of replacement Ig therapy. (Clinical Pearls)

SCIG is a suitable alternative to IVIG and may present certain opportunities for optimizing care for patients with PIDD at home.[9] SCIG is an excellent alternative to IVIG for those patients unable to tolerate the IV route of Ig replacement due to systemic side effects with IVIG administration. SCIG also leads to higher serum IgG steady-state levels[10] and may eliminate the wear-off effects of IVIG replacement therapy.[11,12]

TABLE 84.2 Characteristics of Immunoglobulin (Ig) Preparations

Brand (Manufacturer)	Manufacturing Process	pH	Stabilizer/ Sodium	Parenteral Form and Final Concentrations	IgA Content (µg/mL)	Antiviral Steps
Intravenous IgG						
Gammagard S/D (Shire)	Cohn-Oncley cold ethanol fractionation, followed by ultracentrafiltration and ion exchange chromatography	6.4–7.2	Albumin, glycine, glucose; sodium 8.5 mg/mL / 17 mg/mL	Lyophilized powder 5%, 10%	<2.2 µg/mL (5% solution)	PEG precipitation, solvent detergent treated, low pH, depth filtration
Gammagard Liquid (Shire)	Cohn-Oncley cold ethanol fractionation, followed by ultracentrafiltration and ion exchange chromatography	4.6–5.1	Glycine; trace sodium	10% liquid solution	37 µg/mL	S/D treatment, nanofiltration, low pH incubation
Gamunex -C (Instituto Grifols, SA)	Cohn-Oncley cold ethanol fractionation, diafiltration and depth filtration; caprylate/ anion-exchange chromatography	4.0–4.5	Glycine; trace sodium	10% liquid solution	46 µg/mL	Caprylate precipitation, depth filtration, low pH incubation
Privigen (CSL Behring AG)	Cold alcohol fractionation, octanoic fractionation, chromatography and depth filtration	4.6–5.0	L-Proline trace sodium	10% solution	≤25 µg/mL	Depth filtration, pH4 incubation, nanofiltration, TSE reduction steps
Gammaked (Grifols -Distributed by: Kedrion Biopharma)	Cohn-Oncley cold ethanol fractionation, diafiltration and depth filtration; caprylate/ anion-exchange chromatography	4.6–4.5	Glycine	10% liquid solution	46 µg/mL	Caprylate precipitation, depth filtration, low pH incubation
Flebogamma DF (Instituto Grifols, SA)	Cold alcohol fractionation, ion exchange chromatography, polyethylene glycol precipitation,	5.0–6.0	D-Sorbitol	5% and 10% liquid solution	<50 µg/mL for the 5%	PEG precipitation, pasteurization, S/D treatment, nanofiltration, pH4 treatment
Octagam (Octapharma)	Cohn-Oncley cold ethanol fractionation, chromatography	5.1–6.0	Maltose	5% and 10% solution	<100 µg/mL	Solvent/detergent, pH4 incubation, ultrafiltration
Gammaplex (Bio Products Laboratory)	Kistler & Nitschmann fractionation, chromatography, DEAE Sephadex chromatography, CM-Sepharose chromatography	4.8–5.1	D-sorbitol, glycine and polysorbate 80 Sodium 30–50 mmol/L	5% solution	<10 µg/mL	S/D treatment, nanofiltration, low pH incubation
Bivigam (Biotest Pharmaceuticals Corporation)	Cohn-Oncley cold ethanol fractionation, ion-exchange chromatography	4.0–4.6	Glycine 0.1–0.14 M sodium	10%	≥ 200 µg/mL	S/D, nanofiltration, fraction precipitation
Carimmune NF (CSL Behring AG)	Kistler & Nitschmann fractionation, Cold alcohol fractionation, Chromatography, nanofiltration	6.4–6.8	Sucrose <20 mg sodium/g protein	Lyophilized preparation—3%, 6%, 9%, or 12%	720 µg/mL	pH4, trace pepsin, depth filtration, nanofiltration, TSE reduction steps
Subcutaneous IgG						
Hizentra (CSL Behring AG)	Cold alcohol fractionation, octanoic acid fractionation, anion exchange chromatography	4.6–5.2	L-Proline Trace sodium	20%	≤50 µg/mL	Depth filtration, pH4 incubation, depth filtration, nanofiltration, TSE reduction steps
Gamunex-C (Instituto Grifols, SA)	Cold ethanol fractionation, diafiltration and depth filtration; caprylate/anion exchange chromatography	4.0–4.5	Glycine; trace sodium	10%	46 µg/mL	Caprylate precipitation, depth filtration, low pH incubation
Gammaked (Grifols -Distributed by: Kedrion Biopharma)	Cold ethanol fractionation, Depth filtration, caprylate/ anion-exchange chromatography	4.6–4.5	Glycine	10%	46 µg/mL	Caprylate precipitation, depth filtration, low pH incubation
Gammagard Liquid (Shire)	Cohn-Oncley cold ethanol fractionation, followed by ultracentrafiltration and ion exchange chromatography	4.6–5.1	Glycine; trace sodium	10% solution	37 µg/mL	S/D treatment, nanofiltration, low pH incubation
HyQvia (Shire)	Cohn-Oncley cold ethanol fractionation, followed by ultracentrafiltration and ion exchange chromatography	4.6–5.1	Glycine	10% plus human recombinant hyaluronidase	37 µg/mL	S/D treatment, nanofiltration, low pH incubation
CUVITRU (Shire)	Cohn-Oncley cold ethanol fractionation, followed by ultracentrafiltration and ion exchange chromatography	4.6–5.1	Glycine	20%	80 µg/mL	S/D treatment, ultra/ diafiltration, low pH incubation

Contents of table summarized from manufacturers' literature.

PEG, polyethylene glycol; *S/D*, solvent detergent; *TSE*, transmissible spongiform encephalopathy.

CLINICAL PEARLS 1

Comparison of Routes of Administration of Immunoglobulin Therapy in Patients With Primary Immune Deficiency

Intravenous (IV) Route

Advantages

- Achieve rapid plasma levels
- Can use this route in patients with bleeding disorders
- 3–4 week intervals

Disadvantages

- Need IV access
- Interrupt patient's schedule for 3–5-hour period
- Often needs to come to a hospital or infusion center
- System side effects may be more frequent in some patients

Subcutaneous (SC) Route

Advantages

- IV access not needed for those patients with poor venous access
- Eliminate trough levels
  - Achieve a stable serum level of immunoglobulin G (IgG)
  - May eliminate 3rd–4th week fatigue prior to next infusion (wear off)
- Less systemic adverse effects than IV route
- More flexibility for patient's (parent's) schedule
  - Distance from infusion center or hospital
  - Young adults going to college

Disadvantages

- Minor local reactions at the site of infusion
- Patient reliability
- Need for a pump

SCIG products are 10% or 20% formulations, the former being similar in composition to IVIG products. The calculated dose for SC administration is generally 100–150 mg/kg weekly. Depending on the weight or body mass of the patient and the concentration of the SCIG (*i.e.,* 10% or 20%), infusions may have to be given more frequently than every 7 days. Pharmacokinetics studies in clinical trials have suggested that upward adjustments in the dose of 37% of the IVIG dose may be needed to achieve comparable bioavailability, defined as the area under the serum concentration curve (AUC).[13] This adjustment in dosing between IVIG and SCIG has not been mandated by European regulatory agencies, and in the United States some clinical immunologists select dosing based on optimization of prevention of infections, as noted above for the IVIG dosing.[5,6] Each subcutaneous infusion requires a small portable syringe driver–type pump together with a 10–20-mL syringe and an infusion set with a specialized SC 25–27-gauge needle. The length of the needle may have to be adjusted for the thickness of the subcutaneous tissue of each patient. Before infusion, the line needs to be checked to ensure that there is no blood return. Infusions can be given weekly at multiple sites or more often (*e.g.,* daily), if needed, to maintain adequate serum IgG levels. The 20% SCIG product can be given every 2 weeks. Infusion sites are usually on the abdominal wall and lateral thigh. In adults, 20–35 mL can be infused into a single site, depending on the amount of subcutaneous tissue. A general guideline for infusion rate is 15–35 mL/hr per site depending on the weight and subcutaneous tissue of the patient. The SCIG schedule should be started 1 week after the last dose of IVIG, or in a new patient loaded with 4 or 6 doses of SCIG. Before home treatment, patients need to be instructed on the correct technique, made to apply it under close supervision, and educated about the recognition of possible side effects. SCIG infusion is safer, better tolerated, and preferred by some patients. Several studies in Europe and Canada have shown that health care costs of SCIG therapy are lower compared with IVIG.[14-16] It should be considered as an alternative, especially in those patients with systemic adverse reactions from IV administration. Recently, an alternative method for the administration of Ig has been approved and has been referred to as *facilitated SC infusion.*[17,18] In this approach, human recombinant hyaluronidase is used to enhance the volume of Ig (10% liquid Ig) that can be infused in the SC space, allowing for monthly doses of Ig to be administered in one or two sites. The hyaluronidase depolymerizes the hyaluronan temporarily allowing the Ig greater access to the lymphatics of the SC space, facilitating the absorption of Ig. This method of administration of 10% Ig has been shown in pharmacokinetic studies to result in bioavailability of 93% of the IVIG dose and thus does not require an upward adjustment factor in calculating the dose of Ig replacement. More details about SC infusions can be found in several reviews.[11,19,20]

ADVERSE EVENTS ASSOCIATED WITH IVIG THERAPY

Rate-Related Adverse Events

Typical rate-related adverse reactions with IVIG include tachycardia, dyspnea, chest tightness, back pain, arthralgia, myalgia, hypertension or hypotension, headache, pruritus, rash, and low-grade fever. Mild to moderate reactions occur in 5–15% of infusions; severe reactions occur in <1% of patients. Of course, in patients with autoimmune disorders, reaction rates are higher with higher doses. Patients with more profound immunodeficiency or patients with active infections also tend to have more severe reactions. Other factors that contribute to adverse reactions include change of IVIG products, concomitant infections, higher concentrations or lyophilized products, and rapid infusion rates (reviewed in the report by Stiehm[21]). The cause of the reactions is thought to be related to the anticomplementary activity of IgG aggregates in the IVIG in which immune complexes form between infused antibodies and antigens of infectious agents in the patient. The other possible mechanism is that the formation of oligomeric or polymeric IgG complexes interacts with Fc receptors and triggers the release of inflammatory mediators. These rate-related reactions occur less frequently with the newer IVIG products that are liquid and isoosmolar. Headaches are the most frequent symptom associated with IVIG infusions occurring in 5–20% of infusions and one-third of patients. Slowing the infusion rate or discontinuing therapy until symptoms subside may diminish the reaction. Pretreatment with nonsteroidal antiinflammatory drugs (NSAIDs), acetaminophen (15 mg/kg/dose), diphenhydramine (1 mg/kg/dose), and/or hydrocortisone (6 mg/kg/dose, maximum 100 mg) 1 hour before the infusion may prevent adverse reactions. Oral hydration prior to the infusion is often helpful. Switching products may also lead to adverse events in 15–18% of patients and should be discouraged.

Central Nervous System–Related Adverse Events

Aseptic meningitis has been reported as one of the complications of IVIG, especially with large doses and rapid infusions and in the treatment of patients with autoimmune disease.[21] Interestingly, this rarely occurs in subjects with immunodeficiencies. Symptoms,

CLINICAL PEARLS 2

Adverse Events Associated With Immunoglobulin Therapy

Rate-Related

- Infusion site erythema, swelling, pain, itching
- Headache
- Myalgia, back pain, arthralgia
- Malaise, fatigue
- Chills, Fever
- Pruritus
- Rash—urticaria
- Nausea, vomiting
- Tachycardia
- Dyspnea, chest pain, or tightness
- Hypotension/hypertension

Central Nervous System

- Severe headaches
- Trigger migraine headaches
- Aseptic meningitis*

Renal

- Azotemia
- Renal failure

Thromboembolic Events*

- Thrombosis/cerebral infarction
- Myocardial infarction
- Pulmonary embolism

Anaphylaxis From Anti-Immunoglobulin E (IgE) Antibodies to IgA

Other (Isolated Reports)

- Cardiac rhythm abnormalities
- Coagulopathy
- Serum sickness
- Hemolysis—alloantibodies to blood type A/B
- Cryoglobulinemia
- Neutropenia
- Alopecia
- Uveitis
- Noninfectious hepatitis
- Progressive neurodegeneration

*See text for predisposing risk factors.

including headache, stiff neck, and photophobia, usually develop within 24 hours after completion of the infusion and may last 3–5 days. Spinal fluid pleocytosis occurs in most patients. Long-term complications are minimal. The etiology of aseptic meningitis is unclear, but migraine has been reported as a risk factor and may be associated with recurrence despite the use of different IVIG preparations and slower rates of infusion. (Clinical Pearls 2)

Renal Adverse Events

Acute renal failure is a rare but significant complication of IVIG treatment and has been associated in the past with products containing sucrose. Histopathological findings of acute tubular necrosis, vacuolar degeneration, and osmotic nephrosis were suggestive of osmotic injury to the proximal renal tubules. Most patients (95%) had received large doses for treatment of autoimmune diseases. The majority of the cases were treated successfully with conservative therapy, but deaths were reported in 17 patients who had serious underlying conditions. Most cases of this adverse event were associated with IVIG products containing sucrose as a stabilizer. Risk factors for this adverse reaction include preexisting renal insufficiency, diabetes mellitus, dehydration, patient age >65 years, sepsis, paraproteinemia, and concomitant use of nephrotoxic agents. Newer IVIG products are using alternative stabilizers (*e.g.*, amino acids) instead of sucrose.

CLINICAL PEARLS 3

Risk Factors for Adverse Events

- Infusion issues
  - Prior history of an infusion reaction with an immunoglobulin (Ig) product
  - First infusion in a patient with active infection or inflammation
  - Changing immunoglobulin products
  - Rapid infusion and/or large dose
- Patient factors
  - Preexisting renal insufficiency
  - Prior history of thrombotic event
  - Autoimmune disorder
  - Diabetes mellitus
  - Age—older age
  - Hyperlipidemia or elevated cholesterol
  - Dehydration with volume depletion
  - Hypercoagulable state
  - Indwelling catheters
  - Paraproteinemia or other causes of hyperviscosity
  - Cardiac or peripheral vascular disorders
  - Estrogen use
  - Smoking

Thromboembolic Events

This adverse effect was observed mainly in patients receiving large doses of IVIG for autoimmune diseases. Patients with elevated serum viscosity (*e.g.*, cryoglobulinemia, hypergammaglobulinemia, and hypercholesterolemia) are at risk for developing a critical increase in serum viscosity with IVIG, especially high doses that predispose them to thromboembolic events, such as myocardial infarction, stroke, deep vein thrombosis, or pulmonary embolism. (Clinical Pearls 3) Recently, a contaminating procoagulant factor (*e.g.*, factor XIa) has been implicated in these thromboembolic events.[22] Patients at risk are older (>65 years), on multiple drugs, and have comorbidities, such as diabetes, hypertension, and so on. Patients at risk should be well hydrated, IVIG should be administered at lower rates, and products with low sodium and an osmolality in the physiological range should be selected. Ig manufacturers have taken steps to remove or reduce this procoagulant activity from their products.[23]

Transfusion Reaction Caused by Antibodies Against IgA

True anaphylaxis is rare in patients with selective IgA deficiency and CVID who develop IgE antibodies to IgA after treatment with Ig; this adverse event appears to be much less frequent than originally thought. Patients with CVID have IgG antibodies (10–22% in various studies) to IgA, but there is no correlation to the presence of these antibodies and adverse reactions.[24] Patients with anti-IgA antibodies who have had reactions to IVIG have tolerated SCIG.[25]

Other Adverse Reactions

A number of other adverse reactions have been reported in association with IVIG infusions. (Clinical Pearls 2) These side effects are usually less common and are discussed in more detail elsewhere.[21]

Summary: Ig Replacement in Treatment of Immune Deficiency

Ig replacement is the mainstay of treatment for primary humoral immune deficiency. The goal of treatment is to provide a broad spectrum of antibodies to prevent infections and chronic long-term complications. The usual dose is 400–600 mg/kg/month, but this may vary among individuals, and some patients may require higher doses. A serum trough level above 500 mg/dL has been shown to be effective in the prevention of serious bacterial infections. However, recent studies have suggested that even higher doses (*e.g.,* a “biological” trough level and achieving trough levels of IgG in the range of 750–900 mg/dL) may be desirable.[6] Ig can be given intramuscularly, intravenously, or subcutaneously. SC administration has been proven to be safe and is a good alternative in some patients, especially those experiencing side effects of administration by the IV route. Generally, IVIG replacement therapy is considered safe in the majority of patients. Side effects are usually mild and treatable with premedication. Good manufacturing practices, improved screening of plasma donors, testing of the source plasma, and additional viral inactivation and removal steps have made IVIG a better and safer plasma-derived product. To optimize the care of patients with PIDs, there are eight guiding principles for the use of Ig replacement therapy. (see Therapeutic Principles)

MECHANISMS OF ACTION OF IG THERAPY IN AUTOIMMUNE AND INFLAMMATORY DISEASES

Although we still do not understand *all* the mechanisms by which IVIG has immunomodulatory effects, knowledge of the actions of IVIG in these diseases will allow for definition of appropriate indications and schedules of administration of IVIG and the design of new-generation IVIG products that are better able to target the immune perturbations in autoimmune and inflammatory processes. (Key Concepts 3) Furthermore, additional multicenter placebo-controlled clinical trials are needed to confirm clinical efficacy. A more detailed discussion of the evidence-based treatment of autoimmune and inflammatory disorders can be found elsewhere.[1,2] This chapter will not review the indications for IVIG therapy; the reader is directed to other works, including comprehensive reviews and practice-based guidelines, for more details about the various indications for IVIG therapy in autoimmune and inflammatory disorders. This section will address the possible mechanisms of actions of Ig therapy; more than one mechanism(s) may be important in the immune modulation of these autoimmune disorders.

KEY CONCEPTS 3

Mechanisms of Action of Intravenous Immunoglobulin (IVIG) in Autoimmune and Inflammatory Diseases

- Blockade of Fc receptors on macrophages of the reticuloendothelial system of liver and spleen
- Restoration of idiotypic/antiidiotypic network
- Suppression or neutralization of cytokines by specific antibodies in the IVIG
- Block binding of adhesion molecules on leukocytes to vascular endothelium
- Inhibition of complement uptake on target tissues
- Neutralization of microbial toxins
- Blockade of Fas ligand mediated apoptosis by anti-Fas antibodies in the IVIG
- Induction of apoptosis with anti-Fas antibodies at high concentrations of IVIG
- Neutrophil apoptosis by anti-Siglec-9 antibodies in IVIG
- Saturation of FcRn receptors to enhance the clearance of autoantibodies
- Induction of inhibitory FcγInd receptors on effector macrophages
- Neutralization of growth factors for B cells (*e.g.,* B cell–activating factor ([BAFF])
- Inhibition of T-cell proliferative responses
- Expansion and/or activation of a population of regulatory T cells (Tregs)
- Enhancement of the cyclooxygenase 2 (COX-2) pathway by increasing prostaglandin E2 (PGE_2) from dendritic cells (DCs)
- Downregulation of the T-helper 17 (Th17) pathway
- Modulation of DC function through C-type lectin receptors (DCLRs) DC-specific intercellular adhesion molecule-grabbing nonintegrin (DC-SIGN; spleen) and DCLR (lung and lymph nodes)

Blockade of Fc Receptors of the Reticuloendothelial System

Idiopathic thrombocytopenic purpura (ITP) results from accelerated platelet destruction attributable to an immunological process resulting in bleeding that may be life threatening. Studies in ITP have shown that IVIG leads to the rapid reversal of the thrombocytopenia. Autoantibody-opsonized platelets are destroyed in the spleen and the liver by FcγR-mediated phagocytic clearance. (Key Concepts 4)

Interactions of Idiotype and Antiidiotype as Immune Modulation

Idiotype–antiidiotype interactions between antiplatelet glycoprotein IIb (GPIIb)/IIIa autoantibodies and IVIG may be another mechanism by which IVIG could affect autoantibody production and effector function in ITP and may be playing a role in the long-term immune effects in ITP. IVIG contains antiidiotypic antibodies to anti-DNA, anti–factor VIII, antineutrophil cytoplasmic antibody (ANCA) autoantibodies, antithyroid autoantibodies, and others. A proposed mechanism of action of IVIG in ANCA-positive vasculitis is binding or neutralization of the ANCA autoantibodies by antiidiotypic antibodies in IVIG. Studies have shown that IVIG contains antibodies with idiotypic specificities that can bind and neutralize potentially pathogenic autoantibodies in autoimmune neurological diseases, such as Guillain-Barré syndrome (GBS), chronic inflammatory demyelinating polyneuropathy (CIDP), and myasthenia gravis (MG).[26] Antiidiotypic antibodies in IVIG directed against idiotopes on the anti-GM_1 Ig molecule block the binding of the anti-GM_1 antibodies to its target antigen. Support for a similar mechanism in MG comes from the fact that IgG or $F(ab')_2$ fragments in IVIG preparations are capable of binding to anti–acetylcholine receptor antibodies *in vitro*. Additionally, using *in vitro* nerve-muscle preparations, Buchwald et al.[27] showed that the $F(ab')_2$ portion of IVIG neutralized the “blocking” effect of serum from patients with acute GBS. (Key Concepts 5)

These antiidiotypic antibodies may also act in concert with their effects on the FcγRIIB receptor on B cells to produce a negative "off signal" to the B cells synthesizing these autoimmune antibodies. Thus the antiidiotypic antibodies in IVIG may be beneficial by restoring the idiotypic control network in these autoimmune diseases.

The Role of the FcRn Receptor on Immune Modulation

The FcRn receptor (neonatal Fc receptor) was identified as the receptor responsible for protecting IgG from catabolism in the endocytic vesicles of the endosome, and this explains the relatively long half-life of this plasma protein. IVIG may accelerate the catabolism of IgG autoantibodies by saturating these protective receptors in direct proportion to the relative concentration of exogenous plasma levels of IgG from IVIG. In mouse models of bullous pemphigoid and arthritis, IVIG treatment resulted in a reduction in pathogenic antibodies to levels below the disease-causing threshold, and this effect was attenuated in FcRn-deficient mice.[28] Hansen and Balthasar[29] showed in a rat model of immune thrombocytopenia that IVIG enhanced the clearance of antiplatelet antibodies in a dose-dependent manner by saturation of the FcRn receptor. These same investigators[30] estimated that approximately 50% of the overall effect of IVIG in ITP may be attributed to IVIG-mediated acceleration of the elimination of antiplatelet antibodies by the FcRn saturation mechanism and showed that in FcRn knock-out mice, IVIG failed to increase the clearance of antiplatelet antibodies. However, very little data, if any, are available in human autoimmune disease to support this mechanism.

KEY CONCEPTS 4

Proposed Mechanisms of Action of intravenous immunoglobulin (IVIG) in Autoimmune Idiopathic Thrombocytopenic Purpura

- Fc receptor blockade of reticuloendothelial system
- Fcγ receptor downregulation
- Idiotype–antiidiotype interaction between antiplatelet GPIIb/IIa autoantibodies and the antiidiotypic antibodies in IVIG
- Activation of inhibitory receptor FcγRIIB
- Saturation of FcRn receptor to accelerate the catabolism of antiplatelet autoantibodies

Modulation of Immunoregulatory Function Through the Fc Receptor

The FcγRIIB receptor provides an inhibitory signal to cells through a pathway mediated by an immunoregulatory tyrosine-based inhibition motif (ITIM). Studies have shown that IVIG stimulates these inhibitory FcγRIIB receptors found on a variety of cell types, including macrophages, B cells, and a subpopulation of T cells. Samuelsson et al.,[31] using a murine model of ITP, showed that the administration of IVIG prevented platelet destruction by a pathogenic monoclonal autoantibody. Protection was associated with the induction of the expression of FcγRIIB receptors on splenic macrophages. This inhibitory FcγRIIB receptor was required for the protection of the animals against the monoclonal autoantibody, since disruption of the receptor by either genetic deletion or a blocking monoclonal antibody (mAb) reversed the therapeutic effects of IVIG. Kaneko et al.[32] showed that that the inhibitory properties of IVIG were linked to the sialylation of the glycan component of the Fc fragment. The important glycan moiety in the IgG molecule is located at the asparagine (Asn^{297}) site in the second domain of the constant region. Using a K/BxN serum-induced arthritis model in mice, Kaneko et al.[32] showed that IVIG at 1 g/kg inhibited the inflammatory arthritic process. Deglycosylated or neuraminidase-treated IVIGs were unable to inhibit this inflammation. IVIG enriched for the sialylated glycan moiety had comparable inhibitory effects on the inflammatory process at only one-10th of the dosage used with intact IVIG. This inhibitory activity was dependent on FcγRIIB expression on effector macrophages. Anthony et al.[33] engineered a recombinant/sialylated human IgG1 Fc protein that had 35-fold enhanced immune-modulating activity compared with native IVIG. Anthony et al.[34] performed studies to examine the mechanism by which the sialylated Fc fragment could mediate its antiinflammatory activity and to identify the target cell that initiates this antiinflammatory pathway. A splenic marginal zone macrophage expressing the C-type lectin receptor (*i.e.*, SIGN-R1) was required for the antiinflammatory activity of IVIG in concert with its ability to bind to sialylated Fc domains. These authors proposed that the interaction between sialylated IgG Fc and SIGN-R1 leads to the upregulation of the inhibitory FcγRIIB receptor on effector cells. They suggested that DC-SIGN (dendritic cell–specific intercellular adhesion molecule-grabbing nonintegrin), the human orthologue of SIGN-R1, has a comparable role in the antiinflammatory effects of IgG Fc fragment on human macrophages and dendritic cells (DCs). Anthony et al.[35] presented data in their murine model that this immune-modulating pathway may be mediated by the production of interleukin-33 (IL-33) by DCs, which, in turn, produce IL-4 from basophils, leading to the increased expression of the FcγRIIB receptor on effector macrophages. Of note, observations in patients with chronic inflammatory demyelinating polyneuropathy (CIDP) treated with high-dose IVIG demonstrated upregulation of the FcγRIIB receptor on peripheral blood monocytes.[36] Others, however, have questioned whether this pathway of immune modulation is important in humans.[37-39] Nevertheless, these exciting studies define, at least in mice, an important mechanism by which IVIG modulates immune processes mediated through sialylated Fc on the IgG molecule and the receptors on DCs and effector macrophages (*i.e.*, SIGN-R1 and FcγRIIB) involved in this antiinflammatory process.

KEY CONCEPTS 5

Proposed Mechanisms of Action of IVIG in Neuromuscular Diseases

- Modulation of proinflammatory cytokines, *e.g.*, TNF-GR α, IL-1GR β in GBS
- Reduction in muscle levels of ICAM-1, TGF-GR β, and TGF-GR β mRNA in dermatomyositis
- Neutralization by antiidiotypic antibodies in IVIG of potentially pathogenic autoantibodies, e.g., anti-Gm1 antibodies in GBS and CIDP, and anti-AchR in MG
- IVIG saturation of FcRn receptors to accelerate degradation of IgG autoantibodies
- Inhibition of complement deposition and formation of MAC on endomysial capillaries in dermatomyositis

AchR, acetylcholine nicotinic receptor; CIDP, chronic inflammatory demyelinating polyneuropathy; FcRn, neonatal Fc receptor; GBS, Guillain-Barré syndrome; MAC, membranolytic attack complex; MG, myasthenia gravis.

Neutralizing Antibody Activity in IVIG Against Bacterial Toxins

Kawasaki syndrome (KS), an acute multisystem disease of unknown etiology, primarily affects infants and young children. Although KS occurs worldwide in children of all racial groups, it is most prevalent in Japan and in children of Japanese ancestry. Although the acute illness is generally self-limiting, coronary artery abnormalities related to a generalized inflammation and immune activation of small- and medium-sized blood vessels develop in up to 25% of untreated patients. Although the etiology remains unknown, the clinical features and laboratory findings suggest an infectious or postinfectious process. The administration of high-dose IVIG, together with aspirin, is the standard of care in the treatment of KS.

KS is associated with marked activation of T cells and monocytes–macrophages. On the basis of immunological and clinical features of KS overlapping with those of bacterial toxic shock–like syndrome, studies were carried out to determine if KS is associated with exposure to a superantigen, such as a bacterial toxin. Acute KS is associated with marked immune activation and increased circulating cytokine levels. Some of these cytokines elicit proinflammatory and prothrombotic responses by inducing the expression of leukocyte adhesion molecules, which localize inflammatory cells to vascular endothelial cells. The expression of endothelial–leukocyte adhesion molecules has been demonstrated in acute KS, and its downregulation correlates with favorable response to IVIG treatment. The magnitude and persistence of proinflammatory cytokine synthesis have been reported to constitute a risk for the development of coronary artery abnormalities.

IVIG has been shown to contain high titers of specific antibodies inhibitory to the activation of T cells by staphylococcal and streptococcal superantigens (Chapter 6). These findings may account for the observation that treatment of acute KS with IVIG results in a marked reduction of macrophage and T-cell activation.[40] In this regard, the efficacy of IVIG in suppressing the immune activation associated with KS and, more importantly, its ability to prevent the development of coronary artery abnormalities in this illness may relate to the neutralizing antibody activity of IVIG against these bacterial toxins. Toxin neutralization is not likely the only beneficial effect of IVIG in KS. Blocking or neutralizing cytokines by the anticytokine antibodies in IVIG may modulate the local inflammatory responses of blood vessels in KS by modifying leukocyte adhesion after increasing the expression of cell-surface determinants on vascular endothelial cells.

IVIG may have therapeutic value in the treatment of patients with toxic shock syndrome secondary to *Staphylococcus aureus* or *Streptococcus pyogenes* exotoxins. In an open study by the Canadian Streptococcal Study Group[41], IVIG appeared to be beneficial in patients with streptococcal toxic shock syndrome. In a meta-analysis of IVIG treatment of neonatal sepsis, there was a sixfold decrease in mortality. IVIG inhibits *Staphylococcus* exotoxin–induced T-cell activation and contains antibodies against exotoxins responsible for toxic shock syndrome. Great variations in neutralizing activity against streptococcal pyrogenic exotoxins can be found in different brands and even among different lots of IVIG. However, these findings suggest that it is possible to select one IVIG preparation that contains high levels of neutralizing activity against a wide variety of Group A streptococcal superantigens, which could be used in the treatment of streptococcal toxic shock syndrome. The neutralizing capacity of IVIG against these bacterial superantigens is important because of their potential to stimulate the production of proinflammatory cytokines that lead to clinical disease. A number of *in vitro* studies have shown that IVIG can inhibit the production of, or bind to and neutralize, a number of cytokines and growth factors from various cell types.[42-44] Thus IVIG may exert its antiinflammatory effects in many different types of inflammatory diseases by interrupting or modifying a number of different steps in the inflammatory cascade, from the inhibition of effector cell function to reduction in cytokine-induced endothelial cell activation, or the “neutralization” of proinflammatory cytokines.

Modulation of Adhesion Molecules on Endothelial Cells and Antibodies in IVIG to Cell Surface Receptors

IVIG contains a number of natural autoantibodies that may have immune-modulating activities (*e.g.*, antibodies to CD4, major histocompatibility complex [MHC] class I molecules, cytokines, adhesion molecules, and Siglec molecules on leukocytes and other cell surface molecules). The “natural” antibodies in IVIG have also been shown to bind to a number of plasma and tissue proteins, including B cell–activating factor (BAFF), granulocyte macrophage–colony-stimulating factor (GM-CSF), liver antigens, and beta-amyloid peptide.[42,45] The binding of native IgG can be significantly increased by mild denaturing conditions (*e.g.*, mild pH treatment and cold ethanol precipitation used during the manufacturing of IVIG). St-Amour et al.[46] suggested that these commercial fractionation processes for IVIG could contribute to its therapeutic antiinflammatory effects.

Vassilev[47] showed that IVIG contains specific antibodies to a 10-peptide sequence, including the RGD (Arg-Gly-ASP) motif that is expressed in adhesion molecules on a variety of cell surfaces. Most integrins bind to this RGD sequence. Gill et al.[48] reported that IVIG inhibits leukocyte recruitment into inflammatory tissues by inhibiting selectin and integrin binding. In a mouse model of sickle cell vaso-occlusive crisis, IVIG was shown to inhibit neutrophil adhesion to the vascular endothelium, resulting in an increase in capillary blood flow and reversal of the vessel occlusion.[49] IVIG could modulate this cytokine-mediated endothelial cell activation by neutralizing the effects of the cytokines, inhibiting endothelial cell responses to the cytokines, or inhibiting the production of cytokines and growth factors. These mechanisms of IVIG may be playing an important role in preventing coronary artery abnormalities in patients with KS.

Toxic epidermal necrolysis (TEN) and Stevens-Johnson syndrome (SJS) are severe drug-induced skin diseases (reviewed in Letko et al.[50]). TEN results in apoptotic epidermal cell death, in which there is separation of large areas of the skin at the epidermal junction, producing the appearance of scalded skin. Keratinocyte apoptosis that precedes epidermal detachment is an early event in TEN. A number of drugs, including sulfonamides, anticonvulsants, and NSAIDs, can cause TEN and SJS. The mortality rate can be as high as 30%. Viard et al.[51] studied serum samples from patients with TEN and found that the sera of these patients had very high levels of soluble Fas ligand (sFasL). Keratinocytes normally express the death receptor Fas. The keratinocytes of patients with TEN also express very high levels of active Fas ligand. In their small pilot study, Viard et al.[51] administered IVIG (0.2–0.75 g/kg for 4 consecutive days) to 10 patients with TEN. In all 10 patients, the progression of skin disease was rapidly arrested within 24–48 hours, with rapid skin healing and no adverse effects. In *in vitro* studies, IVIG completely inhibited Fas-mediated keratinocyte apoptosis. This effect was related to the presence of naturally occurring Fas-blocking antibodies in IVIG, which inhibit Fas-mediated keratinocyte cell death.

Modulation of Complement Effector Function

The principal inflammatory mechanism in dermatomyositis (DM) is complement (C)-dependent microangiopathy with activation of C3 and deposition of the complement C5b-9 membrane attack complex (MAC) on the endomysial capillaries.[26,52] Basta et al. showed that IVIG can inhibit the uptake of C components on target tissues.[53] In patients with DM who were treated with IVIG, C3 deposition was reduced with corresponding decreases in complement expression on endomysial capillaries.[26,52] IVIG prevents the uptake of complement components and formation of the MAC on the endomysial capillaries in the muscle tissues of patients with DM. Consequently, IVIG allowed neovascularization to occur with reversal of the ischemic process, resulting in muscle tissue healing. This effect of IVIG on complement deposition may be relevant to other autoimmune neurological diseases, such MG, GBS, and CIDP, in which complement may be playing a role in the tissue damage.[26,52] Arumugam et al.[54] showed that IVIG protects the brain against ischemic injury mediated by complement in a mouse model of experimental stroke.

Effects of Ig on the Regulatory T-Cell Pathways

In a mouse model of multiple sclerosis (MS), IVIG has been shown to expand and enhance the function of FoxP3$^+$ regulatory T cells (Tregs)[55] while inhibiting the differentiation of T-helper 17 (Th17) and Th1 cells. This protective effect of IVIG was lost in mice that were depleted of Tregs.[56] These changes were independent of FcγRIIB and the Fc domain, since F(ab')$_2$ fragments led to similar changes in Th17 cells, Tregs, and clinical efficacy in this experimental allergic encephalomyelitis (EAE) model. Desialylated IVIG had the same immune-modulating effects as "native" IVIG.[57] Investigations by Trinath et al.[58] suggested that the mechanism by which IVIG induces Tregs was the enhancement of the cyclooxygenase 2 (COX-2) pathway via increased expression of prostaglandin E2 (PGE$_2$) from human DCs.

Kaufman et al. have extensively studied the effects of IVIG therapy in an ovalbumin-sensitized mouse model of asthma. They reported that IVIG markedly attenuated lung inflammation, decreased bronchial hyperresponsiveness to methacholine, and suppressed the Th2 pathway.[59] The draining pulmonary lymph nodes of IVIG-treated mice showed a significant increase in CD4$^+$CD25$^+$FoxP3$^+$ regulatory cells. IVIG-primed DCs on adoptive transfer to ovalbumin (OVA)–sensitized and challenged mice abrogated airway hyperresponsiveness and induced Tregs.[60] In their model system, Massoud et al.[61] reported that sialylated IgG bound to a novel C-type lectin receptor (*i.e.,* dendritic cell immunoreceptor [DCIR]) induced Tregs. Thus a number of studies have demonstrated the importance of the induction of FoxP3$^+$ Tregs by IVIG in modulating the autoimmune/antiinflammatory process. In contrast, Ig therapy in this murine model downregulated the Th-17 pathway.[62] In patients with KS and GBS, clinical improvement with IVIG therapy correlated with increased number and function of Tregs.[63,64]

Summary: IVIG in Treatment of Autoimmune and Inflammatory Diseases

Clearly, IVIG (IgG) has a number of immune-modulating effects and has been found to be an effective treatment for a wide spectrum of autoimmune and inflammatory diseases.[44,65] At present, IVIG is FDA approved for only a few autoimmune and inflammatory diseases. In autoantibody-mediated disease, the Fc domain appears to be the important IgG moiety that leads to immune modulation. The importance of sialylation of the Fc fragment remains controversial, as do the mediators involved (*e.g.,* IL-33 and IL-4). Differences in animal models, IVIG source, route and timing of the administration of the IVIG, mouse strain, and other variables may account for the differences in the experimental observations of laboratories. In T-cell mediated animal models of disease, such as EAE, there is strong evidence that there is upregulation of Tregs and inhibition of the Th17 pathway. These effects may be mediated by the F(ab')2 portion of the IgG molecule, and not the Fc domain. Furthermore, there is also controversy over the receptor on antigen-presenting cells (APCs), such as macrophages and DCs, which are involved in the immune-modulating process mediated by IVIG. In certain animal models, the SIGN-R1 (or in humans DC-SIGN) is important, and in other models (*e.g.,* murine asthma), a novel C-type lectin receptor (DCIR) appears to be important. These differences undoubtedly relate to the disease model. Two studies have demonstrated the importance of the PGE$_2$ pathway on IVIG-induced immune modulation mediated by the F(ab')2 portion of the IgG molecule. These observations may point to the possibility of alternative treatment regimens that employ the selective increase in PGE$_2$ in certain autoimmune or inflammatory disorders to increase Tregs and inhibit the production of certain cytokines. Further clarification in human disease models using *in vitro* human cells is important. These mechanisms are not mutually exclusive, and probably more than one mechanism is playing a role in the efficacy of Ig therapy in an autoimmune disease process. A better understanding of the pathogenic mechanisms involved in these diseases will undoubtedly lead to a more effective therapy with IVIG and more specific, modified forms of this biological product.

ON THE HORIZON

Translational Research Opportunities Related to Immunoglobulin Therapy

- Elucidation of mechanism(s) of action will lead to more precise bioengineered molecular products to treat autoimmune and inflammatory diseases.
- Clarification of the mechanisms of action of intravenous immunoglobulin (IVIG) in autoimmune and inflammatory diseases should lead to better understanding of the pathobiology of these diseases.
- Enhanced product purification—identification of the minor components in Ig products that may contribute to adverse reactions should lead to improved manufacturing processes and improved product tolerability for patients (*e.g.,* procoagulant factors, isohemagglutinins)
- Ig delivery—development of innovative approaches to the delivery of Ig products to patients to enhance safety and compliance.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Orange JS, Hossny EM, Weiler CR, et al. Use of intravenous immunoglobulin in human disease: a review of evidence by members of the Primary Immunodeficiency Committee of the American Academy of Allergy, Asthma and Immunology. J Allergy Clin Immunol 2006;117:S525–53.
2. Perez E, Bonilla T, Chinen J, et al. Update on the use of Immune Globulin (IG) in human disease: a review of evidence by members of the primary immunodeficiency committee of the American Academy of Allergy, Asthma and Immunology. J Allergy Clin Immunol 2017;139(3):S1–46.
3. Quartier P, Debre M, DeBlie J, et al. Early and prolonged intravenous immunoglobulin replacement therapy in childhood agammaglobulinemia: a retrospective survey of 31 patients. J Pediatrics 1999;134:589–96.
4. Orange JS, Grossman WJ, Navickis RJ, et al. Impact of trough IgG on pneumonia incidence in primary immunodeficiency: a meta-analysis of clinical studies. Clin immunol 2010;137(1):21–30.
5. Bonagura VR, Marchlewski R, Cox A, et al. Biologic IgG level in primary immunodeficiency disease: the IgG level that protects against recurrent infection. J Allergy Clin Immunol 2008;122(1):210–12.
6. Lucas M, Lee M, Lortan J, et al. Infection outcomes in patients with common variable immunodeficiency disorders: relationship to immunoglobulin therapy over 22 years. J Allergy Clin Immunol 2010;125(6):1354–60 e4.
7. Jiang F, Torgerson TR, Ayars AG. Health-related quality of life in patients with primary immunodeficiency disease. Allergy Asthma Clin Immunol 2015;11:27–38.
8. Routes J, Costa-Carvalho B, Grimbacher B, et al. Health-related quality of life and health resource utilization in patients with primary immunodeficiency disease prior to and following 12 months of immunoglobulin G treatment. J Clin Immunol 2016;36(5):450–61.
9. Jolles S, Orange JS, Gardulf A, et al. Current treatment options with immunoglobulin G for the individualization of care in patients with primary immunodeficiency disease. Clin Exp Immunol 2015;179: 146–60.
10. Berger M, Rojavin M, Kiessling P, et al. Pharmacokinetics of subcutaneous immunoglobulin and their use in dosing of replacement therapy in patients with primary immunodeficiencies. Clin immunol 2011;139(2):133–41. PubMed PMID: 21353644.
11. Berger M. Subcutaneous administration of IgG. Immunol Allergy Clin North Am 2008;28(4):779–802, viii.
12. Lawo J-P, Hubsch A, Rojavin M. Quantification of the wear-off effect towards the end of the IVIG infusion interval:Pooled data analysis. J Allergy Clin Immunol 2014;133(2):AB179.
13. Berger M, Jolles S, Orange JS, et al. Bioavailability of IgG administered by the subcutaneous route. J Clin Immunol 2013;33(5):984–90.
14. Ducruet T, Levasseur M-C, Roches A, et al. Pharmacoeconomic advantages of subcutaneous versus intravenous immunoglobulin treatment in a Canadian pediatric center. J Allergy Clin Immunol 2013;131(2):585–7.
15. Fasth A, Nystrom J. Quality of life and health-care resource utilization among children with primary immunodeficiency receiving home treatment with subcutaneous human immunoglobulin. J Clin Immunol 2008;28(4):370–8.
16. Misbah S, Sturzenegger MH, Borte M, et al. Subcutaneous immunoglobulin: opportunities and outlook. Clin Exp Immunol 2009;158(Suppl. 1):51–9.
17. Ponsford M, Carne E, Kingdon C, et al. Facilitated subcutaneous immunoglobulin (fSCIg) therapy- practical considerations. Clin Exp Immunol 2015;182:302–13.
18. Wasserman RL, Melamed I, Stein MR, et al. Recombinant human hyaluronidase-facilitated subcutaneous infusion of human immunoglobulins for primary immunodeficiency. J Allergy Clin Immunol 2012;130:951–7.
19. Wasserman RL. Progress in gammaglobulin therapy for immunodeficiency: from subcutaneous to intravenous infusions and back again. J Clin Immunol 2012;32(6):1153–64.
20. Bonilla FA. Pharmacokinetics of immunoglobulin administered via intravenous or subcutaneous routes. Immunol Allergy Clin North Am 2008;28(4):803–19, ix.
21. Stiehm ER. Adverse effects of human immunoglobulin therapy. Transfus Med Rev 2013;27:171–8.
22. Daniel GW, Menis M, Sridhar G, et al. Immune globulins and thrombotic adverse events as recorded in a large administrative database in 2008 through 2010. Transfusion 2012;52:2113–21.
23. Farrugia A, Quinti I. Manufacture of immunoglobulin products for patients with primary antibody deficiencies- the effect of processing conditions on product safety and efficacy. Front Immunol 2014;5: 1–8.
24. Horn J, Thon V, Bartonkova D, et al. Anti-IgA antibodies in common variable immunodeficiency (CVID): diagnostic workup and therapeutic strategy. Clin immunol 2007;122(2):156–62.
25. Rachid R, Bonilla FA. The role of anti-IgA antibodies in causing adverse reactions to gamma globulin infusion in immunodeficient patients: a comprehensive review of the literature. J Allergy Clin Immunol 2012;129(3):628–34.
26. Dalakas M. Mechanistics effects of IVIG in neuroinflammatory diseases: conclusions based on clinicopathologic correlations. J Clin Immunol 2014;34(Suppl1):S120–6.
27. Buchwald B, Ahangari R, Weishaupt A, et al. Intravenous immunoglobulins neutralize blocking antibodies in Guillain-Barre syndrome. Ann Neurol 2002;51:673–80.
28. Li N, Zhao M, Hilario-Vargas J, et al. Complete FcRn dependence for intravenous Ig therapy in autoimmune skin blistering diseases. J Clin Invest 2005;115:3440–50.
29. Hansen RJ, Balthasar JP. Effects of intravenous immunoglobulin on platelet count and antiplatelet antibody disposition in a rat model of immune thrombocytopenia. Blood 2002;100:2087–93.
30. Hansen RJ, Balthasar JP. Mechanisms of IVIG action in immune thrombocytopenic purpura. Clin Lab 2004;50:133–40.
31. Samuelsson A, Towers TL, Ravetch JV. Anti-inflammatory activity of IVIG mediated through the inhibitory Fc receptor. Science 2001;291:484–6.
32. Kaneko Y, Nimmerjahn F, Ravetch JV. Anti-inflammatory activity of immunoglobulin G resulting from Fc sialylation. Science 2006;313:670–3.
33. Anthony RM, Nimmerjahn F, Ashline D, et al. Recapitulation of IVIG anti-inflammatory activity with a recombinant IgG Fc. Science 2008;320(373-376).
34. Anthony RM, Wermeling F, Karlsson M, et al. Identification of a receptor required for the anti-inflammatory activity of IVIG. Proc Natl Acad Sci USA 2008;105:19571–8.
35. Anthony RM, Kobayashi T, Wermeling F, et al. Intravenous gammaglobulin suppresses inflammation through a novel T(H)2 pathway. Nature 2011;475(7354):110–13.
36. Tackenberg B, Jelcic I, Baerenwaldt A, et al. Impaired inhibitory Fcgamma receptor IIB expression on B cells in chronic inflammatory demyelinating polyneuropathy. Proc Natl Acad Sci USA 2009;106(12):4788–92.
37. Campbell IK, Miescher S, Branch DR, et al. Therapeutic effect of IVIG on inflammatory arthritis in mice is dependent on the Fc portion and independent of sialylation or basophils. J Immunol 2014;192(11): 5031–8.
38. Bayry J, Bansal K, Kazatchkine M, et al. DC-SIGN and a2,6-sialylated IgG Fc interaction is dispensable for the anti-inflammatory activity of IVIG. Proc Natl Acad Sci USA 2009;106:E24.
39. Yu X, Vasiljevic S, Mitchell DA, et al. Dissecting the molecular mechanism of IVIg therapy: the interaction between serum IgG and DC-SIGN is independent of antibody glycoform or Fc domain. J Mol Biol 2013;425(8):1253–8.
40. Burns J, Franco A. Ihe immunomodulatory effects off intravenous immunoglobulin therapy in Kawasaki disease. Expert Rev Clin Immunol 2015;11(I7):819–25.

41. Kaul R, McGeer A, Norrby-Teglund A, et al. Intravenous immunoglobulin therapy for streptococcal toxic shock syndrome—a comparative observational study. Clin Infect Dis 1999;28:800–7.
42. Bayry J, Negi VS, Kaveri SV. Intravenous immunoglobulin therapy in rheumatic diseases. Nat Rev Rheumatol 2011;7:349–59.
43. Tha-In T, Bayry J, Metselaar HJ, et al. Modulation of the cellular immune system by intravenous immunoglobulin. Trends Immunol 2008;29: 608–15.
44. Ballow M. The IgG molecule as a biological immune response modifier:mechanisms of action of intravenous immune serum globulin in autoimmuen and inflammatory disorders. J Allergy Clin Immunol 2011;127:315–23.
45. Kaveri SV, Lacroix-Desmazes S, Bayry J. The antiinflammatory IgG. N Engl J Med 2008;359:307–9.
46. St-Amour I, Laroche A, Bazin R, et al. Activation of cryptic IgG reactive with BAFF, amyloid beta peptide and GM-CSF during the industrial fractionation of human plasma into therapeutic intravenous immunoglobulins. Clin immunol 2009;133:52–60.
47. Vassilev TL, Kazatchkine MD, Van Huyen J-PD, et al. Inhibition of cell adhesion by antibodies to Arg-Gly-Asp (RGD) in normal immunoglobulin for therapeutic use. Blood 1999;11:3624–31.
48. Gill V, Doig C, Knight D, et al. Targeting adhesion molecules as a potential mechanism of action for intravenous immunoglobulin. Circulation 2005;112:2031–9.
49. Chang J, Shi PA, Chiang EY, et al. Intravenous immunoglobulins reverse acute vaso-occlusive crises in sickle cell mice through rapid inhibition of neutrophil adhesion. Blood 2008;111:915–23.
50. Letko E, Papaliodis DN, Papaliodis GN, et al. Stevens-Johnson syndrome and toxic epidermal necrolysis: a review of the literature. Ann Allergy Asthma Immunol 2005;94:419–36.
51. Viard I, Wehrli P, Bullanim R, et al. Inhibition of toxic epidermal necrolysis by blockade of CD95 with human intravenous immunoglobulin. Science 1998;282:490–3.
52. Lunemann J, Nimmerjahn F, Dalakas MC. Intravenous imunoglobulin in neurology-mode of action and clinical efficacy. Nat Rev Neurol 2015;11(2):80–9.
53. Basta M. Modulation of complement-mediated immune damage by intravenous immune globulin. Clin Exp Immunol 1996;104(Suppl. 1): 21–5.
54. Arumugam TV, Tang S, Lathia JD, et al. Intravenous immunoglobulin protects the brain against experimental stroke by preventing complement-mediated neuronal cell death. Proc Nat Acad Sci 2007;104:14104–9.
55. Ephrem A, Chamat S, Miquel C, et al. Expansion of CD4+CD25+ regulatory T cells by intravenous immunoglobulin: a critical factor in controlling experimental autoimmune encephalomyelitis. Blood 2008;111: 715–22.
56. Othy S, Hegde P, Topcu S, et al. Intravenous gammaglobulin inhibits encephalitogenic potential of pathogenic T cells and interferes with their trafficking to the central nervous system, implicating sphingosine-1 phosphate receptor 1-mammalian target of rapamycin axis. J Immunol 2013;190(9):4535–41.
57. Othy S, Topcu S, Saha C, et al. Sialylation may be dispensable for reciprocal modulation of helper T cells by intravenous immunoglobulin. Eur J Immunol 2014.
58. Trinath J, Hegde P, Sharma M, et al. Intravenous immunoglobulin expands regulatory T cells via induction of cyclo-oxygenase-2-dependent prostaglandin E2 in human dendritic cells. Blood 2013;122(8):1419–27.
59. Kaufman GN, Massoud AH, Audusseau S, et al. Intravenous immunoglobulin attenuates airway hyperresponsiveness in a murine model of allergic asthma. Clin Exp Immunol 2011;41(5):718–28.
60. Massoud AH, Guay J, Shalaby KH, et al. Intravenous immunoglobulin attenuates airway inflammation through induction of forkhead box protein 3-positive regulatory T cells. J Allergy Clin Immunol 2012;129(6): 1656–65 e3.
61. Massoud AH, Yona M, Xue D, et al. Dendritic cell immunoreceptor: a novel receptor for intravenous immunoglobulin mediates induction of regulatory T cells. J Allergy Clin Immunol 2014;133(3):853–63 e5.
62. Maddur MS, Vani J, Hegde P, et al. Inhibition of differentiation, amplification, and function of human TH17 cells by intravenous immunoglobulin. J Allergy Clin Immunol 2011;127(3):823–30 e1-7.
63. Olivito B, Taddio A, Simonini G, et al. Defective FoxP3 expression in patients with acute Kawasaki disease and restoration by intravenous immunoglobulin therapy. Clin Exp Rheumatol 2010;28(Suppl. 57): 93–7.
64. Tsurikisawa N, Saito H, Oshikata C, et al. High dose intravenous immunoglobulin treatment increases regulatory T cells in patients with eosinophilic granulomatosis with polyangiitis. J Rheumatol 2012;39: 1019–25.
65. Ballow M. Mechanisms of immune regulation by IVIG. Curr Opin Allergy Clin Immunol 2014;14:509–15.

MULTIPLE-CHOICE QUESTIONS

1. The BEST strategy for dosing of immunoglobulin (Ig) replacement therapy in patients with antibody deficiency diseases is by achieving a serum IgG level:
   A. Of 500 mg/dL
   B. Of 1000 mg/dL
   C. Of 200 mg/dL over baseline
   D. That minimizes the number of infections

2. Thromboembolic events associated with intravenous immunoglobulin (IVIG) therapy in autoimmune diseases is thought to be caused by:
   A. Osmolality of the product
   B. Sodium content
   C. Factor XIa
   D. IgA in the product

3. Which of the following receptors during IVIG treatment of autoimmune disease is enhanced on effector macrophages?
   A. CD25
   B. FcγRI
   C. FcγRIIB
   D. TACI

4. Which mechanism of action of IVIG is important in the rapid increase in platelet counts in patients with autoimmune thrombocytopenic purpura (ITP)?
   A. Saturation of the FcRn receptors
   B. Neutralization of cytokines
   C. Blocking the uptake of complement on platelets
   D. Blocking the Fc receptors on macrophages

85

Gene Therapy for Primary Immune Deficiency Diseases

Caroline Y. Kuo, Donald B. Kohn

Gene therapy as it is being applied for primary immune deficiencies (PIDs) represents an autologous hematopoietic stem cell transplant (HSCT), in which a patient's own stem cells are genetically corrected and transplanted back (Fig. 85.1). Thus, gene therapy for PID builds upon decades of experience using allogeneic HSCT from a healthy donor, where replacement of some or all of a PID patient's bone marrow hematopoietic stem cells (HSCs) with HSC from a healthy donor can be curative of the disease. Although initial efforts in gene therapy were not beneficial for the patients involved, there has been steady progress with the methods, and there is now clear-cut therapeutic efficacy using gene therapy/autologous HSCT for the three most commonly transplanted PIDs: severe combined immune deficiency (SCID), Wiskott-Aldrich syndrome (WAS), and chronic granulomatous disease (CGD)(Chapters 22, 34). Most gene therapy efforts to date have used **gene addition** methods in which a normal copy of the relevant gene is added to the patient's cells, usually using a viral vector. More recently, methods are under development to perform **gene correction**, using homologous recombination (HR) DNA repair of double-stranded breaks induced near the site of genome mutations by site-specific endonucleases.

KEY CONCEPTS

- Some severe primary immune deficiencies (PIDs) can be treated by transplantation of hematopoietic stem cells (HSCs), but this is optimal with a well-matched immune-compatible donor and may entail immune complications.
- Gene therapy using autologous HSCs that have been gene-corrected (either added or endogenously-corrected) may avoid the immune complications of allogeneic transplant and confer similar benefits.
- Stable gene addition to HSC can be done using integrating viral vectors, from retrovirus and lentivirus.
- Early-phase clinical trials have been performed by applying gene therapy to PIDs.
- Gene therapy using gamma-retroviral vectors led to immune reconstitution for several forms of severe combined immune deficiency, Wiskott-Aldrich syndrome and chronic granulomatous disease, but some patients developed leukoproliferative complications.
- More recent trials using safer vectors are continuing to yield clinical efficacy with good safety profiles.
- New techniques are being developed for precise gene editing in HSCs, which may allow application to a wider spectrum of PIDs.

The key potential advantage of autologous transplant with gene therapy, compared to allogeneic HSCT, is that it may have reduced risks and a better safety profile, because it eliminates the need for pre-transplant immune-suppressive conditioning and post-transplant immune suppression; essentially, there should be no risk of graft-versus host-disease (GvHD) from the autologous graft (Table 85.1). Additionally, gene therapy could have increased efficacy compared to allogeneic HSCT in some conditions, due to the potential to over-express the relevant gene product (*e.g.* adenosine deaminase enzyme) and lead to a supra-physiological effect from the engineered graft.

However, gene therapy may have unique risks: the potential for the genetic manipulation of the genome of the stem cells by either gene addition or gene correction causing malignant transformation and leukoproliferative complications. Also, gene therapy could have decreased efficacy if the percentage of transplanted stem cells that are successfully gene-corrected is low, if the level of expression of the inserted transgene is suboptimal, or if the process of *ex vivo* gene manipulation impairs the stem cell's capacity for long-term hematopoiesis. Additionally, developing gene therapy for PID will require a separate research endeavor for each genetic etiology (*e.g.*, ADA SCID, XSCID, Rag1 SCID, Rag2 SCID, etc.), whereas allogeneic HSCT is a more "one size fits all" approach requiring less individualization for similar classes of disorders.

GENE TRANSFER TO HEMATOPOIETIC STEM CELLS

For gene therapy to have an enduring effect in PID, the gene addition or correction must occur in the long-term pluripotent HSCs by some method that will lead to the corrective gene being passed on to the billions of progeny blood cells made from each HSC. Gene modification of the far more numerous, but short-lived, progenitor cells would only lead to transient presence of gene-corrected blood cells. The technical challenge is for the gene delivery to HSCs to be efficient (high percentage of the cells modified), to yield persistent expression, and to have minimal immediate cytotoxicity or long-term genotoxicity.

Many of the gene transfer methods used for research purposes, such as transfection, electroporation, and nano-particle delivery are themselves transient, and the inserted gene would be lost by dilution as the stem cells proliferate. Thus, most of the studies to date have used viral vectors derived from the Retroviridae family that integrate their genomes into the chromosomes of the target cell to achieve persistence of the normal gene (Fig. 85.2). Genes delivered by murine gamma-retroviral, human lentiviral (HIV), spumaviral (foamy) or alpha-retroviral vectors remain permanently covalently linked to the cellular chromosomal DNA for stable transmission to progeny cells. The steps involved in using these types of vectors to produce a gene-modified HSC graft are described in the Therapeutic Principles box.

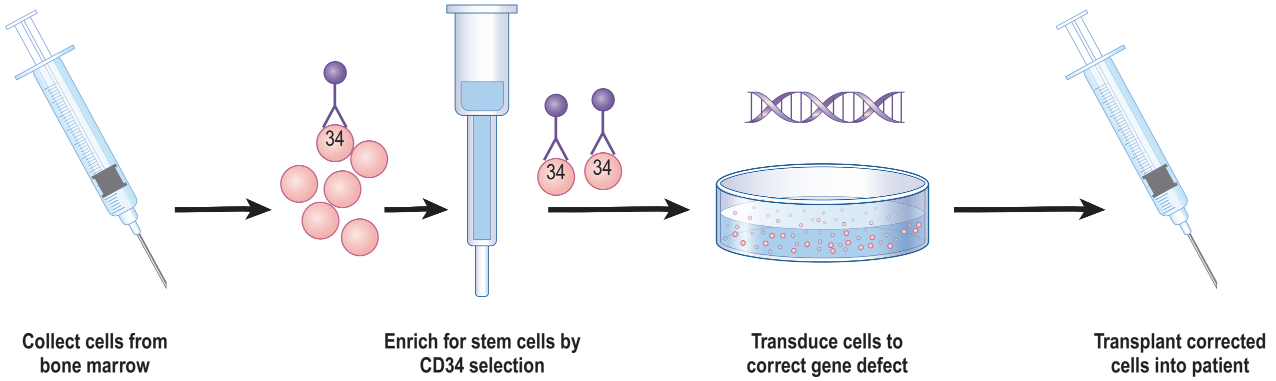

34
34
34
Collect cells from bone marrow
Enrich for stem cells by CD34 selection
Transduce cells to correct gene defect
Transplant corrected cells into patient

FIG 85.1 **Gene Therapy to Treat Primary Immune Deficiencies (PIDs) by Autologous Transplantation of Gene-Corrected Hematopoietic Stem Cells (HSCs).** HSC may be obtained from the bone marrow, by granulocyte–colony-stimulating factor (G-SCF) mobilization and leukapheresis (peripheral blood stem cells [PBSC]), or from umbilical cord blood of a PID patient. The HSC may be enriched by isolating the $CD34^+$ fraction of cells using immunomagnetic separation methods. The $CD34^+$ cell population is then cultured for gene manipulation (gene addition or gene correction). The gene-corrected autologous HSCs are then transplanted back to the patient.

TABLE 85.1 Advantages and Disadvantages of Allogeneic vs. Autologous (Gene Therapy) HSCT for PIDs

Transplant Type	Advantage	Disadvantage
Allogeneic HSCT	Normal function of relevant gene assured; benefit expected if sufficient donor chimerism	Need suitable matching donor
	Well-established and long-term experience with benefits/risks	Immune: GvHD and rejection risks.
	Excellent outcomes with matched sibling donor	Requires immune modulation: immune ablative conditioning; graft manipulation; and GvHD prophylaxis and treatment.
		May contribute to morbidity.
Autologous Gene Therapy HSCT	Patient is donor	Potential genotoxicity from gene addition or correction causing cell loss, dysfunction or transformation.
	No risks of GvHD.	Need to gene-modify high percentage of primary HSC with minimal toxicity.
	May not require immune suppression before (*e.g.*, flu/ATG) and after (*e.g.*, corticosteroids, calcineurin inhibitor) HSCT.	Low fractional correction may blunt efficacy.
	Risks of rejection may be less than with allogeneic cells.	Prior chemotherapy or marrow dysfunction may preclude use of autologous stem cells
	Transgene over-expression may have benefits (*e.g.*, adenosine deaminase).	Immunogenicity of transgene products is not well defined.
Gene Addition (*e.g.*, lentiviral vector)	Currently showing efficacy for SCID, WAS, and CGD (and for non-PIDs; *e.g.*, X-ALD, MLD, β-thalassemia).	Risks for insertional mutagenesis (gene disruption, gene activation) from semi-random insertions into target cell genomes.
	Current generation of SIN lentiviral vectors has significantly reduced genotoxicity, compared to earlier-used gamma-retroviral vectors.	Transgene not under normal transcriptional control. Function of transgene (level, lineage, longevity) may vary and be variegated.
Gene Correction (*e.g.*, nuclease/HDR)	Corrected endogenous gene should have normal function.	Risks of local or off-target gene disruption (insertion/deletion).
		Risk of translocations.

CGD, chronic granulomatous disease; *GvHD,* graft-versus-host disease; *HSC,* hematopoietic stem cells; *HSCT,* hematopoietic stem cell transplantation; *PID,* primary immune deficiencies; *SCID,* severe combined immune deficiency; *WAS,* Wiskott-Aldrich syndrome.

CLINICAL TRIALS OF GENE THERAPY FOR PRIMARY IMMUNE DEFICIENCIES

To date, clinical trials of autologous transplant/gene therapy have been performed for five PID disorders — multiple trials for ADA-deficient SCID, XSCID, CGD, and WAS, and one trial for leukocyte adhesion defect (LAD) (Table 85.2). These have occurred over roughly demarcated eras as the technology and clinical approaches advanced in the past three decades (Table 85.3). Because each genetic form of PID requires a separate development path from pre-clinical studies to define efficacy and safety through clinical trials, each individual PID will be discussed separately; although they share a common progression through the different eras.

Over these eras, the preferred vector for gene delivery to HSCs has progressed from murine gamma-retroviral vectors (γ-RV) to HIV-1–based lentiviral vectors in which the enhancers in the

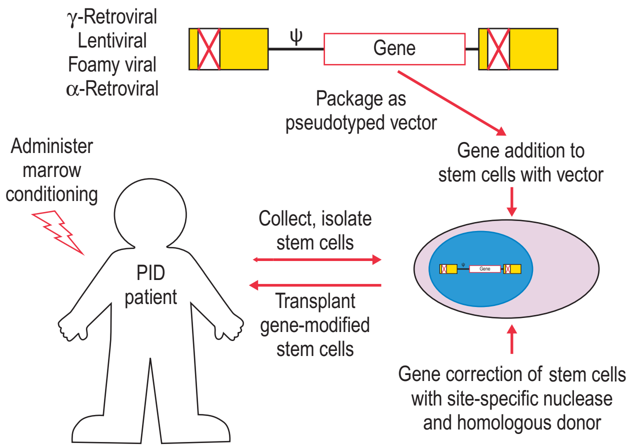


FIG 85.2 Autologous Transplantation of Gene-Corrected Hematopoietic Stem Cells (HSC). Gene addition may be performed using Retroviridae-derived vectors (from gamma-retroviruses, lentiviruses, foamy viruses or alpha-retroviruses) to transfer a normal copy of a relevant gene into HSCs collected and isolated from a primary immune deficiency (PID) patient. The gene-containing vector is packaged in a suitable envelope (pseudotyped) for gene addition to human HSC. Alternatively, the HSC may be gene-corrected using site-specific endonucleases to augment homologous recombination–directed gene correction. The patient may receive marrow cytoreductive conditioning to enhance engraftment after transplant of the gene-modified stem cells.

THERAPEUTIC PRINCIPLES

Steps for Gene Transfer to Hematopoietic Stem Cells for Clinical Transplantation

Package as Pseudotyped Vector

Transfect packaging cell with vector plasmid and virion/envelope plasmids.
Collect released vector from cell culture medium (DNAse).
Purify and concentrate vector (*e.g.,* ion-exchange chromatography, tangential flow filtration).
Aliquot. Certify. Release.

Transduce Stem Cells With Vector

Isolate CD34$^+$ stem/progenitor cells from clinical source (bone marrow, mobilized peripheral blood, cord blood).
Grow CD34$^+$ cells in serum-free medium plus recombinant cytokines (*e.g., ckit* ligand, flt-3 ligand, thrombopoietin).
Add vector to cells and culture for transduction.
Link vector sequences covalently into stem cell's chromosomal DNA.
Formulate and characterize cell product for release.

Administer Marrow Cytoreductive Conditioning

May deliver single or combination chemotherapeutic agents or monoclonal antibodies to "make space" in marrow.

Transplant Gene-Modified Stem Cells

Infuse gene-modified cell product intravenously.
Stem cells engraft and transmit transgene to all progeny blood cells.
Transgene produces necessary gene product to correct genetic deficiency.

long-terminal repeats are self-inactivated (SIN vectors). Lentiviral vectors have several potential advantages over γ-RV, including the ability to more effectively transduce human HSCs in a shorter period of *ex vivo* culture, the capacity to carry longer and more complex genetic sequences (such as cellular gene enhancers and promoters), and less tendency to insert near the 5′ ends of genes which may decrease risks for *trans*-activating expression of the adjacent cellular genes. The combination of recombinant cytokines used to activate the HSCs to facilitate transduction was advanced as factors were identified that acted on the earliest HSCs, including flt-3 ligand and thrombopoietin, combined with c-*kit* ligand.

Role of Cytoreductive Conditioning to Facilitate Engraftment

Initial trials of gene therapy for PID did not administer pre-transplant cytoreductive conditioning, due to potential risks of chemotherapy or radiation with unknown prospects of benefit from the gene transfer procedure. It is well-known from multiple transplant studies in mice and large animal models that there is minimal if any engraftment of autologous HSCs when given without prior conditioning, unless extraordinary numbers of cells (*e.g.*, 30–50x higher than standard cell dose/kg) are given; even mega-dose transplants without conditioning lead to only low levels of engraftment (*e.g.*, 1%), although this may be persistent.[1] While there was initial reluctance to use conditioning in gene therapy trials when there had not previously been efficacy, now with the clear-cut benefits that may be obtained from gene therapy, the necessity to use conditioning for autologous gene therapy HSCT (which is less than that needed for allogeneic HSCT) is becoming recognized as the standard. The amount and types of chemotherapy drugs used for conditioning has varied, depending on the disease setting. For SCID, where a relatively low level of gene-corrected HSCs may support immune reconstitution, reduced-intensity conditioning (RIC), such as relatively low dose busulfan alone (*e.g.* 4–6 mg/kg), may be sufficient. For other disorders where less of a selective advantage may exist for the gene-corrected cells, a higher level of engraftment of modified HSC may be needed, and thus stronger conditioning regimens have been used, reaching levels for myeloablative conditioning (MAC) (*e.g.*, busulfan, 12–16 mg/kg). For the WAS where it may be necessary to ablate the pre-transplant immunity to reduce risks for post-transplant autoimmunity, immune-suppressive drugs (*e.g.*, fludarabine, rituximab) have been added to conditioning regimens. These combined iterative approaches to improving gene therapy have led to the current state, where clinical benefits are being routinely achieved, as detailed below for each disorder.

New methods are being explored to "make space" in the marrow using alternatives to chemotherapy and radiation, such as monoclonal antibodies to markers present on HSCs. These are moving toward clinical assessments and may in the future facilitate engraftment with lower short-term and long-term risks than with chemotherapeutic agents.

Adenosine Deaminase (ADA)-Deficient Severe Combined Immune Deficiency (SCID; Chapter 35)

The first clinical trial of gene therapy for an inherited disorder (other than a premature attempt at gene therapy for beta-thalassemia) was directed against adenosine deaminase (ADA)-deficient Severe Combined Immune Deficiency (SCID). ADA-SCID was the first form of human SCID for which the responsible gene was identified and cloned, allowing gene therapy

TABLE 85.2 Clinical Trials of Gene Therapy for Primary Immune Deficiencies

PID	Investigator(s)	Year	Vector/Target	Conditioning	NCT #
ADA SCID	Blaese, Anderson, Culver	1990	LASN (MLV LTR) PBMC	None	—
	Bordignon, Mavilio	1992	G1ADA (MLV LTR) PBMC and BM	None	—
	Hoogerbrugge, Valerio	1993	MLV-ADA BM CD34$^+$	None	—
	Kohn, Parkman	1993	G1NA-ADA (MLV LTR) UCB CD34$^+$	None	—
	Onodera/Sakiyama	1995	LASN (MLV LTR) PBMC	None	—
	Aiuti/Roncarolo	1998	G1ADA (MLV-LTR) BM CD34$^+$	Busulfan 4 mg/kg	00599781
	Gaspar/Thrasher	1999	SFFV-ADA-wpre BM CD34$^+$	Melphalan; busulfan	—
	Otsu/Ariga	2003	GCsap-M-ADA BM CD34$^+$	None	—
	Kohn/Candotti	2001	MND-ADA and GCsap-M-ADA BM CD34$^+$	None (4). Busulfan (6)	—
	GlaxoSmithKlein	2008	G1ADA BM CD34$^+$	Busulfan	00598481
	Kohn/Candotti	2009	MND-ADA BM CD34$^+$	Busulfan 90 mg/m2	00794508
	Gaspar	2012	EFS-ADA CD34$^+$ BM/PBSC	Busulfan	01380990
	Kohn	2013	EFS-ADA BM CD34$^+$	Busulfan	01852071
XSCID	Cavazzana-Calvo/ Fischer	1998	MFG-IL2Rg BM CD34$^+$	None	—
	Thrasher/Gaspar	1999	MLV-IL2Rg BM CD34$^+$	None	—
	Thrasher/Fischer/ Cavazzana/Pai/ Williams	2011	SIN γRV EFS-IL2Rg BM CD34$^+$	None	Fr: 01410019 UK: 01175239 US: 01129544
	Sorrentino	2012	EFS-IL2Rg-Ins BM CD34$^+$	None/ Busulfan	01512888
	Malech/DeRavin	2011	EFS-IL2Rg-Ins CD34$^+$ PBSC	Busulfan	01306019
LAD	Hickstein, Bauer	1999	MLV-CD18 BM	None	—
CGD	Malech	1995	MLV-p47 CD34$^+$ PBSC	None	—
	Malech	1998	MLV-gp91phox	None	—
	Ott/Grez	2009	SFFV-gp91phox CD34$^+$ PBSC	8 mg/kg	00927134
	Kang/Malech	2010	MT-gp91phox CD34$^+$ PBSC	10 mg/kg	—
	Thrasher/ Serve/ Reichenbach	2013	pChim-gp91phox (G1XCGD) CD34$^+$ PBSC	pK adjusted, ~12 mg/kg	01855685
	Kohn/Williams/Kang	2015	pChim-gp91phox (G1XCGD) CD34$^+$ PBSC	pK adjusted, ~12 mg/kg	02234934
WAS	Klein	2008	SFFV-WAS gRV	Busulfan	—
	Aiuti	2011	1.6hWASP-WPRE LV	Bu/Flu/Rituximab	01515462
	Hacein-Bey Abina… Cavazzana	2011		Bu/flu (rituximab or alemtuzimab for auto-immunity)	02333760
	Thrasher	2011		Bu/Flu	01347242
	Pai	2011		Bu/Flu	01410825

CGD, chronic granulomatous disease; *LAD*, leukocyte adhesion defect; *PID*, primary immune deficiencies; *SCID*, severe combined immune deficiency; *WAS*, Wiskott-Aldrich syndrome; *XSCID*, X-linked SCID.

TABLE 85.3 Eras in the Advancement of Clinical Gene Therapy for PIDs

Era	Predominant Vector	Growth Factors	Conditioning Regimen	PIDs Treated	Typical Outcomes
Early (~1990–1999)	γ-RV	IL3/IL6/SCF	None	ADA XSCID CGD LAD	Insufficient engraftment. No efficacy.
Middle (~1998–2006)	γ-RV	SCF/FLT3L/TPO	RIC-MAC	ADA XSCID CGD WAS	Efficacy. Genotoxicity in some.
Current (~2007–present)	SIN LV	SCF/FLT3L/TPO	RIC-MAC	ADA XSCID CGD WAS	Efficacy. Good safety profile.

to be approached.[2] As a ubiquitously-expressed housekeeping enzyme, the non-regulated expression of an introduced ADA gene was expected to be tolerated and potentially beneficial. Also, the ability to treat ADA-SCID with a bone marrow transplant from a matched sibling donor without the use of cytoreductive or immune suppressive conditioning was taken to imply that there is a selective advantage for ADA-replete T lymphocytes over ADA-deficient cells and that only a modest number of engrafted gene-corrected HSC may provide clinical benefit.

A series of clinical trials were performed using gamma-retroviral vectors and targeting either peripheral blood T cells or bone marrow hematopoietic stem cells. While initial studies in the early 1990's did not achieve clinical efficacy, subsequent trials that applied reduced-intensity condition prior to transplant to increase engraftment of gene-corrected HSCs have led to immune restoration in the majority of more than 40 treated patients, without vector-related complications.[3-6] The relatively low dose chemotherapy and absence of graft versus host disease makes these autologous transplants well-tolerated with essentially none of the clinical side-effects seen in allogeneic HSCT with conditioning. One of the retroviral vectors used in studies at Telethon Gene Therapy Program for Genetic Diseases, of the Hospital San Raffaele, in Milan Italy was licensed by GSK and has received licensure approval in the European Union, a major advance for gene therapeutics which is available for patients. Other ongoing clinical trials in the US and the UK are using a lentiviral vector for ADA SCID with excellent initial clinical results in terms of efficacy and safety. Thus, gene therapy for ADA SCID has become a major treatment option for patients, as the efficacy and safety profile have been favorable and may exceed those of matched unrelated or haplo-identical transplants, although there have not been randomized controlled trials.

X-linked SCID (Chapter 35)

The second genetic form of human SCID for which the relevant gene was identified and cloned is the X-linked form (XSCID).[7,8] The responsible gene, $IL2R_\gamma$, encodes the common cytokine receptor chain (or γ_c), a component of several multimeric receptors for a family of lymphopoietic cytokines: IL-2, IL-4, IL-7, IL-9, IL-15 and IL-21. Defects of γ_c severely impair the development of multiple components of the immune system and patients typically have severely reduced numbers of T and NK cells, and may have immature but non-functional B cells. Retroviral-mediated gene transfer of a normal human $IL2R_\gamma$ cDNA was shown to restore cytokine-induced signaling activity and lymphocyte function in patients' cells and in murine models.[9-12]

XSCID patients were first treated by gene therapy using gamma-retroviral vectors.[5,13] No conditioning was given, relying on the potent selective survival and proliferation advantage gene-corrected lymphoid cells were expected to have in lymphopenic SCID patients. Indeed robust reconstitution of T cell immunity was achieved, although there were variable responses of B cells. However, after 2-3 years from treatment, several of the patients developed a severe complication from the retroviral vector of leukemia-like leukoproliferation. New vectors were developed that lacked the strong enhancer elements of the retroviral vectors that caused the insertional oncogenesis (self-inactivating or "SIN" vectors). One published study using a SIN gamma-retroviral vector demonstrated immune restoration without any evidence of leukoproliferation. New studies are in progress using lentiviral vectors.[14] The first of these studies treated a group of XSCID patients who had undergone allogeneic transplant some years earlier, but who had not achieved complete immune reconstitution and had significant morbidity from poor growth and development.[15] Following gene therapy using a lentiviral vector and non-myeloablative conditioning, all have improved general well-being, with development of sufficient antibody production to be able to stop immunoglobulin replacement therapy. Studies treating more typical XSCID infants are now underway at several centers in the US and Europe. Thus, the ability of gene therapy to provide clinical benefit to these difficult patients is indeed gratifying.

Leukocyte Adhesion Defect (LAD; Chapter 22)

Bauer and Hickstein performed gene therapy for two LAD patients using a gamma-retroviral vector expressing human CD18 to transduce bone marrow cells without conditioning.[16] There were no beneficial effects of gene therapy and both later succumbed to complications of their immune deficiency. These investigators have developed a highly effective vector from the spumavirus called Human Foamy Virus and a clinical trial is under development.[17]

Chronic Granulomatous Disease (CGD; Chapter 22)

Chronic Granulomatous Disease (CGD) is a rare primary immunodeficiency (1 in 200,000 live births in the US), first termed Fatal Granulomatous Disease of Childhood in 1959 to describe individuals affected by recurrent, invasive bacterial and fungal infections complicated by granuloma formation.[18] Clinical presentation is explained by defects in any of the components that comprise the nicotinamide adenine dinucleotide phosphate (NADPH) oxidase complex leading to phagocyte dysfunction. In its fully assembled form, the NADPH oxidase complex is composed of five proteins, two of which are membrane-bound (gp91phox and p22phox) and three of which are cytosolic (p47phox, p67phox, and p40phox). All five components are necessary for proper NADPH oxidase function without which affected individuals are particularly predisposed to infections with *Aspergillus* species, *Staphylococcus aureus*, *Burkholderia cepacia*, *Serratia marcescens*, *Nocardia* species, and less commonly, *Salmonella* and BCG infections outside of North America. Defects in gp91phox encoded by *CYBB* account for the most common X-linked form of CGD while the remaining defects are inherited in an autosomal recessive (AR) pattern.[19] Since its initial description almost six decades ago, CGD has evolved from a disease with early mortality to one with relatively good outcomes and multiple treatment options.

While prophylactic antimicrobial and immunomodulatory agents have dramatically decreased infection rates in CGD patients, curative therapy can only be achieved with HSCT. Transplant outcomes have continued to improve over the years, with good results reported even in patients with high-risk features such as intractable infections and auto-inflammation using reduced-intensity conditioning protocols.[13] Nevertheless, allogeneic HSCT can still be complicated by graft-versus-host disease and pre-existing infections and is not preferred for those without an HLA-matched stem cell donor. For these patients, autologous HSCT with gene-modified cells is becoming a more viable and realistic option.

The first gene therapy trials for CGD began in the mid-1990's again using gamma-retroviral vectors. As with ADA SCID, initial trials in which pre-transplant conditioning was not used did not lead to efficacy as there was minimal if any engraftment of gene-corrected stem cells. In subsequent studies where non-myeloablative conditioning with busulfan was used, increasingly higher levels of engraftment of corrected stem cells was achieved. The best and worst of these studies was one using a gammaretroviral vector with a potent transcriptional control element to drive high level expression of the *gp91phox* gene. The three treated subjects had initial development of neutrophils with restored oxidase function and cleared severe resistant infections. However, insertional oncogenesis occurred in this study also, and myelodysplasia or frank myeloid leukemia developed. More recently, trials are using a lentiviral vector with a myeloid-specific transcriptional control element intended to drive *gp91phox* expression in mature myeloid cells, where it is needed, but to not have activity in HSCs that are the likely targets for transformation. Initial results are showing safety and evidence of efficacy.[14]

Recently, the group at NIH has also investigated the use of site-specific endonucleases (a zinc-finger nuclease, discussed below) to target integration of a normal gp91phox cDNA expression cassette to the AAVS1 gene "safe harbor site".[15] Good levels of engraftment were demonstrated in mouse models, with up to 11% of human cell expressing gp91phox *in vivo*. This work represents the first step towards targeted gene editing for X-CGD and suggests that it can also be possible to integrate the transgene in its natural location in the genome under control of its endogenous promoter.

WISKOTT-ALDRICH SYNDROME (WAS; Chapter 35)

The Wiskott Aldrich Syndrome (WAS), initially described as an X-linked syndrome in kindreds in Germany and the US, presents with multiple clinical manifestations, including the classical triad of immune deficiency, eczema and thrombocytopenia. The complex immune deficiency involves defects of T, B, NK and antigen-presenting cells. The identified WAS gene encodes a 501 amino acid proline-rich protein that has multiple identified functional domains, placing it at the nexus of intracellular signaling and cytoskeleton control. Once the WAS cDNA was obtained, gamma-retroviral vectors were constructed and shown to correct several manifestations of the disorder in patient derived cells and in murine models. However, it is not fully known what levels of WAS transgene expression are needed to safely and effectively correct the major manifestations of lymphocyte and platelet dysfunction. A concern has been raised that sub-optimal levels or low frequency of expression of WAS protein (WASP) could allow auto-immunity to develop. If correction is only partial, (*e.g.*, some proportion of B cells are not gene corrected and have defective auto-regulatory function), auto-immunity may occur.[16]

A first trial of gene therapy for WAS used a gamma-retroviral vector with a very potent long terminal repeat promoter enhancer (from the Spleen Focus Forming Virus). G-CSF-mobilized peripheral blood stem cells were used as the HSC source and conditioning with myeloablative dosages of busulfan was given prior to transplant. There were excellent initial results in terms of immune reconstitution and platelet counts, demonstrating that gene therapy can be therapeutic for this disorder. However, there was subsequently a very high incidence of acute leukemia among these patients, developing within a few years from treatment, with ALL, AML or both occurring in a shocking 7 of 9 subjects.[17] The mechanism of insertional oncogenesis from the gamma-retroviral vector was clearly the cause, with the leukemias showing clonal vector integrations adjacent to known proto-oncogenes, such as *LMO2* and *MDS1*.

While that trial was being initiated, work was ongoing through a multi-center collaboration in the EU to develop a SIN lentiviral vector for WAS that uses the promoter from the WAS gene *per se* to drive expression of the WAS cDNA. Extensive pre-clinical studies showed the efficacy of this vector to improve immunologic and hematologic parameters in murine and human disease models.[18,19] And, it displayed significantly lower risks for genotoxicity than the type of gamma-retroviral vector used in the first trial.

The lentiviral vector with the WAS promoter has now been used in parallel but independent phase I/II clinical trials performed in several countries.[20,21] Relatively high intensity conditioning was administered to attempt to obtain high-level engraftment of gene-corrected HSCs, with some differences among the centers in the precise conditioning regimen used. All used moderate doses of busulfan (pharmacokinetically [pK]-adjusted) and fludarabine, and variably, rituximab or other serotherapy agents. These trials have uniformly demonstrated efficacy and safety. Findings have included clinical improvement in general health, bleeding incidents, and eczema, and as expected for an autologous transplant, no problems from graft versus host disease. Most have had good recovery of T, B, NK cell numbers and functions. However, there has been only modest and variable improvements for platelet levels (*e.g.* 20-60,000/uL), with evidence that higher transplanted dose of transduced cells led to higher platelet counts.[20] No new onset of auto-immunity has been reported, although in some cases pre-existing problems have persisted. In fact, decreases of indices of auto-immunity and improved B cell tolerance have been documented after gene therapy for WAS.[22] Polyclonal vector integration patterns were seen, with no reported development of clonal expansions or frank leukoproliferative complications. Integration site patterns resemble those seen in other trials using lentiviral vectors into human HSCs, with highly diverse vector integration sites and no predilection for cancer-related genes, significantly different than the pattern seen in the gamma-retroviral vector trials.

These results are promising and of significant clinical benefit and safety, but would be better if more normal levels of platelets could be achieved. It is possible that the level of expression of WASP from the WAS gene promoter in the lentiviral vector is inadequate to support normal platelet production or survival. If so, introducing more than one copy of the vector per cell could lead to higher levels of WASP and higher platelet counts. The WAS gene promoter that has been used as a regulatory element (a 1.6 kb fragment from the 5′ end of the gene) may be sub-optimal to drive sufficient expression; there may be other regulatory elements from the WAS genomic locus not included in the vectors that are needed for some essential aspect of expression (lineage, level, longevity). Astrakhan et al[23] reported that a lentiviral vector using a stronger internal promoter (a retroviral LTR) to drive the WASP cDNA led to better restoration of T and B cell activity in the WAS gene knock-out mouse model. Alternatively, the absolute number of gene-corrected HSCs engrafted may mediate the platelet levels and a higher percentage of transduced cells or higher numbers of gene-corrected cells may support higher platelet levels. Relative risks of multiple integrants versus a stronger promoter are not known. As for the other disorders discussed here, direct gene correction should yield normal, physiologic expression of the WAS gene.

GENE THERAPY CONSIDERATIONS FOR OTHER PID

Currently, a new development project is needed to bring gene therapy to clinical application for each individual PID-causing locus, such as the more than 20 human genes that may cause SCID, 5 or more CGD loci, several for Hemophagocytic Lymphoistiocytosis (HLP), X-linked lymphoproliferative disease (XLP), etc. Each gene and disease setting poses different challenges in terms of necessary gene transfer efficiency, level of expression, need for regulation of the transferred gene, safety considerations, and measurable end-points.

X-linked agammaglobulinemia (XLA; Chapter 34) is another logical disease to consider treating by autologous transplant with gene therapy, since normal B cell development from HSC with a normal Bruton tyrosine kinase (BTK) gene should correct the immune deficiency. Because of the good clinical effects from immunoglobulin replacement therapy for XLA and the toxicities from HSCT, especially from chemotherapy and GVHD, HSCT is rarely done for XLA patients. A few XLA patients have had allogeneic transplants from healthy donors and have developed B cell reconstitution.

Gene therapy studies in *BTK* gene knock-out mice have shown that it can lead to immune reconstitution, using lentiviral vectors with B lymphoid-specific promoters.[24,25] While no adverse effects were seen from constitutive expression of the *BTK* gene in these murine studies, these do not constitute formal toxicology studies which would be needed before clinical applications. It is likely that regulated expression of *BTK*, rather than constitutive, ubiquitous expression, is needed for highest efficacy and safety. In theory, lentiviral viral vectors using components of the *BTK* gene transcriptional regulatory sequences could yield vectors with the desired expression specificity. Alternatively, gene correction of the *BTK* gene, using the methods discussed below, could restore precise *BTK* expression regulation and may have the greatest safety profile.

Common Variable Immune deficiency (CVID) comprises the most common severe human PID (Chapter 34). While immunoglobulin replacement therapy can ameliorate the immune deficiency caused by hypogammaglobulinemia, the high rate of other clinical complications that CVID patients may experience necessitate new treatments. However, gene therapy requires knowing the responsible pathogenic gene, and has been limited to monogenic disorders. To date, a single gene defect has not been identified in the majority of patients with CVID. There are some known CVID genes, including *TNFRSF13B* (encoding TACI 8–10%), *TNFRSF13C* (encoding BAFF-R), COS, CD19, and MSH5, which in total may be responsible for 10–15% of all CVID patients. A separate gene therapy project would be needed to develop treatment for each causal gene, through the full spectrum from pre-clinical activities to clinical trial performance. And, because most of these known CVID-causing genes are involved in cell stimulation and signaling, they may require regulated, rather than ubiquitous, constitutive expression for safety. Because of these constraints, it is not currently possible to apply gene therapy for the majority of patients with CVID.

There has been growing recognition of numerous immune dysregulation and auto-inflammatory syndromes due to autosomal dominant gain-of-function mutations (*e.g.* STAT3, MEFV, IL-1, NFkB, interferon pathways).[26] The pathology attributed to blood cells (as opposed to those effects on other tissues) may be corrected by HSCT, and here too autologous transplant/gene therapy may have advantages over allogeneic. Gene modification of a patient's HSC by addition of a shRNA cassette or gene disruption using site-specific endonucleases may suppress expression of the dominant gene. And, direct correction of the pathogenic mutation could also be beneficial, as discussed below. It is likely that the gene modification would need to be efficient to yield a high fractional correction of the engrafting stem cells. In this setting, some pre-transplant immune suppression may be needed to ablate pre-existing auto-immunity. And, perhaps most challenging for gene therapy are the PIDs that also have major somatic or developmental problems, such as chromosomal abnormalities, ataxia-telangiectasia, and others. Here the gene therapy with the HSC may benefit that component of the disorder but would not address the others; systemic delivery of genes or delivery to the CNS is under study, but it is not yet sufficiently efficient for most clinical needs.

GENE CORRECTION (EDITING) FOR GENE THERAPY OF PID

A major paradigm shift in gene therapy is under way as methods to perform precise edits of cells' genomes are being developed. As an alternative to the semi-random insertion of normal copies of the relevant gene delivered by a viral vector, as discussed for all of the studies above, techniques are being established to either correct specific bases in the DNA, or to insert (or remove) gene sequences at specific sites by harnessing cellular DNA repair pathways. These DNA repair mechanisms normally correct the many double-stranded DNA breaks that occur during DNA replication or from environmental genotoxic agents (ionizing radiation, chemicals). To simplify a highly complex topic, there are two major DNA repair pathways for rejoining the sequences that flank a double-stranded DNA break: non-homologous end joining (NHEJ) and homologous recombination (HR). NHEJ reconnects the broken ends of the chromosomes in a way that often leads to insertion or deletion of DNA bases (indels) at the junctional site. This is, in essence, a mutagenic process and may be used to disrupt genes to knock-out their activity, examples being the HIV-1 co-receptor CCR5, dominant-active

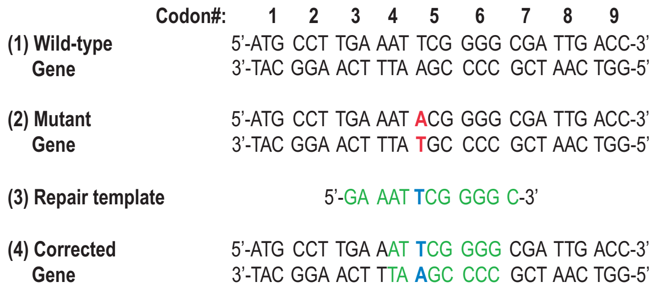

Codon#: 1 2 3 4 5 6 7 8 9
(1) Wild-type Gene
5'-ATG CCT TGA AAT TCG GGG CGA TTG ACC-3'
3'-TAC GGA ACT TTA AGC CCC GCT AAC TGG-5'
(2) Mutant Gene
5'-ATG CCT TGA AAT ACG GGG CGA TTG ACC-3'
3'-TAC GGA ACT TTA TGC CCC GCT AAC TGG-5'
(3) Repair template
5'-GA AAT TCG GGG C-3'
(4) Corrected Gene
5'-ATG CCT TGA AAT TCG GGG CGA TTG ACC-3'
3'-TAC GGA ACT TTA AGC CCC GCT AAC TGG-5'

FIG 85.3 Site-Specific Gene Repair by Homologous Recombination (HR). In this example, instead of the wild-type gene sequence (1), a patient's gene (2) has a mutation from a base-pair substitution of an A for a T at the start of the 5th codon (red). An artificial donor template (3), here as a single-stranded deoxyoligonucleotide of 12 bases in length (green), is provided with the correct base present at the site of the patient's mutation (blue). If the donor template is used to repair a double-stranded DNA break induced near the mutation by a site-specific endonuclease (red arrow), sequences from the donor (green) will be incorporated into the patient's gene (4), introducing the normal corrective base pair (blue).

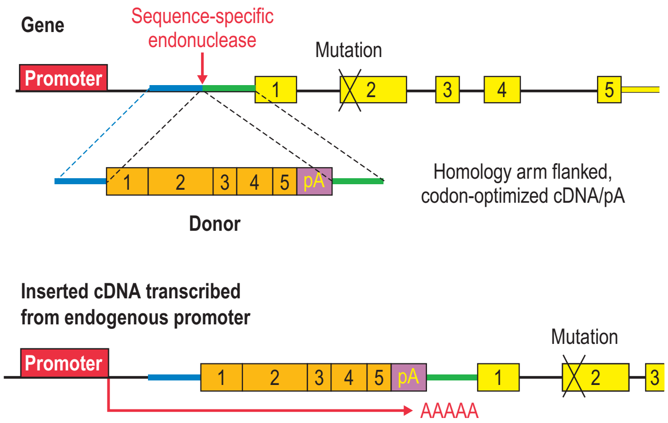

Gene
Sequence-specific endonuclease
Mutation
Promoter
1
2
3
4
5
1
2
3
4
5
pA
Homology arm flanked, codon-optimized cDNA/pA
Donor
Inserted cDNA transcribed from endogenous promoter
Mutation
Promoter
1
2
3
4
5
pA
1
2
3
AAAAA

FIG 85.4 Site-Specific Insertion of a cDNA to Override Down-Stream Gene Mutations. A prototypical gene is shown with 5 exons (yellow boxes 1–5) and an upstream promoter (red box). A mutation in exon 2 (black X) inactivates the gene. A sequence-specific endonuclease (red arrow) is designed to introduce a double-stranded DNA break, in this case in the 5′ untranslated region of the gene (blue and green lines). A cDNA molecule contains the contiguous exons of the gene (orange 1–5), codon-optimized to increase expression and to eliminate homology with the endogenous exons to eliminate illegitimate recombination events, and a polyadenylation signal (pA) to terminate transcription. The cDNA is flanked by the sequences from the endonuclease cleavage site (blue and green lines). The donor can be inserted into the nuclease target site by homologous recombination, placing the cDNA under transcriptional control of the endogenous gene promoter; the cDNA transcript (red arrow) would override any downstream mutations in the gene.

transcriptional factors such as CTLA-4, or STAT3 Gain of Function alleles, a transcriptional repressor of fetal hemoglobin, etc. There have been clinical trials for patients with HIV infection in which the gene for CCR5 HIV co-receptor was disrupted using a zinc finger nuclease (ZFN) to make a double stranded DNA break while NHEJ was allowed to repair the break, leading to indels that inactivated the gene and co-receptor expression.[27]

HR is a more precise repair mechanism that normally uses a copied sister chromosome or the other homologue as a template to repair the break; sequences of the template are copied into the repair site and if there are differences, the template acts as a donor for the new sequences. For gene correction, an artificial donor template is provided to the cells to instruct the introduction of the intended sequence changes (Fig. 85.3). Beyond its use for modification of single base pairs, as illustrated here, HR can be used to introduce whole gene sequences into the target site by flanking a gene cassette with "homology arms" that consist of the DNA sequences homologous to the target site (Fig. 85.4).

Although HR has been used to introduce genes into cells, such as murine embryonic stem cells to make gene knockout and knockin mice, it is generally a low-frequency event (occurring in $1/10^6$–$1/10^4$ cells) and requires the use of selectable markers to isolate the rare desired recombinant. Whereas cloning and expansion of murine (and human) embryonic stem cells can be done to produce populations of the rare recombinant cells, this is not possible with primary HSCs, which cannot be expanded to any great degree from single cells. Methods achieving high efficiency of gene modification with low cytotoxicity in large numbers of primary stem cells are needed for clinical applications to autologous HSCT. The major breakthrough in this area comes from the observation that HR is vastly more frequent when a double-stranded DNA break is introduced close to the target site;[28] then, the repair donor molecule can be used at efficiencies in the range ~1–50% of treated cells.

Several classes of designer site-specific endonucleases have been derived, including ZFN, introduced above; Transcription Activator-Like Effector Nucleases (TALENs); and more recently Clustered, Regularly Interspaced, Short Palindromic Repeats (CRISPRs) that allow the introduction of a double-stranded DNA break at unique sites in the mammalian genome with high specificity. Current methods introduce the nuclease into HSCs in a method where it will only be present transiently, such as electroporation of *in vitro* transcribed messenger RNA encoding the nuclease proteins. The homologous donor template is introduced as either a short (*e.g.* 50–100 bases) single-stranded deoxy-oligonucleotide that is co-electroporated with the nuclease, or as a longer sequence carried by a viral vector that does not integrate into the target cell chromosome, including adeno-associated virus (AAV) or an integrase-defective lentiviral vector (IDLV). The development and application of this technology has evolved at an incredibly rapid pace and is approaching the levels of efficiency needed for clinical applications.

One caveat is that potential off-target activity of the nucleases could cause genotoxicity from either disruption of unintended genes or even introduction of chromosomal translocations between two cut sites produced simultaneously, *e.g.*, one on-target and one off-target. Current studies are assessing the consequences of this potential genotoxicity in human HSCs, while ongoing work is seeking to improve the specificity of the nucleases to eliminate or greatly minimize off-target activity.

Gene correction will have several advantages over gene addition. It would avoid the potential problems from randomly inserting vectors that may disrupt or over-stimulate adjacent cellular genes, as discussed above for the retroviral vectors. Crucially, correcting the endogenous gene keeps its expression under normal physiological control. This will be essential for gene therapy of many PIDs, where it could be problematic if the gene is expressed constitutively from a viral vector. For example, it was shown that retroviral vector delivery of a normal CD40 ligand gene corrected murine models of X-linked hyper-IgM syndrome (XHIM), but it subsequently led to the development of lymphomas in the mice from constitutive rather than regulated CD40 ligand expression.[29] Bruton tyrosine kinase (BTK), defective in X-linked agammaglobulinemia, similarly may need to be expressed specifically at certain stages of B-lymphocyte development for safety and efficacy. A long list of other loci involved in PIDs may similarly be best approached by gene correction, including SLAM, XIAP, JAK3, FOXP3, IL-10, IL7Ralpha, TACI, CTLA4, etc. While current gene correction techniques are likely below the frequency of efficiency needed to yield clinical benefits for most disorders, this field is moving at a lightning pace with new advances being reported weekly in the activity and specificity of a whole host of nucleases and other genome editing tools.

USE OF PLURIPOTENT STEM CELLS AS A SOURCE OF HSC FOR GENE THERAPY OF PID

The establishment of human pluripotent stem cells (hPSCs), initially as human embryonic stem cells (hESCs) and subsequently as induced pluripotent stem cells (iPSCs), has brought the promise of novel models to study human diseases (“disease in a dish”) and to provide renewable sources of patient-compatible cells for cellular therapies. The essentially unlimited ability to expand hPSCs and their capacity to produce any of the cell-types in the body has led to investigations to harness them for regenerative medicines. In the treatment of PIDs, hPSC could provide an ideal target to produce autologous HSCs with precise gene correction. While techniques have been developed to perform the gene modification strategies that may be sufficiently robust for clinical applications, the major roadblock is the current inability to produce clinically-relevant numbers of transplantable HSCs from hPSCs. The HSC is a relatively evanescent state and it has not been possible to “freeze” differentiation from hPSC at that state, although it has been possible to proceed right through the HSC stage to produce relatively pure populations of individual mature blood cells. As with gene correction, the pace of scientific progress with these phenomenal cells is proceeding rapidly, and it is likely that new sources of gene-corrected autologous HSCs will be available for clinical transplants in the near future.

GENE THERAPY FOR PID INVOLVING SERUM PROTEIN DEFICIENCIES

Whereas many of the severe PIDs are due to blood cell defects and thus responsive to HSCT, others result from deficiencies of serum proteins, such as complement components. Here, gene therapy may be approached, as is being done for hemophilia, by direct *in vivo* administration of a gene-containing vector that can permanently insert the gene into target tissues–such as the liver, skeletal muscle, or endothelium–that can serve as protein sources. *In vivo* gene delivery is being studied using retroviral and lentiviral vectors or adeno-associated virus (AAV) vectors, and methods for *in vivo* gene correction using site-specific nucleases and homologous donor sequences are also under development. To this end, Crystal and co-workers have recently reported studies in a murine model of hereditary angioedema using an AAV vector to express the *SERPING1* gene encoding the C1 inhibitor.[30]

ADVANCING GENE THERAPY FOR PID FROM EXPERIMENTAL TO STANDARD OF CARE

Initial trials of gene therapy for PIDs were all performed at tertiary-level academic medical centers, often with federal or disease-foundation research grant funding, and they tested initial hypotheses concerning safety and efficacy. These centers may continue to perform early phase clinical trials for specific disorders, especially where there is local expertise on the disease being studied. However, there is an ongoing transition to industry-sponsored trials focused on drug product development, as the effective vectors and the gene-modified stem cell products they compose are advanced to licensure and marketing as pharmaceutical agents. A common model used by these new gene therapy companies is one of central processing, at one or a few high-grade commercial GMP facilities. The autologous patient cells are procured at their home institution (by leukopheresis or bone marrow harvest), shipped to the central processing site for genetic manipulation and cryopreservation, and then returned to be administered locally. Currently, gene therapy transplants have been performed at a limited number of clinical trial sites due to the cell processing expertise needed; presumably under the commercial model, it may be done at any suitable hematopoietic stem cell transplant center, with the company selling the processed cell product, like a medical device or an unrelated stem cell product. Alternatively, self-contained cell processing and manipulation devices are being developed that could allow gene modification of stem cells to be done at individual institutions, without requiring highly trained staff or high-grade GMP facilities.

Major issues remain to be determined about cost and reimbursement for gene therapy. Effective gene therapy for the severe diseases being approached would be expected to lead to large lifetime savings in medical costs. The one-time price can be compared to the costs faced by the patient encumbered by the progressive nature of the underlying disease, to the costs for long-term protein-based therapies and possibly even to the costs for allogeneic transplantation. However, those one-time costs will be expensive for clinical gene therapy transplants using pharmaceutical-grade vectors and commercial-level cell processing with the attendant quality control. Thus, a single large expenditure for gene therapy may eventually be cost-effective compared to ongoing medical costs; however, the method of financing the large up-front charge, at least in the United States with its multiple insurance companies, remains to be determined.

CONCLUSION

In the past decades, gene therapy for PIDs has gone from a dream for the future to a clinical reality. Gene therapy for ADA-deficiency SCID has been safe and effective in most treated patients. Although initial trials of gene therapy for XSCID, CGD and WAS showed efficacy, they were marred by an unacceptably high rate of complications from genotoxicity associated with gamma-retroviral vectors. New vectors are now in clinical trials and are showing excellent safety profiles and higher efficacy in clinical trials for ADA-deficient SCID, XSCID, WAS, and CGD as well as other non-PID disorders. Approaches to direct gene correction are being developed that may broaden indications for PIDs that may be treated by gene therapy (*e.g.*, XHIM, XLA, etc.). Thus, the continued efforts of scientists and physicians to develop gene therapy are leading to a new therapeutic modality, ideally to permanently and safely cure these diseases.

ON THE HORIZON

- Continued expansion of the genetic types of PID being treated by gene addition using integrating vectors (*e.g.* other forms of SCID, other forms of CGD, LAD, HLH, etc.).
- Development of safe and effective marrow conditioning regimens that are not chemotherapy-based (*e.g.* monoclonal antibodies)
- Application of gene editing to broader spectrum of PIDs: *e.g.* X-linked agammaglobulinemia, X-linked Hyper-IgM syndrome (CD40 ligand deficiency), RAG1-deficient SCID, Gain of Function STAT1, IPEX).
- Understand the molecular pathogenesis of more PID (*e.g.* CVID, autoimmune) and develop effective gene therapy approaches. (Some may require more than just HSC correction, *e.g.* ataxia telangiectasia).
- Production of autologous HSC by cellular reprogramming coupled to effective expansion of HSC

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Stewart FM, Crittenden RB, Lowry PA, et al. Long-term engraftment of normal and post-5-fluorouracil murine marrow into normal nonmyeloablated mice. Blood 1993;81:2566–71.
2. Blaese RM, Culver KW, Miller AD, et al. T lymphocyte-directed gene therapy for ADA- SCID: initial trial results after 4 years. Science 1995;270:475–80.
3. Candotti F, Shaw KL, Muul L, et al. Gene therapy for adenosine deaminase – deficient severe combined immune deficiency : clinical comparison of retroviral vectors and treatment plans. Blood 2012;120:3635–46.
4. Cicalese MP, Ferrua F, Castagnaro L, et al. Update on the safety and efficacy of retroviral gene therapy for immunodeficiency due to adenosine deaminase deficiency. Blood 2016;128:45–54.
5. Shaw KL, Sokolic R, Davila A, et al. Phase II clinical trial of gene therapy for adenosine deaminase-deficient severe combined immune deficiency (ADA-SCID). Mol Ther 2014;22(Suppl. 1):S107.
6. Candotti F, Johnston JA, Puck JM, et al. Retroviral-mediated gene correction for X-linked severe combined immunodeficiency. Blood 1996;87:3097–102. Available from: http://www.ncbi.nlm.nih.gov/pubmed/8605322.
7. Cavazzana-Calvo M, Hacein-Bey S, de Saint Basile G, et al. Role of interleukin-2 (IL-2), IL-7, and IL-15 in natural killer cell differentiation from cord blood hematopoietic progenitor cells and from gamma c transduced severe combined immunodeficiency X1 bone marrow cells. Blood 1996;88:3901–9.
8. Hacein-Bey-Abina S, Hauer J, Lim A, et al. Efficacy of gene therapy for X-linked severe combined immunodeficiency. N Engl J Med 2010;363:355–64. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2957288&tool=pmcentrez&rendertype=abstract.
9. Gaspar HB, Cooray S, Gilmour KC, et al. Long-term persistence of a polyclonal T Cell pepertoire after gene therapy for X-linked severe combined immunodeficiency. Sci Transl Med 2011;3:97ra79.
10. Hacein-Bey-Abina S, Pai S-Y, Gaspar HB, et al. A modified γ-retrovirus vector for X-linked severe combined immunodeficiency. N Engl J Med 2014;371:1407–17. Available from: http://www.nejm.org/doi/abs/10.1056/NEJMoa1404588.
11. De Ravin SS, Wu X, Moir S, et al. Lentiviral hematopoietic stem cell gene therapy for X-linked severe combined immunodeficiency. Sci Transl Med 2016;8:335ra57. Available from: http://stm.sciencemag.org/content/8/335/335ra57.figures-only.
12. Bauer TR Jr, Hickstein DD. Gene therapy for leukocyte adhesion deficiency. Curr Opin Mol Ther 2000;2:383–8.
13. Güngör T, Teira P, Slatter M, et al. Reduced-intensity conditioning and HLA-matched haemopoietic stem-cell transplantation in patients with chronic granulomatous disease: a prospective multicentre study. Lancet 2014;383:436–48. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24161820.
14. Santilli G, Almarza E, Brendel C, et al. Biochemical correction of X-CGD by a novel chimeric promoter regulating high levels of transgene expression in myeloid cells. Mol Ther 2011;19:122–32.
15. De Ravin SS, Reik A, Liu P-Q, et al. Targeted gene addition in human CD34+ hematopoietic cells for correction of X-linked chronic granulomatous disease. Nat Biotechnol. Nature Publishing Group 2016;(February):1–8. Available from: http://www.nature.com/doifinder/10.1038/nbt.3513.
16. Becker-Herman S, Meyer-Bahlburg A, Schwartz MA, et al. WASp-deficient B cells play a critical, cell-intrinsic role in triggering autoimmunity. J Exp Med 2011;208:2033–42.
17. Boztug K, Schmidt M, Schwarzer A, et al. Stem-cell gene therapy for the Wiskott-Aldrich syndrome. N Engl J Med 2010;363:1918–27.
18. Charrier S, Dupré L, Scaramuzza S, et al. Lentiviral vectors targeting WASp expression to hematopoietic cells, efficiently transduce and correct cells from WAS patients. Gene Ther 2007;14:415–28. Available from: http://www.nature.com/doifinder/10.1038/sj.gt.3302863%5Cnpapers3://publication/doi/10.1038/sj.gt.3302863.
19. Scaramuzza S, Biasco L, Ripamonti A, et al. Preclinical safety and efficacy of human CD34+ Cells transduced with lentiviral vector for the treatment of Wiskott-Aldrich syndrome. Mol Ther 2013;21:175–84. Available from: http://dx.doi.org/10.1038/mt.2012.23/nature06264.
20. Hacein-Bey Abina S, Gaspar HB, Blondeau J, et al. Outcomes following gene therapy in patients with severe Wiskott-Aldrich syndrome. JAMA 2015;313:1550–63. Available from: http://jama.jamanetwork.com/article.aspx?doi=10.1001/jama.2015.3253.
21. Aiuti A, Biasco L, Scaramuzza S, et al. Lentiviral hematopoietic stem cell gene therapy in patients with Wiskott-Aldrich syndrome. Science 2013;341:1233151. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23845947.
22. Pala F, Morbach H, Castiello MC, et al. Lentiviral-mediated gene therapy restores B cell tolerance in Wiskott-Aldrich syndrome patients. J Clin Invest 2015;125:3941–51. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4607131&tool=pmcentrez&rendertype=abstract.
23. Astrakhan A, Sather BD, Ryu BY, et al. Ubiquitous high-level gene expression in hematopoietic lineages provides effective lentiviral gene therapy of murine Wiskott-Aldrich syndrome. Blood 2012;119:4395–407.
24. Ng YY, Baert MRM, Pike-Overzet K, et al. Correction of B-cell development in Btk-deficient mice using lentiviral vectors with codon-optimized human BTK. Leukemia 2010;24:1617–30.
25. Kerns HM, Ryu BY, Stirling BV, et al. B cell-specific lentiviral gene therapy leads to sustained B-cell functional recovery in a murine model of X-linked agammaglobulinemia. Blood 2010;115:2146–55.
26. Hagin D, Burroughs L, Torgerson TR. Hematopoietic Stem Cell Transplant for Immune Deficiency and Immune Dysregulation Disorders. Immunol Allergy Clin North Am 2015;35:695–711.
27. Tebas P, Stein D, Tang WW, et al. Gene editing of CCR5 in autologous CD4 T cells of persons infected with HIV. N Engl J Med 2014;370:901–10. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4084652&tool=pmcentrez&rendertype=abstract.
28. Porteus MH, Baltimore D. Gene Targeting in Human Cells. Science 2003;300(May):75390.
29. Brown MP, Topham DJ, Sangster MY, et al. Thymic lymphoproliferative disease after successful correction of CD40 ligand deficiency by gene transfer in mice. Nat Med 1998;4:1253–60.
30. Pagovich OE, Qui T, Whaley AS, et al. One-time gene therapy for hereditary angioedema. Mol Ther 2016;24(Suppl. 1):S298.

MULTIPLE-CHOICE QUESTIONS

3-5 multiple choice review questions - these questions should have explanations for the correct answer.

1. Which of the following vector types is effective for high efficiency permanent gene addition to hematopoietic stem cells:
   A. Adeno-associated virus (AAV)
   B. Gamma-retroviral
   C. Plasmid – by transfection or electroporation
   D. Lentiviral
   E. Adenovirus
   1. A, B, C
   2. B and D
   3. A-D
   4. All are correct

2. Which of the following primary immune deficiencies have shown good clinical responses to gene therapy:
   A. Wiskott-Aldrich Syndrome (WAS)
   B. X-linked Severe Combined Immune Deficiency (XSCID)
   C. X-linked Chronic Granulomatous Disease (XCGD)
   D. X-linked Agammaglobulinemia (XLA)
   E. Common Variable Immune Deficiency (CVID)
   1. A, B, C
   2. B and D
   3. A-D
   4. All are correct

3. Which are true about gene editing:
   1. Gene editing may be preferred to gene addition for primary immune deficiencies where the relevant gene needs to be expressed under precise regulation.
   2. Cells can repair double-stranded DNA breaks by two predominate methods: Non Homologous End Joining (NHEJ), which tends to introduce insertions and deletions (indels) at the chromosome repair junction, and Homologous Recombination (HR), which can copy genetic information from a homologous donor sequence into the repair junction.
   3. The first human clinical trial using gene editing disrupted the gene encoding an essential HIV-1 co-receptor protein in autologous T cells.
   4. The introduction of a double-stranded DNA break increases the frequency of homologous recombination for site-specific gene editing.
   5. Gene editing methods can site-specifically disrupt, correct or introduce a gene.
      1. A, B, C
      2. B and D
      3. A-D
      4. All are correct

Glucocorticoids

Anthony J. Frew

Glucocorticoids (GCs) are among the most commonly prescribed drugs and are used for a wide range of medical conditions.[1] More than 60 years after their introduction into clinical practice, they remain the most important and most frequently employed class of antiinflammatory drugs, and their use continues to increase, with about 10 million new prescriptions issued for oral GCs each year in the United States.[2] Community survey data suggest that 0.5% of the general population and 1.75% of women aged over 55 years are taking oral GCs.[3,4] About 56–68% of patients with rheumatoid arthritis (RA) are treated more or less continuously with GCs.[1] Although GCs are relatively inexpensive, total market volume is valued at ≈US$10 billion per year.[2]

GCs are widely used because they are the most effective (and cost-effective) antiinflammatory and immunomodulatory drugs available. However, GCs can cause serious adverse effects, especially when used incorrectly.

MECHANISMS OF ACTION

The way that GCs are used in different clinical conditions is essentially empirical, as there is only limited evidence to support current practice in specific clinical settings.[1] In general, GC dosages are increased in parallel with the clinical activity and severity of the disease under treatment. The rationale for this approach is that higher dosages increase GC receptor saturation in a dose-dependent manner (Table 86.1), which intensifies the therapeutically relevant, genomic actions of GCs. Moreover, with increasing dosages, additional and qualitatively different, nonspecific, nongenomic actions of GCs come into play (see Table 86.1).

KEY CONCEPTS

Characteristics Applying to Genomic Actions

- Physiologically relevant
- Therapeutically effective at all dosages, even very small ones (low dose therapy).
- Slow; significant changes in the regulator protein concentrations are not seen within less than 30 minutes because of the time required for cytosolic glucocorticoid receptor (cGCR) activation/translocation, transcription, and translation effects.
- The glucocorticoid (GC)–induced synthesis of regulator proteins can be prevented by inhibitors of transcription (*e.g.,* actinomycin D) or translation (*e.g.,* cycloheximide).

Genomic Actions of Glucocorticoids

The antiinflammatory and immunomodulatory effects of GCs are mainly mediated by genomic mechanisms (Figs. 86.1 and 86.2). Their lipophilic structure and low molecular mass allow GCs to pass easily through the cell membrane and bind to cytosolic glucocorticoid receptors (cGCRs). Ultimately this either induces the synthesis of regulatory proteins (“transactivation”) or inhibits their synthesis (“transrepression”).[5] About 10–100 genes per cell are directly regulated by GCs, but many other genes are regulated indirectly through interaction with transcription factors and coactivators (see below).[6] It has been estimated that GCs influence the transcription of approximately 1% of the entire genome.

Structure of the Cytosolic Glucocorticoid Receptor

The nonactivated cGCR (cGCRα) is a 94-kilodalton (kDa) protein held in the cytoplasm as a multiprotein complex, consisting of several heat shock proteins (hsps), including hsp90, hsp70, hsp56, and hsp40 (chaperones) (Fig. 86.3). The cGCR interacts with immunophilins, p23, and several kinases of the mitogen-activated protein kinase (MAPK) signaling system, including Src, which also act as molecular (co)chaperones (see Figs. 86.1 and 86.3).[1,7] The general function of molecular (co)chaperones is to bind and stabilize proteins at intermediate stages of folding, assembly, translocation, and degradation. With regard to cGCR, they also regulate cellular signaling, which includes *(i)* stabilizing a high-affinity conformational state of cGCR; *(ii)* opening the GC-binding cleft to be accessed by GCs; and *(iii)* stabilizing the binding of the GCR to the promoter.[1]

The first step in assembling the multiprotein cytosolic complex is adenosine triphosphate (ATP)–dependent and hsp40(YDJ-1)–dependent formation of a cGCR–hsp70 complex that primes the receptor for subsequent ATP-dependent activation by hsp90, Hop, and p23.[8] The GCR consists of different domains with distinct functions: an N-terminal domain; a DNA-binding domain (DBD); and a ligand-binding domain (LBD) (see Fig. 86.3). The N-terminal domain serves transactivation functions, especially within the “τ1” region. A zinc finger motif, a sequence common to many DNA-interacting proteins, is found twice within the DBD. The LBD consists of 12 α helices, several of which help form a hydrophobic ligand-binding pocket.[7] The cGCR contains another major transactivation region (“τ2”) that can interact with the above-mentioned cofactors (see Fig. 86.3). Following

TABLE 86.1 **Current Knowledge on the Relationship Between Clinical Glucocorticoid (GC) Dosing and Cellular GC Actions**

Terminology (Mg Prednisone Equivalent Per Day)	Clinical Application	Genomic Actions (Receptor Saturation)	Unspecific Nongenomic Actions	Cytosolic Glucocorticoid Receptor (cGCR)–Mediated Nongenomic Actions
Low dose (≤7.5)	Maintenance therapy for many rheumatic diseases	+ (<50%)	–	?
Medium dose (>7.5 – ≤30)	Initially given in primary chronic rheumatic diseases	++ (>50% – <100%)	(+)	(+)
High dose (>30 – ≤100)	Initially given in subacute rheumatic diseases	++(+) (almost 100%)	+	+
Very high dose (>100 mg)	Initially given in acute and/or potentially life-threatening exacerbations of rheumatic diseases	+++ ([almost] 100%)	++	+(+?)
Pulse therapy (≥250 mg for 1 or a few days)	Particularly severe and/or potentially life-threatening forms of rheumatic diseases	+++ (100%)	+++	+(++?)

From Buttgereit F, Straub RH, Wehling M, Burmester GR. Glucocorticoids in the treatment of rheumatic diseases. An update on mechanisms of action. Arthritis Rheum 2004;50: 3408–17, with permission.)

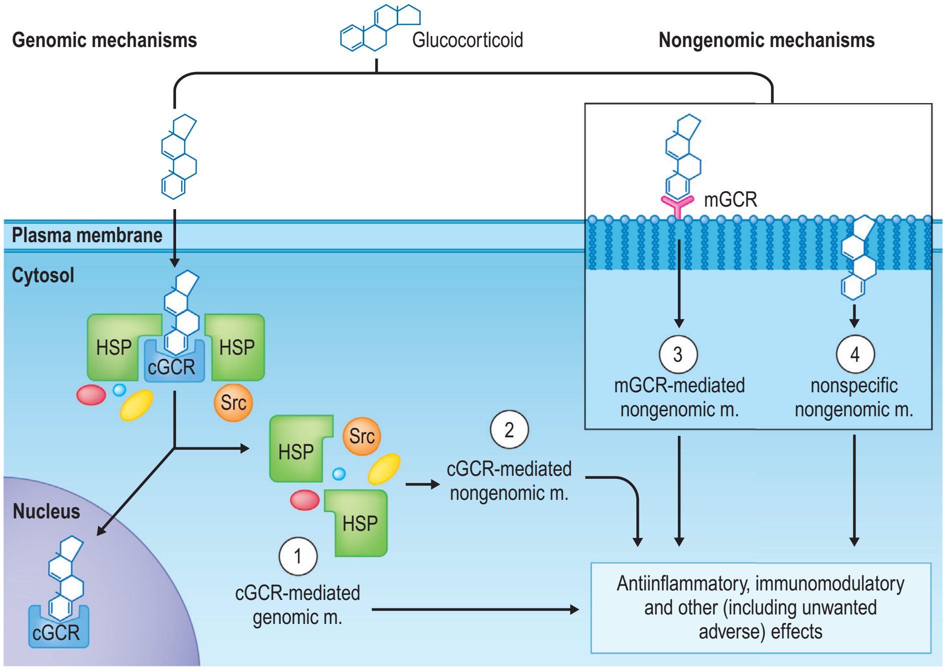

Genomic mechanisms
Glucocorticoid
Nongenomic mechanisms
mGCR
Plasma membrane
Cytosol
HSP
cGCR
HSP
Src
HSP
Src
HSP
2
cGCR-mediated nongenomic m.
3
mGCR-mediated nongenomic m.
4
nonspecific nongenomic m.
Nucleus
cGCR
1
cGCR-mediated genomic m.
Antiinflammatory, immunomodulatory and other (including unwanted adverse) effects

FIG 86.1 **Mechanisms of the Cellular Action of Glucocorticoids (GCs).** As lipophilic substances, GCs pass very easily through the cell membrane into the cell, where they bind to ubiquitously expressed cytosolic glucocorticoid receptors (cGCRs). This is followed by either the classic cGCR-mediated genomic effects (1) or by cGCR-mediated nongenomic effects (2). Moreover, the GC is very likely to interact with cell membranes, either specifically via membrane-bound glucocorticoid receptors (mGCRs) (3) or via nonspecific interactions with cell membranes (4). (From Buttgereit F, Straub RH, Wehling M, Burmester GR. Glucocorticoids in the treatment of rheumatic diseases. An update on mechanisms of action. Arthritis Rheum 2004;50:3408–17.)

GC/cGCR binding, the hsp90 molecules and other molecular chaperones are rapidly shed. This allows translocation into the cell nucleus, where the GC/cGCR complex binds as a homodimer to consensus palindromic DNA sites (glucocorticoid-responsive elements [GREs]).[5]

Translocation Into the Nucleus

Nuclear translocation of the GC/cGCR complex occurs within 20 minutes. This may be caused by hormone-directed recruitment of FKBP52 and dynein to the GCR.[9] Depending on the target gene, transcription is then either activated (transactivation via positive GRE) or inhibited (transrepression via negative GRE).

Interactions With Transcription Factors

As well as the interactions of GC/cGCR complexes with GRE, a further important genomic mechanism of GC action is the interaction of activated cGCR monomers with transcription factors. Accordingly, although the GC/cGCR complex does

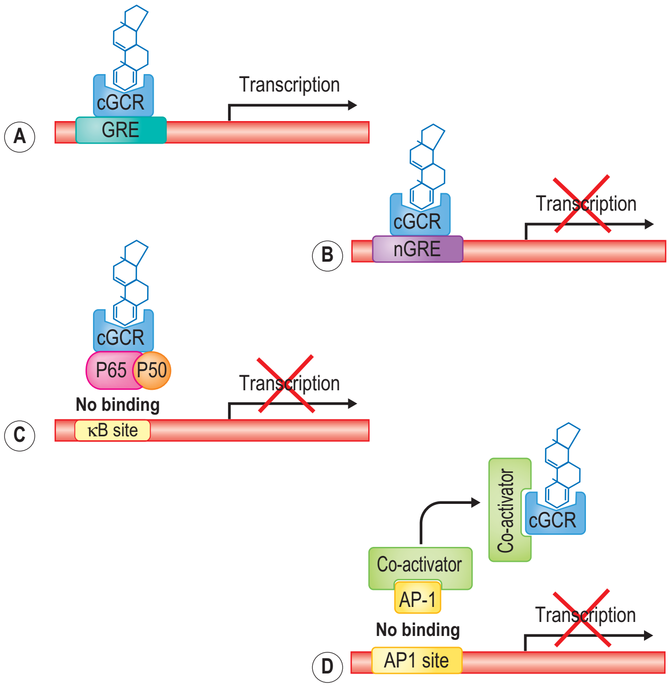


FIG 86.2 **Genomic Mechanisms of Glucocorticoids (GCs).** This figure illustrates the different mechanisms by which the activated glucocorticoid receptor (GCR) complex leads to the induction or to the inhibition of transcription and finally translation/synthesis of specific regulator proteins. Details are given in the text. (From Buttgereit F, Straub RH, Wehling M, Burmester GR. Glucocorticoids in the treatment of rheumatic diseases. An update on mechanisms of action. Arthritis Rheum 2004;50:3408–17.)

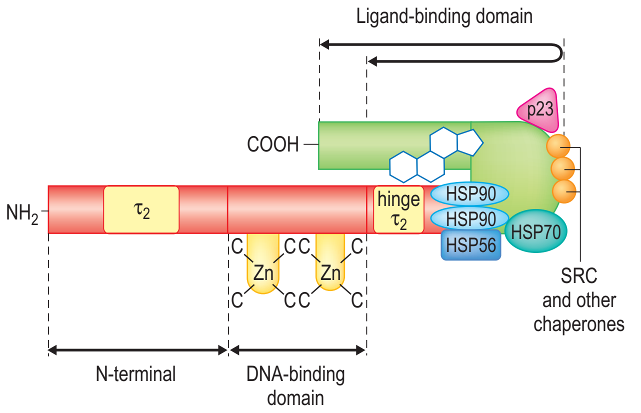


FIG 86.3 **Structure of the Cytosolic Glucocorticoid Receptor (cGCR).** The nonactivated (unligated) cGCR is a 94-kilodalton (kDa) protein retained in the cytoplasm as a multiprotein complex consisting of several heat shock proteins (hsps), including hsp90, hsp70, hsp56, and hsp40 (chaperones). Furthermore, the cGCR interacts with immunophilins, p23, and several kinases of the mitogen-activated protein kinase (MAPK) signaling system, including Src, which also act as molecular (co)chaperones. An important function of molecular (co)chaperones is to stabilize a specific conformational state of the GC which binds ligand with high affinity (see text). The receptor protein itself consists of different domains: an N-terminal, a DNA-binding domain (DBD), and a ligand-binding domain (LBD). The N-terminal harbors transactivation functions, especially within the so-called τ1 region. Another major transactivation region is τ2, which can interact with the above-mentioned cofactors. (From Buttgereit F, Straub RH, Wehling M, Burmester GR. Glucocorticoids in the treatment of rheumatic diseases. An update on mechanisms of action. Arthritis Rheum 2004;50:3408–17.)

not inhibit their synthesis, it modulates the activity of AP-1 (activator protein-1), nuclear factor (NF)-κB (nuclear factor-κB) and NFAT (nuclear factor for activated T cells). This leads to inhibition of nuclear translocation and/or function of these transcription factors and, thus, to inhibition of expression of many immunoregulatory and inflammatory cytokines. Possible mechanisms include[5]:

- Synthesis of IκB (a specific inhibitor of NF-κB) induced through GC/cGCR complex–GRE interaction (see Fig. 86.2A)
- Protein–protein interaction of the GC/cGCR complex with transcription factors through binding to their subunits (see Fig. 86.2C), which prevents their DNA binding
- Competition for nuclear coactivators between the GC/cGCR complex and transcription factors (see Fig. 86.2D)

Inhibition of transcription factor function and the resultant inhibition of protein expression are referred to as *transrepression.* Numerous genes are regulated by this mechanism. Many adverse effects of GCs are caused by *transactivation* (induced synthesis of regulator proteins), whereas most antiinflammatory effects are mediated by transrepression (inhibited synthesis of regulator proteins). This differential molecular regulation underlies current drug-discovery programs aimed at developing dissociated cGCR-ligands.[2]

The cGCRβ Isoform

The cGCRβ isoform is an alternative splicing variant of cGCRα that does not bind GCs or activate gene expression. This isoform is thought to function as a negative inhibitor of cGCRα, and it may play a role in the clinical phenomenon of GC resistance. cGCRβ lacks the GC-binding domain that is needed for activation, and as it does not undergo ligand-dependent downregulation, it has a longer half-life than the active form (cGCRα). It is thought that the likely mechanism of the dominant negative activity of cGCRβ is through the formation of inactive heterodimers with cGCRα.[10]

Posttranscriptional and Posttranslational Mechanisms

GCs also act through posttranscriptional and posttranslational mechanisms, including reduction of the half-life of cytokine messenger RNA (mRNA) and downregulation of the GCR, via reduced mRNA levels and reduced stability of the GCR protein.

KEY CONCEPTS

Glucocorticoid (GC) Effects on Immune Cells

- Inhibit leukocyte traffic and access of leukocytes to the site of inflammation
- Interfere with functions of leukocytes, fibroblasts, and endothelial cells
- Suppress the production and actions of humoral factors involved in the inflammatory process

Glucocorticoid Receptor Resistance

Several mechanisms may explain the clinical finding of GCR resistance, including alterations in the number, binding affinity, or phosphorylation status of GCR. Other possible explanations are polymorphic changes and/or overexpression of (co)chaperones, increased expression of inflammatory transcription factors, overexpression of the GCRβ isoform, the multidrug resistance pump, and altered membrane-bound (mGCR) expression.[11]

Nongenomic Actions of Glucocorticoids

Some regulatory effects of GCs occur within seconds or minutes. These are too rapid to result from genomic actions and must, therefore, be caused by nongenomic mechanisms of action. Three different rapid nongenomic actions of GCs have been described.[1,12,13]

cGCR-Mediated Nongenomic Actions

Dexamethasone can rapidly inhibit epidermal growth factor-stimulated cPLA2 (cytosolic PLA2) activation with subsequent arachidonic acid release.[12] This effect is thought to be mediated by occupation of cGCR, rather than changes in gene transcription, as the observed effect is RU486 sensitive (*i.e.,* GCR dependent) but actinomycin insensitive (*i.e.,* transcription independent). Chaperones or (co)chaperones of the multiprotein complex may act as signaling components to mediate this effect. Following GC binding, the cGCR is released from this complex to mediate classic genomic actions. However, there is also a rapid release of Src and other (co)chaperones of the multiprotein complex, which may cause rapid inhibition of arachidonic acid release. Similarly, dexamethasone has been reported to have cardiovascular protective effects, which are neither genomic (because they occurred too quickly and were actinomycin insensitive) nor nonspecific–nongenomic effects (because they occurred at a very low concentration [100 nM]).[13] These may involve binding of GCs to cGCRs, leading to nontranscriptional activation of phosphatidylinositol 3-kinase, protein kinase Akt, and endothelial NO synthase.

Nonspecific Nongenomic Actions

GCs are sometimes administered at very high doses. Systemic daily dosages >100 mg prednisone equivalent are regarded as "very high dose." "Pulse therapy" is the daily administration of ≥250 mg prednisone equivalent for 1 day or a few consecutive days[14] (see Table 86.1). At a daily dose of 100 mg prednisone equivalent, almost all cGCRs are completely saturated, which means that specificity (*i.e.,* the exclusivity of receptor-mediated effects) is lost. Nonspecific nongenomic actions occur in the form of physicochemical interactions with biological membranes, which probably contribute to the therapeutic effect.[1] Intercalation of GC molecules into cell membranes is thought to alter cell function by influencing cation transport and increasing mitochondrial proton leak. The resulting inhibition of calcium and sodium cycling across the plasma membrane of immune cells is thought to contribute to rapid immunosuppression and to reduced inflammation.

Such high GC doses are only used in a few clinical specialties, and this practice has been criticized by some endocrinologists and pharmacologists. Unfortunately, there are no randomized controlled trials of high-dose GC therapy, but it is often used with clinical success in acute exacerbations of life-threatening diseases and various clinical conditions resistant to other therapies. For example, pulsed intravenous (IV) methylprednisolone is effective in the treatment of systemic lupus erythematosus (SLE) and causes rapid immunosuppression in patients with organ- and/or life-threatening manifestations of SLE. However, the standard regime (1 g/day for 3 consecutive days) is associated with significant risks of infection, and lower doses may be just as effective.[15]

High-dose GCs are also often used in immune thrombocytopenia associated with SLE, although comparative studies are lacking.[16] It has been calculated that in such situations, the concentrations achieved *in vivo* are high enough (around 10–5 mol/L) to cause immediate nonspecific nongenomic effects on immune cells.[1] Intraarticular injections also bring high concentrations of GCs into contact with inflammatory cells, although it is difficult to assess locally achieved concentrations because crystal suspensions are most often used.

Specific Nongenomic Actions

GCs can also induce specific nongenomic actions, mediated through membrane-bound glucocorticoid receptors (mGCRs). The existence and function of membrane-bound receptors have been demonstrated for various steroids (including mineralocorticoids, gonadal hormones, vitamin D, and thyroid hormones).[1,17] Small numbers of mGCRs can be demonstrated by immunofluorescence on human peripheral blood mononuclear cells (PBMCs; monocytes and B cells) obtained from healthy controls.[17] The monoclonal antibody (mAb) used to detect mGCRs also recognizes cGCRs, suggesting that mGCRs are probably variants of cGCRs produced by differential splicing or promoter switching. Immunostimulation with lipopolysaccharide increases the percentage of mGCR$^+$ monocytes, and this can be prevented by inhibiting the secretory pathway with brefeldin A. This suggests that mGCRs are actively upregulated and transported through the cell after immunostimulation. These *in vitro* findings are consistent with observations that the frequency of mGCR$^+$ monocytes is increased in patients with rheumatic disorders and is positively correlated with disease activity in RA.[17] It remains unclear whether mGCRs are involved in pathogenesis. Alternatively, and perhaps more likely, they may cause negative feedback regulation.

GLUCOCORTICOID EFFECTS ON IMMUNE CELLS

Through the above mechanisms, GCs mediate a wide range of antiinflammatory and immunomodulatory effects, with virtually all primary and secondary immune cells affected to some extent (Table 86.2).[18]

KEY CONCEPTS

Definition of Conventional Terms for Glucocorticoid (GC) Dosages

Low dose	≤7.5 mg prednisone equivalent per day
Medium dose	>7.5 mg, but ≤30 mg prednisone equivalent per day
High dose	>30 mg, but ≤100 mg prednisone equivalent per day
Very high dose	> 100 mg prednisone equivalent per day
Pulse therapy	≥ 250 mg prednisone equivalent per day for 1 or a few days

TABLE 86.2 Important Effects of Glucocorticoids (GCs) on Primary and Secondary Immune Cells

Monocytes/Macrophages

↓ Number of circulating cells (↓ myelopoiesis, ↓ release)
↓ Expression of major histocompatibility complex (MHC) class II molecules and Fc receptors
↓ Synthesis of proinflammatory cytokines (*e.g.,* interleukin [IL]-2, IL-6, tumor necrosis factor [TNF]-α) and prostaglandins

T Cells

↓ Number of circulating cells (redistribution effects)
↓ Production and action of IL-2 (most important)

Granulocytes

↓ Number of eosinophil and basophil granulocytes
↑ Number of circulating neutrophils

Endothelial Cells

↓ Vessel permeability
↓ Expression of adhesion molecules
↓ Production of IL-1 and prostaglandins

Fibroblasts

↓ Proliferation
↓ Production of fibronectin and prostaglandins

(From Buttgereit F, Saag K, Cutolo M, et al. The molecular basis for the effectiveness, toxicity, and resistance to glucocorticoids: focus on the treatment of rheumatoid arthritis. Scand J Rheumatol 2005; 34: 14–21, with permission.)

The Role of Endogenous Glucocorticoids in Inflammatory Arthritis

Exogenous (therapeutic) and endogenous (physiological) GCs differ in several respects. The most important differences are in their relative mineralocorticoid and GC (antiinflammatory) activities. Exogenous and endogenous GCs also differ in their pharmacological characteristics, such as plasma kinetics, metabolism, biological half-life, lipophilicity, drug–receptor interactions, and nongenomic potencies.[19]

The actions of exogenous GCs as described above are well established. In contrast, we know relatively little about the role of endogenous GCs in arthritis. Although GC actions on target tissues are thought to be determined by GC plasma concentrations and the tissue-specific density of GCRs, it seems that endogenous GCs are subject to extensive prereceptor metabolism. Within target cells or tissues, GC action depends not only on plasma hormone levels, receptor expression, and receptor-effector coupling but also on local GC metabolism. Specifically, 11β-hydroxysteroid dehydrogenases appear to govern access of GCs to their cognate receptors by changing the balance between active and inactive GCs within the cell (reviewed in [19]). Thus the predominant reductase activity of 11β-hydroxysteroid dehydrogenase type 1 (11βHSD1) catalyzes the formation of bioactive cortisol from inactive cortisone (in humans) and corticosterone from 11-dehydrocorticosterone (in rodents). This nicotinamide adenine dinucleotide phosphate (NADPH)–dependent enzyme is present in many tissues and usually increases the intracellular availability of active GCs. In contrast, 11βHSD2 only possesses dehydrogenase activity: It inactivates active GCs and therefore decreases the intracellular concentration of bioactive GCs.

Proinflammatory cytokines, such as interleukin-1β (IL-1β) and tumor necrosis factor-α (TNF-α), stimulate 11βHSD1 and downregulate 11βHSD2 expression. Hence, specific proinflammatory cytokines can modulate local intracellular GC metabolism, which may affect their own proinflammatory effects. More recently, substantial GC metabolism has been shown in joints, since TNF-α and IL-1β induce 11βHSD1 activity in primary cultures of synovial fibroblasts isolated from synovial tissue biopsy specimens from patients with RA.[20]

In a rodent model of immune-mediated arthritis, targeted disruption of GC signaling in osteoblasts attenuates joint inflammation and cartilage destruction.[21] These results suggest that under the control of endogenous GCs, osteoblasts modulate immune-mediated inflammatory responses and, as a consequence, inflammation-induced cartilage damage and bone integrity. These findings are supported by evidence suggesting that the effects of GCs follow a dose–response curve, with permissive or even stimulatory effects at physiological concentrations and suppressive effects at pharmacological concentrations.[20]

THERAPEUTIC USE

A wide range of GC molecules are available for clinical use: The common basic structure has been modified to improve their usefulness in various clinical applications (Fig. 86.4). Despite their widespread use, the designation of GC treatment regimens is often imprecise. Recommendations for a standardized nomenclature for GC therapy are summarized below.[14]

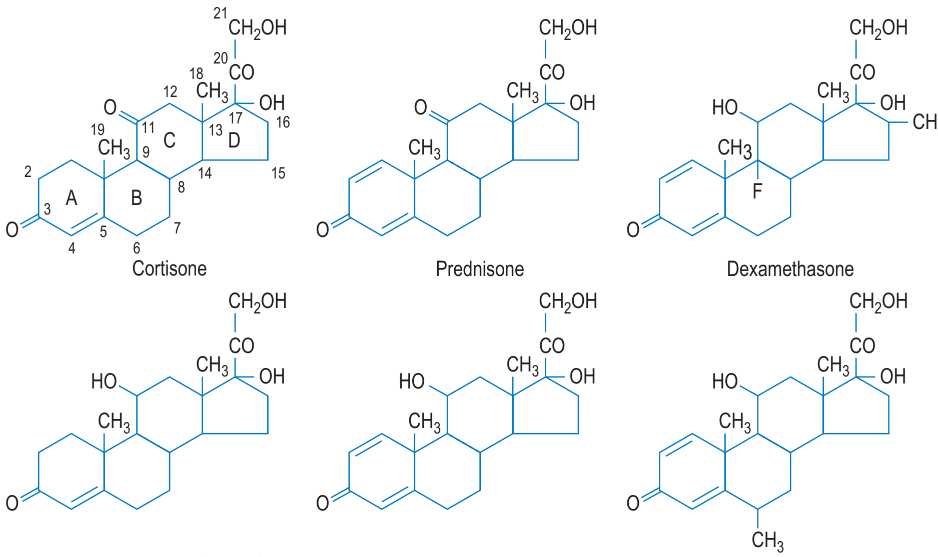
Terminology

Although the term *steroid* is widely used to describe this class of drugs, it is too broad, as it simply describes chemical compounds characterized by a common multiple-ring structure (including cholesterol, vitamin D, and sex hormones). Similarly, the terms *corticosteroid* or *corticoid* are not sufficiently precise, as the adrenal cortex synthesizes not only GCs but also mineralocorticoids and androgens. For these reasons, the terms *glucocorticoid* or *glucocorticosteroid* are preferred; however, *glucocorticoid* is the more widely used term.

When describing the use of GCs, it is necessary to define the drug, the dosage, the route of administration, and the timing, frequency, and duration of treatment.

Different GC drugs have different potencies, and they differ in their ability to produce the distinct therapeutic effects discussed above. Drug potencies are usually described by the equivalent dosages (relative potencies) to produce classic genomic effects (Table 86.3). These values have been used for decades, although the experimental and clinical evidence for their precise relative potency is weak. In practice, relative potencies are useful as a general therapeutic guideline in daily clinical practice, as long as they are not used dogmatically. It has, therefore, been suggested that we should continue to use relative potencies until more exact data are available and that doses of different GCs should be expressed in terms of "prednisone equivalent" (*i.e.,* doses of different GCs are expressed as equivalent to milligrams of prednisone [or prednisolone, as the potency of prednisone is equal to that of prednisolone]).

However, data indicate that the concept of equivalent dosages is only valid for doses less than 100 mg prednisone equivalent because nongenomic effects come into play at higher doses. The reason this is important is that the relative potencies of different GCs to produce these nongenomic effects are completely different



Cortisone

Prednisone

Dexamethasone

Hydrocortisone (cortisol)

Prednisolone

Methylprednisolone

FIG 86.4 Molecular Structures of Cortisone and Commonly Used Glucocorticoids (GCs). Carbon and ring nomenclature is indicated on cortisone molecule.

TABLE 86.3 Drug Potencies of Selected Glucocorticoids (GCs)

Cortisol	1 (per definition)
Prednyliden	3.5
Predniso(lo)ne	4.0
Methylprednisolone	5.0
Dexamethasone	25
Betamethasone	25

Drug potencies describe the potency of the respective drug to produce classical genomic (antiinflammatory) effects relative to cortisol. These potencies provide the basis for calculating equivalent dosages.

(Data from Lipworth BJ. Therapeutic implications of non-genomic glucocorticoid activity. Lancet 2000;356:87–9.)

from their classic genomic effects (see Table 86.3). For example, methylprednisolone is used for pulse therapy of exacerbations of immunologically mediated disorders. Prednisolone and methylprednisolone have similar genomic potencies, but when used for high-dose therapy, the nonspecific nongenomic effect of methylprednisolone is more than three times stronger. This may explain the apparent superiority of high-dose methylprednisolone. In contrast, betamethasone has very low nongenomic potency, which may be why this drug is considered less effective systemically, even though it has the same genomic potency as that of dexamethasone. In summary, the clinical usage of different GCs is clearly determined by the magnitude of their clinical efficacy, but another important factor in selecting which one to use is their nongenomic potency.

Glucocorticoid Treatment Regimens: General Aspects

For many decades, there has been confusion surrounding the terms used to describe dosage (very low, low, mild, mild to moderate, moderate, high, very high, ultrahigh, and mega) and by loose usage of such terms as “low-dose therapy,” “high-dose therapy,” and “pulse therapy.” A 2002 consensus statement has clarified this situation, partly to achieve scientific consistency and partly to recognize that GC actions are strongly dose dependent, both quantitatively and qualitatively. The standardized nomenclature for GC dosages and GC treatment regimens, as discussed below, is now recommended.[14]

Low Dose

Treatment with doses of up to 7.5 mg prednisone equivalent per day is considered low-dose GC therapy because these doses occupy less than 50% of the receptors. Such courses are frequently used for maintenance therapy and have relatively few adverse effects (*e.g.*, osteoporosis). As there may be relative hypocortisolism in chronic inflammatory conditions, such as RA and polymyalgia rheumatica, low-dose GCs act, in part, as replacement therapy for reduced adrenal GC production.

Medium Dose

GC doses of more than 7.5 mg but less than 30 mg prednisone equivalent per day are considered medium-dose therapy because they lead to a significantly higher receptor engagement between 50% and <100%. These doses are effective in modulating disease activity in various rheumatic diseases but may have considerable and dose-dependent adverse effects if given for longer periods.

High Dose

Treatment with doses of more than 30 mg and up to 100 mg of prednisone equivalent per day is considered high-dose therapy because these doses significantly increase receptor saturation, in a dose-dependent manner. At approximately 100 mg prednisone equivalent per day, receptor saturation is almost complete, and it is likely that genomic GC effects are fully exerted (see Table 86.1). High-dose therapy can be successfully given as initial treatment for subacute diseases, such as non-life-threatening exacerbations or visceral complications of RA or other connective tissue diseases, but cannot be administered over the long term because of the danger of severe adverse effects.

Very High Dose

Doses above 100 mg of prednisone equivalent per day are considered “very high.” At this level, there is virtually 100%

saturation of cytosolic receptors, so any further increase in dose may affect the pharmacodynamics (*e.g.,* receptor off-loading and reoccupancy), receptor synthesis and expression. At these doses, nongenomic effects may deliver additional therapeutic benefit, although it remains unclear whether these effects have any direct therapeutic relevance. Experimental data suggest that these differential effects come increasingly into play above ≈100 mg/day (see Table 86.1). Doses of >100 mg of prednisone equivalent per day are frequently (and successfully) given as initial treatment for acute or life-threatening exacerbations of connective tissue diseases, vasculitis, and RA. These doses cannot be administered over the long term because of their severe adverse effects.

Pulse Therapy

Pulse therapy involves administration of ≥250 mg prednisone equivalent per day (usually intravenously) for a short period (typically 1–5 days, rarely longer). At such high doses, the nongenomic potencies of GCs come into play. It is likely that these explain the success of very high doses and pulse therapy in acute exacerbations of immunologically mediated diseases. The immediate effects of very high doses may be additive to the genomic effects mediated by cGCRs. These additional effects may make a crucial contribution to the therapeutic effect by terminating acute exacerbations. Circumstances where very high doses or pulse therapy can be successful include acute episodes or particularly severe forms of rheumatic diseases, such as SLE, vasculitis, polymyositis, and RA (see below).

Alternate-Day Regimens

Alternate-day regimens were introduced for long-term oral GC therapy with the aim of minimizing undesirable adverse effects, such as suppression of the hypothalamo–pituitary–adrenal (HPA) axis. Instead of daily dosing, a single dose is administered every other morning, usually in a dose equivalent to or somewhat higher than twice the usual daily dose. The idea behind this regimen is to allow the HPA axis to remain active by exposing the body to exogenous GC on alternate days instead of suppressing it every day. This strategy only works if the HPA axis is still active, and, unfortunately, patients often experience breakthrough symptoms on the second day of treatment. Alternate-day regimens are used rarely these days except in patients with juvenile idiopathic arthritis, in whom alternate-day GC regimens cause less inhibition of body growth.

Glucocorticoid Withdrawal Regimens

Because of their significant adverse effects, GCs are usually reduced or stopped as soon as the disease is under control. This needs to be done carefully to avoid a relapse in disease activity and to permit recovery of adrenal function. There are no controlled comparative studies to support a specific regimen for weaning patients off GCs, as this process needs to be adjusted according to disease activity, dose/duration of therapy, and clinical response. When patients with RA are treated with up to 10 mg/day of prednisone, the daily dose can be reduced by 2.5 mg every month until 5 mg/day is reached. Thereafter, doses can often be reduced by 1 mg per month. When higher doses have been used, a reduction by 5 mg every 1–2 weeks down to 20 mg/day is often well tolerated, followed by further reductions of 1–2.5 mg/day every 2–3 weeks. Addition of immunomodulatory drugs, such as methotrexate and azathioprine, may allow further dose reductions. Short courses of GCs for some conditions, such as asthma, may be ceased without tapering if the total treatment course is less than 14 days in duration. With longer periods of treatment, the dose should be reduced gradually to allow the HPA axis to recover, but the rate of reduction is usually faster than for rheumatological conditions.

Glucocorticoids in Rheumatoid Arthritis: an Example

GCs are crucially important in the management of RA and are used in various dosages at different disease stages. RA is, therefore, a useful example to discuss GC therapy in more detail.

Low-Dose Maintenance Therapy

In early RA, GC in doses <10 mg are highly effective for relieving symptoms in patients with active arthritis. Many patients are functionally dependent on this low-dose therapy and continue it over the long term.[22] Improvement has been documented not only in all clinical parameters, including pain scales, joint scores, morning stiffness, and fatigue, but also in acute inflammatory markers, such as erythrocyte sedimentation rate (ESR) and C-reactive protein (CRP). After 6 months of therapy, the beneficial effects of GCs seem to diminish, but if therapy is then tapered and stopped, patients often experience an exacerbation of symptoms within a few months.

The disease-modifying properties of GCs were first described in 1995, in a 2-year trial of 7.5 mg prednisolone in patients who had RA of short or intermediate disease duration and were treated with NSAIDs (95%) and disease-modifying antirheumatic drugs (DMARDs) (71%).[23] Nowadays, GCs are considered to have disease-modifying properties in early RA, but it is less clear whether they inhibit the progression of erosions in RA of longer duration. A meta-analysis of 15 studies that included 1414 patients concluded that GCs given in addition to standard therapy can substantially reduce the rate of erosion progression in RA.[24] Another meta-analysis of 70 randomized placebo- or drug-controlled studies found similar effects of DMARDs, GCs, and biologicals on radiographic progression in RA.[25]

Glucocorticoid Pulse Therapy

GC pulse therapy is used to treat some serious complications of RA and to induce remission in active disease, for example, when initiating second-line antirheumatic treatment. Pulse therapy with methylprednisolone 1 g/day, dexamethasone 200 mg/day, or equivalent, given intravenously for 1–3 days, was shown to be effective in most studies, demonstrating a beneficial effect that generally lasted for about 6 weeks, albeit with large variations. This means that it is not wise to apply pulse therapy in active RA, unless the therapeutic strategy is changed (*e.g.,* DMARD treatment introduced to maintain the remission induced by pulse therapy).

Intraarticular Glucocorticoid Injections

Intraarticular injections with GCs are often used in RA. The benefit achieved varies, depending on several factors, such as the joint treated (size, weight-bearing or non–weight-bearing), inflammatory activity, the volume of synovial fluid in the joint, whether or not synovial fluid was aspirated before the injection, the choice and dose of GC preparation, injection technique, and whether the joint is rested after the injection.[26] To prevent GC-induced joint damage, it is recommended that intraarticular GC injections should not be repeated more often than once every 3–4 weeks, and no more than 3–4 times a year in a weight-bearing joint.

KEY CONCEPTS

Glucocorticoid (GC) Therapy in Rheumatoid Arthritis (RA)

Our view of the risk–benefit ratio of low-dose GC has shifted in recent years:

- GCs can now be considered as disease-modifying antirheumatic drugs, especially in early RA.
- Adverse effects of low-dose GCs are less abundant and less severe than previously suggested, and some (*e.g.*, osteoporosis) can be well managed.
- The goal of antirheumatic treatment in early RA is to induce remission of disease by aggressive management: GCs are part of this aggressive strategy.

Adverse Effects

Studies of GC toxicity tend to be retrospective and observational. This can make it difficult to differentiate unfavorable outcomes attributable to GCs from those occurring as a result of the underlying disease or other comorbidities. Furthermore, there is a strong selection bias for GC use, as physicians are more inclined to use them in patients with more severe disease. Frequent, but less serious, adverse effects (*e.g.*, skin thinning, Cushingoid appearance) may be of great concern to patients, whereas more debilitating toxicities, such as osteoporosis, cataracts, and GC-induced hypertension, may initially go unrecognized or be asymptomatic. Interpretation of toxicity data is further confounded by the use of GCs at variable points in the disease course, limited data defining "threshold" doses for particular adverse events, and the fact that toxicity reports cover a heterogeneous group of GC-treated diseases.

Compared with other antirheumatic agents, GCs have a low incidence of short-term symptomatic toxicity, and patients rarely discontinue therapy for this reason. Despite >60 years of use, we still lack robust data on the longer-term toxicities of GCs from large randomized controlled trials with long-term follow-up. The most common GC toxicities are summarized below.

Some progress has been made by formulating recommendations on which adverse effects of GC treatment should be monitored in RA, how to monitor them, and how often. Two levels of monitoring GC adverse events have been proposed: *(i)* for routine clinical practice, details are given on how to identify adverse events in a systematic and practical way, and this should result in preventive and therapeutic measures to minimize the risks of GC therapy; and *(ii)* for clinical trials, recommendations have been made on how to accurately assess the frequency and severity of a wider range of adverse events.[27]

Osteoporosis

Glucocorticoid-induced osteoporosis (GIOP) is the most important potential complication of prolonged GC therapy. Chronic GC treatment results in rapid and profound reductions in bone mineral density, with most bone loss occurring during the first 6–12 months of treatment.[2,28]

GIOP initially affects trabecular bone, but cortical bone is also affected with more chronic use, at such sites as the femoral neck. Precisely how GCs affect bone remains obscure. GCs decrease calcium absorption, increase renal calcium loss, diminish sex and growth hormone production, induce muscle wasting, and modulate RANKL/OPG, NF-κB, and AP-1 signaling in bone.[2] All of these changes lead to enhanced osteoclast function and lifespan and thus to increased bone resorption. Consequently, markers of bone resorption are often increased in patients treated with GCs.[2] However, reduced bone formation as a result of reduced osteoblast function is likely to be a more important effect of GCs on skeletal health. Oral doses of prednisone as low as 2.5 mg/day have been shown to suppress serum osteocalcin, a marker of bone formation. Histologically, mean wall thickness is reduced, reflecting the reduced amount of bone replaced in each remodeling unit. *In vitro,* osteoblasts and their precursors are highly GC responsive. Here, the predominant effect is to promote osteoprogenitor proliferation, lineage commitment, and osteoblast differentiation, resulting in the formation of bone nodules of increased size and numbers. However, GCs also inhibit type I collagen expression and reduce preosteoblastic replication. Finally, GCs promote apoptosis of osteoblasts and osteocytes. The inhibitory effects of GCs on bone formation may partly be a result of downregulation of insulin-like growth factor 1 (IGF-1) expression by osteoblasts. Fortunately, we now have effective strategies for the prevention and treatment of GC-induced osteoporosis, using calcium, vitamin D, and specific osteotropic agents, such as bisphosphonates or parathyroid hormone.[28]

Osteonecrosis

Osteonecrosis has long been considered an important consequence of high-dose GC use. In a Japanese study of femoral head osteonecrosis, 35% of cases were related to GC treatment. Higher average dose may be a more important predictor of avascular necrosis of bone compared with cumulative dose. Osteonecrosis is particularly noted in SLE but rarely occurs in patients with RA receiving low-dose therapy, affecting <3% of patients. Osteonecrosis rarely occurs in patients with SLE on prednisone doses <20 mg/day.

Myopathy

As with osteonecrosis, GC-induced myopathy is rare in patients receiving low-dose GCs. In small studies, myopathy appears more closely associated with fluorinated GC preparations, such as triamcinolone, than with prednisone. Notably, myopathy has been reported after only 3 months' treatment with triamcinolone 8 mg/day. In general, myopathy attributable to prednisone only occurs after higher doses and longer durations of treatment.

Cardiovascular Adverse Effects

GC-induced hypertension seems to be, at least in part, mediated via fluid retention (as a result of mineralocorticoid effects); it is dose related and less likely with medium-dose or low-dose therapy. Individual variation in susceptibility and other factors, such as the starting level of blood pressure, dietary salt intake, functional renal mass, associated diseases, and drug therapy, may also play a role.

Another troublesome potential toxicity of low-dose GC is the development of premature atherosclerotic vascular disease. This has proven difficult to investigate: studies evaluating the effects of GCs on lipids and atherosclerosis in patients with RA have yielded mixed results, with some studies suggesting that GCs may, in fact, reverse unfavorable lipid changes. At present, there is no evidence of a strong association between low-dose

GCs and cardiovascular disease in RA, even though atherosclerotic vascular disease is known to be accelerated in patients with Cushing syndrome. Systemic GC use during pregnancy may also have epigenetic effects leading to adult hypertension in the next generation.[29]

Dermatological Adverse Effects

Skin thinning and ecchymoses are common adverse events with GCs, even at low doses. Cutaneous atrophy results from the catabolic effects of GCs on keratinocytes and fibroblasts. Purpura and easy bruising in GC-treated patients probably result from decreased vascular structural integrity. A Cushingoid appearance is very troubling to patients but is uncommon at doses below the physiological range. One study reported facial fullness ("moon facies") in 13% of patients receiving 4–12 mg triamcinolone for up to 60 days. These adverse effects are observed in over 5% of patients exposed to ≥5 mg prednisone equivalent for ≥1 year. The incidence of iatrogenic Cushing syndrome is dose dependent and generally becomes evident after >1 month of GC therapy.[30] Alternate-day therapy may reduce the incidence, although there are only limited data supporting this concept. GC-induced acne, hirsutism, and striae are other undesirable dermatological effects that occur even with lower doses.

Gastrointestinal Adverse Effects

GCs are considerably less toxic to the upper gastrointestinal (GI) tract compared with NSAIDs. If GCs independently increase the risk of GI events (*e.g.,* gastritis, ulceration, bleeding), the effect is slight, with estimated relative risks varying from 1.1 (not significant) to 1.5 (marginally significant). There have also been anecdotal reports of intestinal rupture, diverticular perforation, and pancreatitis attributed to low-dose GC therapy. GCs are frequently used concurrently with NSAIDs in RA, and meta-analyses have confirmed that the combination of the two drugs synergistically increases the risk of adverse GI events. In a large-scale study based on the UK General Practice Research Database, the risk of upper GI complications was 1.8 (95% confidence interval [CI] 1.3–2.4) times higher in GC users than in nonusers. The risk tended to be greater with higher GC doses, but the dose gradient was not statistically significant. The risk was >12 times higher in those taking both GCs and NSAIDs than in those not using either. No studies have yet looked at the GI effects of combining GCs with cyclooxygenase 2 (COX-2)–selective NSAIDs. Preventive strategies for the adverse effects of GC on the GI tract have been reviewed by Caplan et al.[30]

Infectious Diseases

Medium- to high-dose GC therapy may increase the risk of serious infections leading to hospitalization or surgery, particularly when administered for prolonged periods. However, there is no evidence that infection rates are increased in patients on doses of prednisone below 10 mg/day or cumulative doses below 700 mg. In those taking higher doses, the risk of infection appears to be lessened with alternate-day therapy.

In GC-treated patients, physicians should be aware of the risk of infections with typical and atypical organisms, recognizing that GCs may blunt classic clinical features and delay diagnosis. However, it is difficult to separate the independent effects of GC use from those of other commonly used antirheumatic agents, such as methotrexate and anti-TNF-α agents.[31] At present, the independent role of GCs in facilitating herpes zoster virus (HZV) infection in patients with RA remains uncertain.

Other Adverse Effects

Adverse effects on the HPA axis and on glucose metabolism are reviewed elsewhere, together with neuropsychiatric and ophthalmological adverse effects.[30,32,33]

IMPORTANCE OF TIMING OF GLUCOCORTICOID ADMINISTRATION

In patients with RA, major symptoms, such as pain, inflammation, and stiffness, vary during the course of the day, usually with the highest severity in the morning hours. These symptoms are preceded by elevated levels of proinflammatory cytokines. On the basis of these considerations, it has been proposed that altering the timing of GC administration may help optimize RA therapy.[19,34]

NEW GLUCOCORTICOID RECEPTOR LIGANDS ON THE HORIZON

The various mechanisms of GC action provide interesting opportunities for developing optimized GCs and GCR ligands.

Selective Glucocorticoid Receptor Agonists

The genomic component mechanisms of transactivation and transrepression offer the opportunity to develop GCR ligands that predominantly cause transrepression rather than transactivation. This concept is based on the proposition that the antiinflammatory properties of GCs mostly result from repression of AP-1–stimulated and NF-κB–stimulated synthesis of inflammatory mediators, whereas their adverse effects are associated with transactivation of genes involved in metabolic processes. Investigators have, therefore, sought novel GCR ligands with high transrepression activity but low effect on transactivation. One such compound, A276575, exhibits a high affinity for the GCR and potently represses IL-1α–induced IL-6 production, similar to dexamethasone. However, unlike dexamethasone, A276575 induces little aromatase activity. Other novel nonsteroidal GCR ligands that possess high repression activities against inflammatory mediator production but have lower transactivation activities than traditional GC are being developed. Substances that cause a receptor conformation preferring a GCR–protein interaction as opposed to a GCR/DNA binding–dependent mechanism are called *selective glucocorticoid receptor agonists* (SEGRA) or "dissociated glucocorticoids." The SEGRA concept has, however, been challenged by studies on a mouse knock-in strain with a dimerization-deficient GCR, which demonstrated that some inflammatory processes can be suppressed by GCs and others cannot.[35] Also, these mice exhibited the classic adverse effects of GCs, such as GC-induced osteoporosis. Thus depending on the process being treated, SEGRA could be therapeutically more effective or less effective; moreover, not all adverse effects of GC therapy may be reduced.[35,36] As of now, whether SEGRA will become clinically relevant in practice remains uncertain.

Nitrosteroids

Nitrosteroids are another new class of GC drugs that have been tested in RA and inflammatory bowel disease (IBD). In addition to their enhanced GC properties, nitrosteroids release low levels of NO and have fewer undesirable effects. The prototype of these new steroids, 21-NO-prednisolone (NCX-1015), is much more potent than prednisolone in models of acute and chronic inflammation, including experimental arthritis induced by collagen II.

In contrast, NCX-1015 did not activate primary osteoclast activity *in vitro,* whereas prednisolone did. This lack of adverse effect of NCX-1015 seems to be chiefly caused by NO. It has been suggested that posttranslational modification of GCR (tyrosine nitration) by NCX-1015 may explain its enhanced antiinflammatory activity. Moreover, NCX-1015 potently stimulates IL-10 production, suggesting that nitrosteroids induce Tregs that negatively modulate inflammation. However, further work is needed to confirm the utility of nitrosteroids in clinical practice.

Long-Circulating Liposomal Glucocorticoids

The antiinflammatory efficacy of GCs can be improved by the additional benefits of nongenomic actions at high GC concentrations. This has led to the use of long-circulating liposomal GCs in experimental models. In rats with experimental autoimmune encephalitis, GC-containing liposomes accumulated at sites of inflammation, reaching concentrations >10–5 mol/L for ≥18 hours. These liposomes may be therapeutically superior to conventional intravenous high-dose GC therapy, as evidenced by their successful use in rats with adjuvant-induced arthritis. A single injection of 10 mg/kg liposomal prednisolone phosphate resulted in complete remission of the inflammatory response for almost a week. In contrast, the same dose of unencapsulated prednisolone phosphate did not reduce inflammation and had only a slight effect after repeated daily injections. It may be that preferential delivery of GCs to the site of inflammation leads to very high GC concentrations at the inflamed joint, but lower plasma concentrations with a consequent lower rate of adverse effects. These are very promising developments that exploit the broad spectrum of therapeutically relevant genomic and nongenomic GC actions at the site of inflammation.

Combining Prednisolone and Dipyridamole

Another interesting recent development is the selective amplification of GC antiinflammatory activity through synergistic multitarget action with another drug.[18] The combination of prednisolone and the antithrombotic drug dipyridamole synergistically suppresses the release of proinflammatory cytokines and produces antiinflammatory activity in acute and chronic disease models (including arthritis) at a subtherapeutic dose of prednisolone. This approach may create a dissociated activity profile with an increased therapeutic window through cellular network selective amplification of GC-mediated antiinflammatory signalling.[18] However, the exact molecular mechanism underlying the observed synergy and multitarget action of these two drugs remains obscure.

CONCLUSIONS

GCs are extremely valuable antiinflammatory agents with a range of actions that are useful in the management of inflammatory conditions, including arthritis and asthma. Combination therapy with biological agents may achieve superior benefit in rheumatological disease compared with either class of treatment alone,[37] but GCs will remain important drugs for many years to come. Recent research on GCs has highlighted the effects of cytosolic GCRs on intracellular signaling, transcription processes, and gene expression. Exploration of membrane-bound GCRs, dose–effect relationships, and the timing of GC administration have stimulated intensive research activity aimed at improving the efficacy–risk ratio. Modified-release prednisone has been already approved for the treatment of RA and the associated morning stiffness.[38] Further clinical developments seem likely to follow (*e.g.,* new GCR ligands and liposome encapsulation), all with a view to improving the risk–benefit ratio of GC therapy and the well-being of patients.

ON THE HORIZON

- Selective activators of transrepression (SEGRA)
- Nitrosteroids (with effect of nitric oxide on endothelium)
- Modified release steroids
- Liposomal encapsulation
- Optimized combinations with biologicals

ACKNOWLEDGMENTS

This chapter has been revised and updated from the version in the previous edition. The editors are grateful to Frank Buttgereit, Hans Bilsma and Markus Seibel for their contributions to previous editions and for their permission to let us update the chapter.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Buttgereit F, Straub RH, Wehling M, et al. Glucocorticoids in the treatment of rheumatic diseases. An update on mechanisms of action. Arthritis Rheum 2004;50:3408–17.
2. Caplan A, Fett N, Rosenbach M, et al. Prevention and management of glucocorticoid-induced side effects: a comprehensive review of glucocorticoid pharmacology and bone health. J Am Acad Dermatol 2017;76:1–9.
3. Ramsey-Goldman R. Missed opportunities in physician management of glucocorticoid-induced osteoporosis? Arthritis Rheum 2002;46:3115–20.
4. Walsh LJ, Wong CA, Pringle M, et al. Use of oral corticosteroids in the community and the prevention of secondary osteoporosis: a cross section study. Br Med J 1996;313:344–6.
5. Grbesa I, Hakim O. Genomic effects of glucocorticoids. Protoplasma 2017;254:1175–85.
6. Greulich F, Hemmer MC, Rollins DA, et al. There goes the neighborhood: assembly of transcriptional complexes during the regulation of metabolism and inflammation by the glucocorticoid receptor. Steroids 2016;114:7–15.
7. Petta I, Dejager L, Ballegeer M, et al. The interactome of the glucocorticoid receptor and its influence on the actions of glucocorticoids in combatting inflammatory and infectious diseases. Microbiol Mol Biol Rev 2016;80:495–522.
8. Murphy PJ, Morishima Y, Chen H, et al. Visualization and mechanism of assembly of a glucocorticoid receptor-Hsp70 complex that is primed for subsequent Hsp90-dependent opening of the steroid binding cleft. J Biol Chem 2003;278:34764–73.
9. Davies TH, Ning YM, Sánchez ER. A new first step in activation of steroid receptors. J Biol Chem 2002;277:4597–600.
10. Yudt MR, Jewell CM, Bienstock RJ, et al. Molecular origins for the dominant negative function of human glucocorticoid receptor beta. Mol Cell Biol 2003;23:4319–30.
11. Rodriguez JM, Monsalves-Alvarez M, Henriquez S, et al. Glucocorticoid resistance in chronic diseases. Steroids 2016;115:182–92.
12. Croxtall JD, Choudhury Q, Flower RJ. Glucocorticoids act within minutes to inhibit recruitment of signaling factors to activated EGF receptors through a receptor-dependent, transcription-independent mechanism. Br J Pharmacol 2000;130:289–98.
13. Hafezi-Moghadam A, Simoncini T, Yang E, et al. Acute cardiovascular protective effects of corticosteroids are mediated by nontranscriptional activation of endothelial nitric oxide synthase. Nat Med 2002;8:473–9.

14. Palmowski Y, Buttgereit T, Dejaco C, et al. The “official view” on glucocorticoids in rheumatoid arthritis. A systematic review of international guidelines and consensus statements. Arthritis Care Res (Hoboken) 2017;69:1134–41.
15. Badsha H, Edwards CJ. Intravenous pulses of methylprednisolone for systemic lupus erythematosus. Semin Arthritis Rheum 2003;32:370–7.
16. Vasoo S, Thumboo J, Fong KY. Refractory immune thrombocytopenia in systemic lupus erythematosus: response to mycophenolate mofetil. Lupus 2003;12:630–2.
17. Bartholome B, Spies CM, Gaber T, et al. Membrane glucocorticoid receptors (mGCR) are expressed in normal peripheral blood mononuclear cells and upregulated following in vitro stimulation and in patients with rheumatoid arthritis. FASEB J 2004;18:70–80.
18. Buttgereit F, Saag K, Cutolo M, et al. The molecular basis for the effectiveness, toxicity, and resistance to glucocorticoids: focus on the treatment of rheumatoid arthritis. Scand J Rheumatol 2005;34:14–21.
19. Buttgereit F, Burmester GR, Straub RH, et al. Exogenous and endogenous glucocorticoids in rheumatic diseases. Arthritis Rheum 2011;63:1–9.
20. Hardy R, Rabbitt EH, Filer A, et al. Local and systemic glucocorticoid metabolism in inflammatory arthritis. Ann Rheum Dis 2008;67:1204–10.
21. Buttgereit F, Zhou H, Kalak R, et al. Transgenic disruption of glucocorticoid signaling in mature osteoblasts and osteocytes attenuates K/BxN mouse serum-induced arthritis in vivo. Arthritis Rheum 2009;60:1998–2007.
22. da Silva JAP, Jacobs JWG, Kirwan JR, et al. Long-term glucocorticoid therapy in rheumatoid arthritis: an evidence-based review of potential adverse effects. Ann Rheum Dis 2006;65:285–93.
23. Shane E, Rivas M, McMahon JD, et al. Bone loss and turnover after cardiac transplantation. J Clin Endocrinol Metab 1997;82:1497–506.
24. Kirwan JR, Bijlsma JW, Boers M, et al. Effects of glucocorticoids on radiological progression in rheumatoid arthritis. Cochrane Database Syst Rev 2007;(1):CD006356.
25. Graudal N, Jurgens G. Similar effects of disease modifying anti rheumatic drugs, glucocorticoids and biologic agents on radiographic progression in rheumatoid arthritis: meta-analysis of 70 randomised placebo or drug controlled studies including 112 comparisons. Arthritis Rheum 2010;62:2852–63.
26. Gaffney K, Ledingham J, Perry JD. Intra-articular triamcinolone hexacetonide in knee osteoarthritis: factors influencing the clinical response. Ann Rheum Dis 1995;54:379–81.
27. van de Goes M, Jacobs J, Boers M, et al. Monitoring adverse events of low-dose glucocorticoid therapy: EULAR recommendations for clinical trials and daily practice. Ann Rheum Dis 2010;69:1913–19.
28. Allen CS, Yeung JH, Vandermeer B, et al. Bisphosphonates for steroid-induced osteoporosis. Cochrane Database Syst Rev 2016;(10):CD001347.
29. Anwar MA, Saleh AI, Al Olabi R, et al. Glucocorticoid-induced fetal origins of adult hypertension: association with epigenetic events. Vascul Pharmacol 2016;82:41–50.
30. Caplan A, Fett N, Rosenbach M, et al. Prevention and management of glucocorticoid-induced side effects: a comprehensive review: gastrointestinal and endocrinologic side effects. J Am Acad Dermatol 2017;76:11–16.
31. Singh JA, Hossain A, Kotb A, et al. Risk of serious infections with immunosuppressive drugs and glucocorticoids for lupus nephritis: a systematic review and network meta-analysis. BMC Med 2016;14:137.
32. Joseph RM, Hunter AL, Ray DW, et al. Systemic glucocorticoid therapy and adrenal insufficiency in adults: a systematic review. Semin Arthritis Rheum 2016;46:133–41.
33. Black RJ, Hill CL, Lester S, et al. The association between systemic glucocorticoid use and the risk of cataract and glaucoma in patients with rheumatoid arthritis: a systematic review and meta-analysis. PLoS ONE 2016;11:e0166468.
34. Buttgereit F, Doering G, Schaeffler A, et al. Targeting pathophysiological rhythms: prednisone chronotherapy shows sustained efficacy in rheumatoid arthritis. Ann Rheum Dis 2010;69:1275–80.
35. Kleiman A, Tuckermann JP. Glucocorticoid receptor action in beneficial and side effects of steroid therapy: lessons from conditional knockout mice. Mol Cell Endocrinol 2007;275:98–108.
36. Schacke H, Zollner TM, Docke WD, et al. Characterization of ZK 245186, a novel, selective glucocorticoid receptor agonist for the topical treatment of inflammatory skin diseases. Br J Pharmacol 2009;158:1088–103.
37. Inui K, Koike T. Combination therapy with biologic agents in rheumatic diseases: current and future prospects. Ther Adv Musculoskelet Dis 2016;8:192–202.
38. Krasselt M, Baerwald C. Efficacy and safety of modified-release prednisone in patients with rheumatoid arthritis. Drug Des Devel Ther 2016;10:1047–58.

MULTIPLE-CHOICE QUESTIONS

1. Genomic effects of glucocorticoids (GCs) include ALL of the following except:
   - **A.** Induction of antiinflammatory cytokines
   - **B.** Inhibition of arachidonic acid release
   - **C.** Interaction with transcription factors AP-1 and nuclear factor-κB
   - **D.** Competition for nuclear coactivators

2. The most important complication of GC use is:
   - **A.** Increased risk of infections
   - **B.** Diabetes mellitus
   - **C.** Osteoporosis
   - **D.** Glaucoma

3. Promising future prospects for steroid development include:
   - **A.** Liposomal encapsulation
   - **b.** Nitrogenated steroids
   - **C.** Slow release preparations
   - **D.** Selective inducers of transactivation

87

Immunomodulating Pharmaceuticals

Gideon P. Smith, Edwin S.L. Chan

Excitement over biologic agents and their capacity to regulate immunological reactions that significantly impact on such immunologically mediated diseases, such as rheumatoid arthritis (RA) and inflammatory bowel disease (IBD), has overshadowed the older, small-molecule therapeutic agents. Nonetheless, when tested head-to-head, some small-molecule agents (most notably methotrexate) have proven to be almost as effective as biologics; moreover, combining small-molecule therapies with biological agents generally leads to significantly better outcomes than use of either agent alone.[1] Thus it is likely that small-molecule immunomodulatory drugs will continue to be in wide use. Here, we review the most widely accepted and commonly used immunomodulators currently in clinical use.

METHOTREXATE

Methotrexate (Fig. 87.1) was employed in the treatment of RA as early as 1951, but its popularity with regard to RA did not come until the 1980s. Over the years, extensive experience with its use in inflammatory diseases as diverse as RA (Chapter 52), psoriasis (Chapter 64), and IBD (Chapter 75) has taught us a great deal about its safety, efficacy, and toxicity, as well as its antiinflammatory mechanisms of action. In this respect, methotrexate, much as corticosteroids, can be justly regarded as a cornerstone of immunomodulatory therapy.

Pharmacokinetics of Methotrexate

As an antiinflammatory agent, methotrexate is administered at low doses (usually 10–25 mg/week) once weekly, usually orally, but it can also be given subcutaneously or intramuscularly. At these doses, oral bioavailability is high (60–70%), and although transporters are responsible for its absorption from the gastrointestinal (GI) tract, saturation effect does not occur. A small portion of methotrexate is metabolized by hydroxylation into 7-hydroxymethotrexate. Both compounds have a serum half-life of no more than 8 hours. The much longer antiinflammatory action, which allows for once-weekly dosing, must therefore be mediated by other longer-lasting metabolites, such as polyglutamates. Excretion occurs principally via the urinary tract but also via the biliary tract. Therefore renal function is an important consideration in methotrexate dosing, and any medication that impairs glomerular filtration may also potentiate methotrexate's effectiveness and toxicity.[2]

Mechanisms of Action for Methotrexate

As an analogue of folic acid, methotrexate is an inhibitor of purine and pyrimidine synthesis and thereby suppresses cellular proliferation (Table 87.1). These actions are dependent on inhibition of dihydrofolate reductase; hence toxicities arising from high-dose methotrexate therapy can be treated with folic acid derivatives, such as leucovorin. However, folic or folinic acid, which are often given in conjunction with methotrexate in inflammatory diseases to reduce the incidence of mucositis and bone marrow suppression, do little to inhibit its antiinflammatory efficacy. Decreases in purine and pyrimidine concentrations in the serum have been observed following a single dose of methotrexate, along with decreased proliferation of antigen-stimulated lymphocytes. However, these changes are transient and insufficient to explain the antiinflammatory effectiveness of once-weekly dosing. This, as well as the low doses of methotrexate required to produce an antiinflammatory effect, suggests that the antiinflammatory actions are mediated via different mechanisms.

Methotrexate also blocks intracellular transmethylation reactions and inhibits production of *S*-adenosylmethionine. Since *S*-adenosylmethionine is necessary for formation of the toxic polyamine metabolites spermine and spermidine, their accumulation at the inflammatory site is prevented. This inhibition of transmethylation is associated with an impairment of monocyte and lymphocyte function and thus potentially the synthesis of reactive oxygen species. However, diminution of transmethylation by the use of the *S*-adenosylhomocysteine hydrolase inhibitor, deaza-adenosine, has failed to produce any beneficial clinical effects in RA.

Methotrexate and its long-acting polyglutamate metabolites also exert antiinflammatory effects by releasing the endogenous autocoid adenosine.[3] As potent inhibitors of the enzyme 5-aminoimidazole-4-carboxamide ribonucleotide (AICAR) transformylase, methotrexate polyglutamates promote accumulation of AICAR in tissues. Since AICAR inhibits catabolizing enzymes for both adenosine and adenosine monophosphate (AMP), which can be dephosphorylated to adenosine, the net effect is intracellular and extracellular increases in adenosine levels. These metabolic pathways are pharmacologically relevant, since aminoimidazole carboxamide and adenosine have been shown to be increased in urine following low-dose methotrexate treatment in patients with psoriasis.[4] Adenosine causes diminution of neutrophil accumulation, adhesion, phagocytosis and generation of reactive oxygen species, inhibition of adhesion molecule expression, suppression of proinflammatory cytokines, and induction of antiinflammatory cytokines, as well as modulation of macrophage and endothelial function.[2] Indeed, blockade of adenosine receptors reversed the antiinflammatory effects of methotrexate on animal models. It has also been suggested that

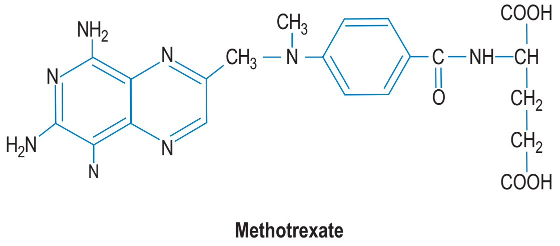


FIG 87.1 Methotrexate—chemical structure.

TABLE 87.1 Methotrexate: Mechanisms of Action

Suggested Mechanism	Rationale
Folate antagonism	Prevents purine and pyrimidine synthesis required for the proliferation of actively dividing immune cells, such as lymphocytes
Inhibition of spermine and spermidine production	Reduces formation of polyamines harmful to tissues
Alteration of cellular redox state	Reversible inhibition of lymphocyte and macrophage functions
Release of adenosine	Generation of a potent endogenous antiinflammatory mediator through inhibition of catabolism of both adenosine and adenosine monophosphate

caffeine, itself a nonselective antagonist of adenosine receptors, may both reduce the effectiveness of methotrexate in RA and protect against the development of cirrhosis of the liver, a major side effect of methotrexate.[5]

Adverse Effects

THERAPEUTIC PEARLS

Methotrexate

Proven safety profile
Concomitant administration of folic acid advisable
Antiinflammatory effects may be reduced by heavy use of caffeine
Hepatotoxicity a rare but real concern
Risk of hepatotoxicity increased with:
- Alcohol use
- Hepatitis
- Diabetes
- Obesity
- Alpha-1-antitrypsin deficiency

Over the years, methotrexate has proven to be one of the safest disease-modifying antirheumatic drugs (DMARDs) in use. Serious side effects such as cirrhosis are much less common than previously thought (Table 87.2). The use of folic acid has decreased the occurrence of mucosal and GI side effects, without limiting its antiinflammatory activity, and cytopenias are managed adequately with regular blood counts. Although side effects, such as nausea and vomiting, may resolve spontaneously or respond to dose reduction or folic acid supplementation, mild transaminasemia has rarely led to discontinuation of the medication. Risk of serious hepatotoxicity over 5 years of use is likely to be less than 1 in every 1000 patients with RA but may be more common among those with psoriasis. Risk factors, such as ethanol consumption, hepatitis B and C, diabetes, obesity, and alpha-1-antitrypsin deficiency, identify patients most likely to develop methotrexate-induced hepatic injury. However, other serious side effects, such as pneumonitis, may be overlooked, since early symptoms (mild cough or shortness of breath) are often ignored. Early identification allows for prompt discontinuation. The risk of developing solid tumors is debated, since the risk of malignancies is intrinsic to some of the conditions, such as RA, for which methotrexate is used. It is likely the risk of drug-induced malignancy is real, since reports have documented tumor regression following discontinuation of methotrexate, but the risk remains extremely small.

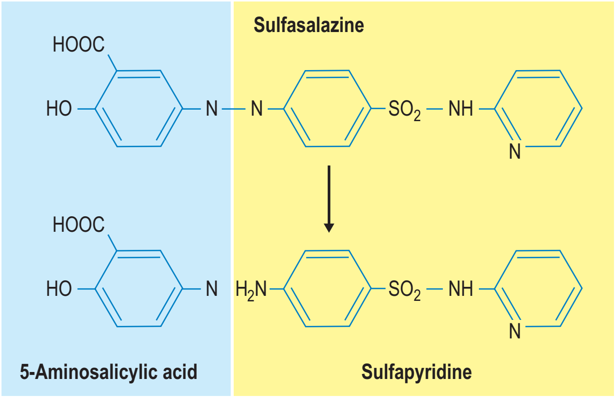


FIG 87.2 Sulfasalazine—chemical structure.

TABLE 87.2 Methotrexate: Adverse Effects

Gastrointestinal
Stomatitis
Anorexia
Nausea
Vomiting
Diarrhea
Cirrhosis
Pancreatitis

Hematological
Leukopenia
Anemia
Thrombocytopenia
Hypogammaglobulinemia
Lymphoma

Cardiovascular
Pericarditis
Thrombosis

Pulmonary
Pulmonary fibrosis
Interstitial pneumonitis

Others
Skin rashes
Renal failure
Abortion
Impotence
Headache
Opportunistic infections

SULFASALAZINE

Sulfasalazine (Fig. 87.2) was originally introduced in the late 1930s for the treatment of RA but is now used in a wide range of inflammatory diseases, in particular, IBD and the seronegative arthritides. It consists of a derivative of the antiinflammatory salicylic acid, 5-amino-salicylic acid, and the antimicrobial sulfapyridine. These two moieties are joined together by an azo bond. Which component is responsible for the drug's antiinflammatory actions is unclear, but it appears to vary according to disease states. For instance, in IBD, 5-amino-salicylic acid is likely the main active component, as it is poorly absorbed following metabolism by intestinal flora. In inflammatory arthritides, sulfapyridine is likely to play a more important role, as it is relatively well absorbed and has a bioavailability in the region of 60%. Since acetylation is the principal route of metabolism of sulfapyridine following absorption, acetylator status is a major determinant of the plasma half-life. For the same reason, slow acetylators are more liable to develop side effects.

Mechanisms of Action for Sulfasalazine

KEY CONCEPT

Sulfasalazine: Mechanism of Action

Suppresses the proliferation of lymphocytes
Suppresses proinflammatory cytokine production
Inhibits activation of nuclear factor (NF)-κB
Promotes adenosine accumulation

Sulfasalazine has a number of immunomodulatory effects. Lymphocyte proliferation is suppressed *in vitro* and involves both B-cell and T-cell populations. *In vivo,* a decrease in activated lymphocytes in the peripheral blood is also seen. Tumor necrosis factor-α (TNF-α) production is suppressed, and receptor binding is inhibited. Sulfasalazine also inhibits activation of the transcription factor, nuclear factor (NF)-κB. Like methotrexate, sulfasalazine inhibits AICAR transformylase and thus promotes accumulation of adenosine and its antiinflammatory actions via the adenosine A_{2A} receptor. Indeed, treatment of animals with an adenosine A_{2A} receptor antagonist reversed sulfasalazine's reduction of leukocyte accumulation in an air-pouch model of inflammation.

Adverse Effects

In a large series, a quarter of those treated over 11 years stopped treatment because of toxicity.[6] Most toxicities occurred early and were both trivial and resolved following therapy withdrawal. Most common are nausea, vomiting, anorexia, and rash. Serious cutaneous reactions, such as exfoliative dermatitis or Stevens-Johnson syndrome, are rare. Transaminasemia and drug-induced hepatitis can occur. Blood dyscrasias with megaloblastic anemia, neutropenia, aplastic anemia, and myelodysplastic syndrome may arise. Neurological adverse effects include headache and dizziness or, more seriously, peripheral neuropathy, Guillain-Barré syndrome, or transverse myelitis. Sulfasalazine should be avoided in patients with sulfa allergy, and glucose-6-phosphate dehydrogenase (G6PD) deficiency screening should be performed before prescribing.

AZATHIOPRINE

Azathioprine, an imidazolyl derivative of 6-mercaptopurine (Fig. 87.3), has been widely used in RA and IBD as well as in solid organ transplantations. Cleavage into 6-mercaptopurine and the imidazole moiety occurs rapidly within erythrocytes, both enzymatically by glutathione transferase and nonenzymatically. Several enzymes (Table 87.3) participate in the metabolism of 6-mercaptopurine into active and inactive compounds. One of these, thiopurine methyltransferase, is associated with genetic polymorphisms; inherited changes in its activity may impact patient response to azathioprine. Xanthine oxidase inactivates 6-mercaptopurine by converting it to 6-thiouric acid. Since this occurs mainly in the liver, toxicity from azathioprine therapy is a danger in enzyme deficiency states, as a result of disease or use of drugs, such as allopurinol.

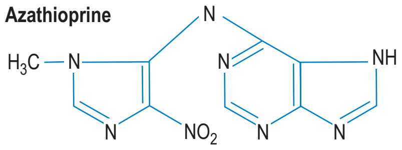


FIG 87.3 Azathioprine—chemical structure.

Proposed Mechanisms of Action for Azathioprine

The immunomodulatory mechanism of azathioprine remains unclear. As purine analogues, the active metabolites interfere with the salvage pathway and *de novo* synthesis of purines, and they are incorporated into RNA and DNA. Proliferation of T and B lymphocytes is inhibited, and function of natural killer (NK) cells is suppressed without any change in cell numbers. The production of antibodies is also suppressed, although it is not known which of these effects predominate *in vivo.* Cellular responses to chemoattractants are altered, and the production of cytokines, such as interleukin-6 (IL-6), is also affected.

Adverse Effects

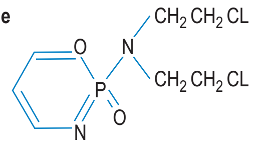
Azathioprine is generally well tolerated. The most common side effects are mild and affect the GI system. Pancreatitis can occur as an idiosyncratic reaction. Hepatotoxicity and cholestasis are not uncommon, and hepatic peliosis and nodular regenerative hyperplasia occur rarely. There have been reports of a possible heightened risk for non-Hodgkin lymphoma, but because of the rarity of these events, no definite link has been established between azathioprine and malignancies. Bone marrow suppression and opportunistic infections pose far greater threats.

CYCLOPHOSPHAMIDE

Alkylating agents were used in the treatment of inflammatory diseases after promising reports on using nitrogen mustard in RA. Cyclophosphamide (Fig. 87.4) is metabolized to produce the alkylating agent phosphoramide mustard as well as acrolein, which, although inactive, results in the hemorrhagic cystitis associated with cyclophosphamide. Cyclophosphamide can be given intravenously, but bioavailability following oral administration is high (>75%). Toxicity has severely limited its use in inflammatory diseases, although its contribution to the management of lupus nephritis cannot be denied. The alkylating actions occur at guanine residues (principally on DNA but also on RNA) resulting in cross-linkable strands and disruption of transcription and translation.

TABLE 87.3 Principal Enzymes Involved in the Metabolism of Azathioprine

Enzyme	Action
Glutathione transferase	Cleaves azathioprine into 6-mercaptopurine and imidazole moieties
Thiopurine methyltransferase	Metabolism of 6-mercaptopurine
Xanthine oxidase	Conversion of 6-mercaptopurine to 6-thiouric acid



Cyclophosphamide

FIG 87.4 Cyclophosphamide—chemical structure.

Mechanisms of Action of Cyclophosphamide

This alkylating process has immunomodulating effects on resting and actively dividing cells. The numbers of circulating CD4 T lymphocytes and, to a lesser extent, CD8 T lymphocytes are reduced, thus reducing the CD4/CD8 ratio. Despite an apparent increase in immunoglobulin-secreting cells, B-cell function is suppressed, and overall immunoglobulin synthesis is reduced.

Adverse Effects

The best known toxicity is hemorrhagic cystitis. Since this occurs more frequently following oral dosing, this route of administration is rarely used. This may relate to continuous exposure of the bladder to acrolein, so the acrolein-neutralizing agent 2-mercaptoethane sulfonate (Mesna) is used prophylactically along with copious hydration. Hemorrhagic myocarditis can also occur and may cause myocardial necrosis, hemopericardium, and congestive cardiac failure. However, survivors of acute cardiac toxicity do not show any residual electrocardiographic or echocardiographic abnormalities.

Besides bone marrow suppression, reduction of fertility, and a heightened risk of infection, cyclophosphamide therapy has been associated with secondary malignancies, which may occur years after drug cessation. Malignancies of the bladder, often of a transitional cell type, tend to occur only in those with a history of treatment-related hemorrhagic cystitis. Myeloproliferative and lymphoproliferative disorders have also been associated with cyclophosphamide use.

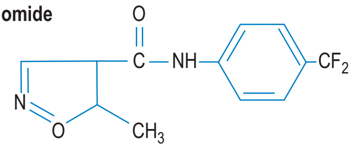
Other Nitrogen Mustard Derivatives

Chlorambucil, or 4-[bis(2chlor-ethyl)amino]benzenebutanoic acid, has wide distribution in tissue and 87% oral bioavailability, but unlike cyclophosphamide, it does not require metabolism by the liver to become metabolically active.[7] The only indication approved by the US Food and Drug Administration (FDA) is for treatment of chronic lymphocytic leukemia (CLL), but like cyclophosphamide, chlorambucil has been reported to be used in treatments for the same wide range of inflammatory conditions. The mechanism of action and side effect profile is also similar to those of cyclophosphamide, but with a higher risk of permanent aplasia.[8]

Melphalan is another phenylalanine derivative of nitrogen mustard, mainly used in treating multiple myeloma. It has been less widely adopted but has been used off-label in a variety of inflammatory conditions. Adverse effects and mechanism are the same as those of cyclophosphamide, as described above.

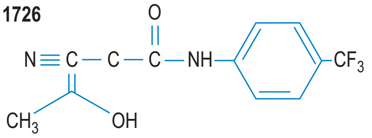
LEFLUNOMIDE

Leflunomide is an inhibitor of *de novo* pyrimidine synthesis. Leflunomide (Fig. 87.5) is converted into the long-acting active compound A77 1726 (2-cyano-3-hydroxy-N-[4-trifluoromethyl]-butenamide) (Fig. 87.6), a reversible inhibitor of the enzyme dihydroorotate dehydrogenase that is involved in pyrimidine synthesis. Because of this long half-life, therapy with leflunomide is usually started with a loading dose to quickly achieve therapeutic levels. A77 1726 is highly plasma-protein bound and undergoes enterohepatic recirculation.



Leflunomide

FIG 87.5 Leflunomide—chemical structure.



A77 1726

FIG 87.6 A77 1726—chemical structure.

Mechanisms of Action of Leflunomide

By inhibiting pyrimidine synthesis, pyrimidine nucleotide availability becomes insufficient for proliferation of immune-response cells. This deficiency is inadequately replenished by the salvage pathways, rendering cell proliferation inefficient and limiting the clonal expansion of T cells. B-cell proliferation is similarly suppressed with reduction of Cdk2, a cyclin-dependent kinase. Leflunomide also inhibits NF-κB activation. Although the effects of moderate concentrations are reversed by uridine *in vitro*, this reversal does not occur at higher concentrations, suggesting possible involvement of other mechanisms. Leflunomide is known to inhibit tyrosine kinase activity at higher concentrations, although the relevance of this effect to therapeutic concentrations achievable *in vivo* remains questionable.

Adverse Effects

GI symptoms are the most common side effect, and hepatic damage is the most important toxicity. Although there are similarities to the toxicity profile of methotrexate, clinical trials have shown that leflunomide and methotrexate can be safely and effectively given together to patients with RA, but transaminasemia occurs more often than with methotrexate alone.[9,10] Fulminant hepatic failure is rare, but fatal cases have occurred. Skin reactions are mostly minor; however, more serious manifestations, such as Stevens-Johnson syndrome and toxic epidermal necrolysis, have been reported.

MYCOPHENOLATE MOFETIL

Mycophenolate mofetil (Fig. 87.7) is widely used in solid organ transplantation and has also been increasingly employed in treatment of autoimmune diseases. It is rapidly absorbed and hydrolyzed into the active compound, mycophenolic acid, which is a reversible inhibitor of inosine monophosphate dehydrogenase. Since inosine monophosphate dehydrogenase is a key enzyme in the *de novo* synthesis of guanine nucleotides, its inhibition is most significant in T and B lymphocytes, which are reliant on this pathway, as they lack the hypoxanthine-guanine phosphoribosyl transferase salvage pathway. The immunological effects are numerous. DNA synthesis in lymphocytes requires the incorporation of guanine nucleotides so that proliferation of

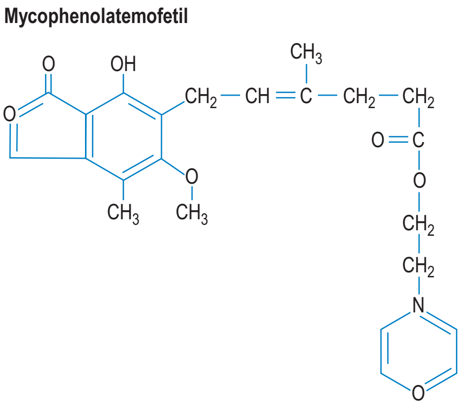


FIG 87.7 Mycophenolate mofetil – chemical structure.

lymphocytes is suppressed. Antibody production and NK-cell activity are also reduced, and *in vitro* cytokine production by activated human mononuclear cells is affected.[11] In addition, delayed-type hypersensitivity responses are suppressed. Although effective in a subset of patients with psoriasis and RA, mycophenolate mofetil has not been widely used in these conditions because other more effective medications are available. It is, however, becoming more popular in the treatment of some diseases, such as myositis, systemic contact dermatitis, severe atopic dermatitis, chronic urticaria, refractory pyoderma gangrenosum, bullous pemphigoid, pemphigus vulgaris, and pemphigus foliaceus, where it is effective with a low risk of side effects.

Adverse Effects

Absolute contraindications for mycophenolate mofetil are drug allergy and pregnancy (Category D). Relative contraindications include lactation; renal, hepatic, or cardiopulmonary disease; and peptic ulcer. It is generally well tolerated when used in autoimmune diseases, such as RA. The most common side effects are nausea, vomiting, abdominal discomfort, diarrhea, fever, headache, skin rash, back pain, and tremor, but these do not usually lead to discontinuation. Rarely reported side effects include leukopenia and other cytopenias, cutaneous and noncutaneous malignancies, and pancreatitis. Toxic doses have not been established for this medication. One patient suffered only moderate leukopenia with no significant GI side effects after ingesting 25g of mycophenolate mofetil. Up to 4 g/day have been used in cardiac and up to 5 g/day in hepatic transplantation patients. However, increased efficacy was not observed above 2 g/day, and patients were more likely to experience GI symptoms and neutropenia at higher doses. For this reason, doses up to 2 g/day are usually employed to treat inflammatory conditions.

HYDROXYUREA

Hydroxyurea is urea with one additional hydroxyl group. It inhibits ribonucleotide reductase, which catalyzes the reduction of ribonucleotides to deoxyribonucleotides, and is thus essential in DNA synthesis. It is effective in the treatment of psoriasis.[12] Hydroxyurea is well tolerated, with the most common side effects being hematological, usually megaloblastic anemia, but also leukopenia and thrombocytopenia. Other significant but rare adverse effects include renal and GI toxicity, a dermatomyositis-like syndrome, leg ulcers, radiation recall, and leukemias.

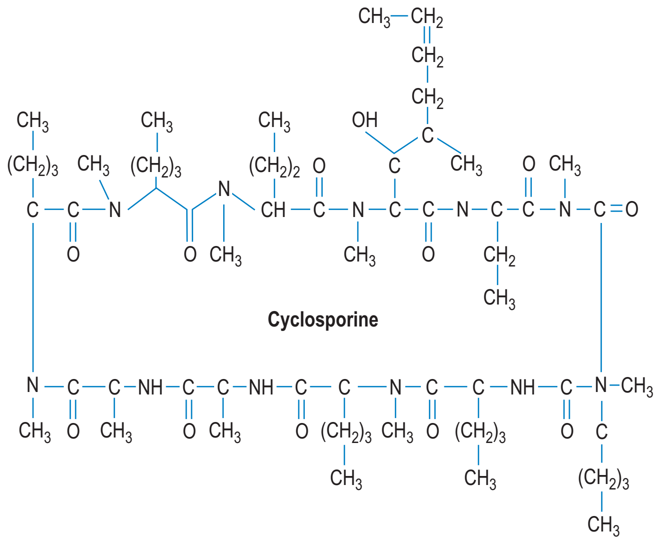


FIG 87.8 Cyclosporine—chemical structure.

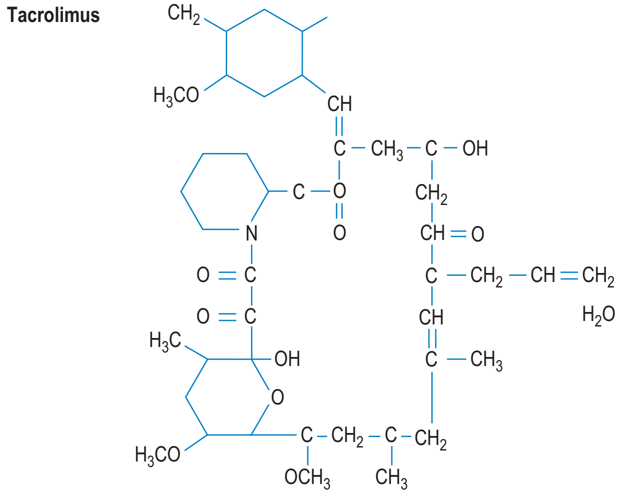


FIG 87.9 Tacrolimus—chemical structure.

ORAL CYCLOSPORINE AND TACROLIMUS (FK506)

Cyclosporine (Fig. 87.8) and tacrolimus (Fig. 87.9) are structurally similar drugs that have been widely used in solid organ transplantation as well as in the treatment of immunological diseases. Cyclosporine has potent inhibitory effects in dampening the production of proinflammatory mediators, such as IL-2, by immunocompetent cells, most importantly T lymphocytes. It does so through binding cyclophilin, which produces a cyclosporine–cyclophilin complex. This complex binds the serine/threonine phosphatase calcineurin.[13] This disrupts the phosphorylation of regulatory proteins for which the nucleated factor of activated T cells (NF-ATs) is a critical component, preventing these proteins from translocating into the nucleus. Thus the transcription of genes, such as *IL-2,* which induces mitogenesis in activated T cells, cannot be effectively activated. A number of other cytokines are affected, including IL-3, IL-6, transforming growth factor (TGF)-β, and interferon (IFN)-γ. Another T cell–specific immunophilin, FK506-binding protein (FKBP), binds tacrolimus to form a FK506-FKBP complex with similar resultant inhibitory activity on calcineurin.[14]

KEY CONCEPT

Cyclosporine: Mechanism of Action

Association with cyclophilin
Formation of cyclosporine–cyclophilin complex
Binds calcineurin
Inactivates calcineurin
Regulatory proteins unable to translocate into nucleus
Transcription of proinflammatory genes affected

Adverse Effects

The most common side effects of cyclosporine are hypertension, hyperkalemia, hypomagnesemia, and hyperlipidemia. More importantly, cyclosporine has well-documented short-term and long-term adverse effects on kidney function.[15] These data come from patients who had received solid organ transplants, most notably renal transplants. In such patients, initial doses of 15–25 mg/kg/day led to a reduction in the glomerular filtration rate (GFR) and rise in serum creatinine in a percentage of patients, as well as histologically proven nephropathy.[16] The pathogenic mechanism of kidney damage is poorly understood but is believed to consist of two phases.[17] The first is a period of partial ischemia secondary to vascular contraction, which is reversible with dose reduction or drug discontinuation. The later, irreversible phase results from chronic scarring of the glomeruli. Treatment recommendations for some diseases, such as psoriasis, have therefore been targeted at much lower maximum daily doses of 5 mg/kg/day, with a reduced dose if creatinine rises 30% above baseline. Otherwise healthy patients treated at these dosage levels have been successfully managed for many years on cyclosporine with no impact on the GFR.[18] Nephrotoxicity has also been a concern during tacrolimus therapy. Other adverse effects include increased rate of infections, malignancy, hepatotoxicity, GI upset, rash, tremor, headache, and insomnia.

TOPICAL PIMECROLIMUS AND TACROLIMUS (FK506)

Tacrolimus is also available in topical formulation. Pimecrolimus is an alternative topical calcineurin inhibitor with similar structure and identical mechanisms of action. The cyclosporine molecule is too large to penetrate skin (1203 Da), whereas tacrolimus and pimecrolimus can, as they are much smaller molecules (804 and 80 daltons [Da], respectively). Both have been approved by the FDA for the treatment of atopic dermatitis but have found widespread use in many other conditions (psoriasis, oral and cutaneous lichen planus, vitiligo, pemphigoid, and pemphigus). They are most often used when long-term topical corticosteroids are contraindicated.

Adverse Effects

The most common side effect is local irritation at the site of application in severely inflamed skin. Therefore initial short-term treatment is often combined with topical steroids. An association with malignancy remains controversial. In 2005, 17 case reports of malignancy in patients using these topicals resulted in a black-box warning, although further studies suggested that the rate of lymphoma from these case reports was lower than the rate observed in the general population in the United States. In 2006, the FDA revised the wording, but a black-box warning about these products remains in place.

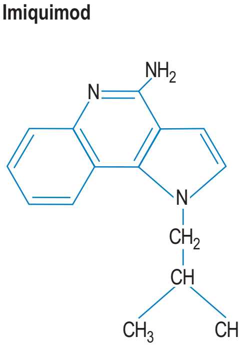


FIG 87.10 Imiquimod—chemical structure.

SIROLIMUS

Sirolimus is a macrolide that binds the cytosolic protein FK-binding protein 12 (FKBP12). In contrast to the tacrolimus–FKBP12 complex, which inhibits calcineurin, the sirolimus–FKBP12 complex directly binds the mammalian target of rapamycin (mTOR) complex1 (mTORC1), thereby inhibiting the mTOR pathway. Thus it inhibits the response to IL-2, blocking activation of T and B cells. It has shown promise in the treatment for systemic lupus erythematosus (SLE); Sjögren syndrome[19]; RA[20]; psoriasis; genetic disorders, such as tuberous sclerosis; and neoplastic disorders, such as Kaposi sarcoma.[21] The main advantage of sirolimus over tacrolimus or cyclosporine is reduced renal toxicity.

IMIQUIMOD

Imiquimod (Fig. 87.10), an imidazoquinoline drug, activates Toll-like receptor (TLR)-7 and possesses both antiviral and antitumor activities.[22] As a cream preparation, it is effective in the treatment of external genital warts caused by infection with human papillomavirus (HPV). Immune amplification responses are induced through stimulation of inflammatory cytokines.[23,24] Production of IFN-α is stimulated, and this suppresses replication of viruses in infected keratinocytes. NK-cell activity is also increased, partly through the induction of oligoadenylate synthase. The increase of dermal IFN-α transcript levels is rapid and dramatic.[25] Other cytokines modulated by imiquimod include TNF-α and IL-12, especially in peripheral blood monocytes.[26,27] The overall effect is a shift from a T-helper-2 (Th2)-cytokine–predominant profile toward a Th1-cytokine–predominant profile. Other conditions where imiquimod is effective include actinic keratosis, herpes simplex, basal cell carcinoma, and molluscum contagiosum.

Adverse Effects

Inflammation at the site of application represents the most common adverse reaction. However, nondermatological side effects, such as fever, fatigue, myalgia, and headache, have also been described.

5-FLUOROURACIL

5-Fluorouracil is a uracil analogue that has two modes of action. First, it inhibits cell proliferation via direct incorporation into RNA causing abnormal base pairing. Second, it binds thymidylate synthetase, blocking the conversion of deoxyuridine monophosphate to deoxythymidine monophosphate, which is essential to DNA synthesis. It may also increase expression of *p53,* a frequently mutated gene in nonmelanoma skin cancers.[28] It can be administered topically, intramuscularly, or intravenously. In topical form, its main uses are to treat actinic keratosis, superficial basal cell carcinoma, Bowen disease, keratoacanthoma, porokeratosis, and verruca vulgaris. However, intravenous treatment of recalcitrant psoriasis, mycosis fungoides, and scleroderma has also been reported.

Adverse Effects

Topical application is associated with an irritant dermatitis, but this is also seen as a marker of clinical efficacy. Parenteral administration for inflammatory conditions is not widespread; adverse effects from parenteral administration are more severe and include clinically significant bone marrow suppression, GI toxicity, and cutaneous reactions.

GLATIRAMER

Glatiramer acetate is a random polymer of glutamic acid, lysine, tyrosine, and alanine, the four amino acids found in myelin basic protein (MBP). The mechanism of action for glatiramer is unknown, but its similarity in structure to MBP may allow it to act as a decoy for immune targeting of myelin. It may also induce expression of glatiramer acetate–specific suppressor T cells, and these have been shown to be present in animal models. In contrast to imiquimod, glatiramer shifts the population of T cells from proinflammatory Th1 cells to regulatory Th2 cells that suppress the inflammatory response. It is FDA approved for the treatment of adults with relapsing–remitting multiple sclerosis (MS), even after only one event.

Adverse Effects

Absolute contraindications include allergy to glatiramer or to mannitol. It is pregnancy Category B, but it is unknown if the drug is secreted in breast milk. The drug is given via subcutaneous injection, and the most common adverse effects are injection site reaction, flushing, rash, dyspnea, and transient chest pain.

FINGOLIMOD (FTY720)

Fingolimod is a relatively new immunomodulator for treating MS. It is a structural analogue of sphingosine and is phosphorylated intracellularly by sphingosine kinases.[29] Signaling via one of the sphingosine 1 phosphate receptors, S1PR1, it is believed to prevent migration of lymphocytes by halting their ability to leave lymph nodes. However, fingolimod has also been reported to act as a cPLA2 inhibitor,[30] a cannabinoid receptor antagonist,[31] and a ceramide synthase inhibitor.[32] So far, the FDA has only approved it for use in MS. However, it has also shown promise in murine models of SLE and RA, with reduced mortality from lupus nephritis[33] and improvement in arthritis, respectively.[34] It also has a potential role in the treatment of cutaneous inflammatory conditions, such as psoriasis and atopic dermatitis.[35]

Adverse Effects

The most common side effects are minor and include headache and fatigue. However, fingolimod has also been associated with serious infections, an increased rate of skin cancers, bradycardia, and one case of focal hemorrhagic encephalitis.

CONCLUSIONS

The field of small molecule immunomodulators contains both traditional and new molecules. The traditional molecules come with a long history of efficacy and side-effect data, and this allows us to safely and accurately utilize these therapies. Despite the introduction of biologics, traditional small-molecule agents, such as methotrexate and cyclosporine, are both clinically efficacious and cost-effective and thus remain the workhorses of our modern pharmaceutical armamentarium. However, as our understanding of the pathways involved in inflammation evolve, our therapeutics are also refined, allowing the production now, for example, of topically effective calcineurin inhibitors, such as tacrolimus and pimecrolimus, which avoid the adverse effects of systemic cyclosporine or topical steroids, and sirolimus, which reduces toxicity profiles of systemic FKBP2 signaling. In addition, newer agents, such as glatiramer and fingolimod, allow treatment of MS, a disease recalcitrant to more traditional therapies, and, although not currently being used in other inflammatory disorders, have shown promise in animal models. The field of small-molecule immunomodulators therefore remains one of both current clinical relevance and continuing, exciting new developments.

ON THE HORIZON

Refinement of our understanding of the pathways and actions of receptors targeted by our current medications is the key to producing more elegant therapies.

Separating therapeutic effects from the pathways leading to side effects would enable much safer and more tolerable small-molecule medications.

Better understanding of the dysregulation that occurs in inflammatory conditions, and the activation of the immune system will be essential to ensuring the efficacy of such new modalities.

The challenge in the next 5–10 years is to continue this refinement in the targeting of small molecules. Many of the agents we currently use target early parts of pathways or indiscriminately trigger multiple receptors. This leads to a greater range of side effects and thus limits our ability to reach adequate dosing levels. A prime example of this is methotrexate, which leads to increased adenosine levels and thus triggers all four adenosine receptors. By developing more selective molecules targeting, for instance, just the adenosine A_{2A} receptor, not only would we avoid the side effects attributable to other adenosine receptors, we might also avoid the effects of methotrexate on bone marrow and mucous membranes via inhibition of dihydrofolate reductase. In other medications, more specific targeting of pathways may similarly reduce side effects and make for more effective therapeutics. This will require a better understanding of the immune system and its dysregulation in disease, as well as development and application of these medications to clinically relevant models

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Pincus T, Furer V, Sokka T. Underestimation of the efficacy, effectiveness, tolerability, and safety of weekly low-dose methotrexate in information

presented to physicians and patients. Clin Exp Rheumatol 2010;28:S68–79.

2. Chan ES, Cronstein BN. Molecular action of methotrexate in inflammatory diseases. Arthritis Res 2002;4:266–73.
3. Cronstein BN, Naime D, Ostad E. The anti-inflammatory mechanism of methotrexate. Increased adenosine release at inflamed sites diminishes leukocyte accumulation in an in vivo model of inflammation. J Clin Invest 1993;92:2675–82.
4. Baggott JE, Morgan SL, Sams WM, et al. Urinary adenosine and aminoimidazole carboxamide excretion in methotrexate-treated patients with psoriasis. Arch Dermatol 1999;135:813–17.
5. Chan ES, Montesinos MC, Fernandez P, et al. Adenosine A(2A) receptors play a role in the pathogenesis of hepatic cirrhosis. Br J Pharmacol 2006;148:1144–55.
6. Amos RS, Pullar T, Bax DE, et al. Sulphasalazine for rheumatoid arthritis: toxicity in 774 patients monitored for one to 11 years. Br Med J (Clin Res Ed) 1986;293:420–3.
7. Lind MJ, Ardiet C. Pharmacokinetics of alkylating agents. Cancer Surv 1993;17:157–88.
8. Wolverton SE, Remlinger K. Suggested guidelines for patient monitoring: hepatic and hematologic toxicity attributable to systemic dermatologic drugs. Dermatol Clin 2007;25:195–205, vi–ii.
9. Kremer JM, Genovese MC, Cannon GW, et al. Concomitant leflunomide therapy in patients with active rheumatoid arthritis despite stable doses of methotrexate. A randomized, double-blind, placebo-controlled trial. Ann Intern Med 2002;137:726–33.
10. Weinblatt ME, Kremer JM, Coblyn JS, et al. Pharmacokinetics, safety, and efficacy of combination treatment with methotrexate and leflunomide in patients with active rheumatoid arthritis. Arthritis Rheum 1999;42:1322–8.
11. Nagy SE, Andersson JP, Andersson UG. Effect of mycophenolate mofetil (RS-61443) on cytokine production: inhibition of superantigen-induced cytokines. Immunopharmacology 1993;26:11–20.
12. Menter A, Korman NJ, Elmets CA, et al. Guidelines of care for the management of psoriasis and psoriatic arthritis: section 4. Guidelines of care for the management and treatment of psoriasis with traditional systemic agents. J Am Acad Dermatol 2009;61:451–85.
13. Liu J, Farmer JD Jr, Lane WS, et al. Calcineurin is a common target of cyclophilin-cyclosporin A and FKBP-FK506 complexes. Cell 1991;66:807–15.
14. Baughman G, Wiederrecht GJ, Campbell NF, et al. FKBP51, a novel T-cell-specific immunophilin capable of calcineurin inhibition. Mol Cell Biol 1995;15:4395–402.
15. Wilkinson A, Ross EA, Hawkins R, et al. Measurement of true glomerular filtration rate in renal transplant patients receiving cyclosporine. Transplant Proc 1987;19:1739–41.
16. Mihatsch MJ, Antonovych T, Bohman SO, et al. Cyclosporin A nephropathy: standardization of the evaluation of kidney biopsies. Clin Nephrol 1994;41:23–32.
17. Gaston RS. Chronic calcineurin inhibitor nephrotoxicity: reflections on an evolving paradigm. Clin J Am Soc Nephrol 2009;4:2029–34.
18. Kessel A, Toubi E. Cyclosporine-A in severe chronic urticaria: the option for long-term therapy. Allergy 2010;65:1478–82.
19. Perl A. Emerging new pathways of pathogenesis and targets for treatment in systemic lupus erythematosus and Sjogren's syndrome. Curr Opin Rheumatol 2009;21:443–7.
20. Laragione T, Gulko PS. mTOR regulates the invasive properties of synovial fibroblasts in rheumatoid arthritis. Mol Med 2010;16: 352–8.
21. Paghdal KV, Schwartz RA. Sirolimus (rapamycin): from the soil of Easter Island to a bright future. J Am Acad Dermatol 2007;57:1046–50.
22. Hemmi H, Kaisho T, Takeuchi O, et al. Small anti-viral compounds activate immune cells via the TLR7 MyD88-dependent signaling pathway. Nat Immunol 2002;3:196–200.
23. Dahl MV. Imiquimod: an immune response modifier. J Am Acad Dermatol 2000;43:S1–5.
24. Skinner RB Jr. Imiquimod. Dermatol Clin 2003;21:291–300.
25. Imbertson LM, Beaurline JM, Couture AM, et al. Cytokine induction in hairless mouse and rat skin after topical application of the immune response modifiers imiquimod and S-28463. J Invest Dermatol 1998;110:734–9.
26. Gibson SJ, Imbertson LM, Wagner TL, et al. Cellular requirements for cytokine production in response to the immunomodulators imiquimod and S-27609. J Interferon Cytokine Res 1995;15:537–45.
27. Tyring S. Imiquimod applied topically: a novel immune response modifier. Skin Therapy Lett 2001;6(6):1–4.
28. Ceilley RI. Mechanisms of action of topical 5-fluorouracil: review and implications for the treatment of dermatological disorders. J Dermatolog Treat 2012;23:83–9.
29. Billich A, Bornancin F, Devay P, et al. Phosphorylation of the immunomodulatory drug FTY720 by sphingosine kinases. J Biol Chem 2003;278:47408–15.
30. Payne SG, Oskeritzian CA, Griffiths R, et al. The immunosuppressant drug FTY720 inhibits cytosolic phospholipase A2 independently of sphingosine-1-phosphate receptors. Blood 2007;109:1077–85.
31. Paugh SW, Cassidy MP, He H, et al. Sphingosine and its analog, the immunosuppressant 2-amino-2-(2-[4-octylphenyl]ethyl)-1,3-propanediol, interact with the CB1 cannabinoid receptor. Mol Pharmacol 2006;70:41–50.
32. Berdyshev EV, Gorshkova I, Skobeleva A, et al. FTY720 inhibits ceramide synthases and up-regulates dihydrosphingosine 1-phosphate formation in human lung endothelial cells. J Biol Chem 2009;284:5467–77.
33. Ando S, Amano H, Amano E, et al. FTY720 exerts a survival advantage through the prevention of end-stage glomerular inflammation in lupus-prone BXSB mice. Biochem Biophys Res Commun 2010;394:804–10.
34. Tsunemi S, Iwasaki T, Kitano S, et al. Effects of the novel immunosuppressant FTY720 in a murine rheumatoid arthritis model. Clin Immunol 2010;136:197–204.
35. Herzinger T, Kleuser B, Schafer-Korting M, et al. Sphingosine-1-phosphate signaling and the skin. Am J Clin Dermatol 2007;8: 329–36.

MULTIPLE-CHOICE QUESTIONS

1. Which of the following medications' antiinflammatory effects may be reduced in heavy coffee drinkers?
   A. Methotrexate
   B. Azathioprine
   C. Mycophenolate mofetil
   D. Cyclosporine
   E. Tacrolimus

2. Which of the following medications is pregnancy Category B?
   A. Methotrexate
   B. Mycophenolate mofetil
   C. Glatiramer
   D. Azathioprine
   E. 5-Florouracil

3. In which of the following medicines should glucose-6-phosphate dehydrogenase (G6PD) screening be considered before use?
   A. Methotrexate
   B. Azathioprine
   C. Sulfasalazine
   D. Mycophenolate mofetil
   E. Imiquimod

88

Protein Kinase Antagonists in Therapy of Immunological and Inflammatory Diseases

Arian Laurence, Massimo Gadina, John J. O'Shea

Reversible phosphorylation is one of the major mechanisms controlling protein activity in all eukaryotic cells and, as such, is involved in all fundamental cellular processes, including cell cycle and cell growth, cell shape and movement, metabolism, differentiation, and apoptosis. This covalent modification is a major means for transmitting information from outside the cell and between the subcellular components within the cell. Phosphorylation is a major mechanism underlying normal signaling, as exemplified by insulin and other growth factors, but in addition, the importance of protein phosphorylation is supported by evidence that mutations and dysregulation of protein kinases play causal roles in human disease. This is especially true in cancer, in which mutant protein kinases or their upstream activators function as oncogenes.

From the point of view of an immunologist, protein phosphorylation is also the major mechanism by which immune receptors, including the T-cell receptor (TCR), B-cell receptor, and natural killer (NK) cell receptor and Fc receptors, trigger signaling. As discussed in Chapter 4, the first step in signaling by multichain immune recognition receptors (*e.g.*, the aforementioned receptors) is tyrosine phosphorylation of the receptor itself and adapter molecules, such as linker of activated T cells (LAT), mediated by Src family protein tyrosine kinases (PTKs). This leads to the recruitment of PTK members spleen tyrosine kinase (Syk) and Zap70 followed by phosphorylation of adapters, such as SH2 domain-containing leukocyte phosphoprotein of 76 kDa (SLP-76) and the activation of Tec family PTKs. These initial steps trigger the activation of serine-threonine kinases, including the mitogen-activated protein kinases (MAPKs) and protein kinase C (PKC) family (Fig. 88.1). We now know a great deal of details about how the cascade of protein phosphorylation links events at the plasma membrane to calcium modulation, cytoskeletal rearrangement, gene transcription, and other canonical features of lymphocyte action (Chapter 12).

Similarly, a critical first step in signaling by many cytokine receptors is the activation of phosphorylation. The receptors for classical growth factor cytokines, including stem cell factor and platelet-derived growth factor (PDGF) are receptor tyrosine kinases (RTKs), whereas the receptors for transforming growth factor family cytokines are receptor serine–threonine kinases. Type I and II cytokine receptors signal via the activation of receptor-associated Janus kinases (JAKs; see below, Table 88.1 and Table 88.2) (Chapter 9). Other cytokines like interleukin (IL)-1 and tumor necrosis factor (TNF) initiate signaling in a kinase-independent manner but, nonetheless, signal through kinase cascades to exert their effects. It is clear that many aspects of protein phosphorylation are of major importance in immune and inflammatory mechanisms.

The nonredundant functions of various kinases in different immune cells are exemplified by both studies in knock-out mice and humans with mutations. On the basis of these findings, targeting protein kinases has been proposed to be a useful strategy in the development of novel immunosuppressant drugs and is one of the most active areas of pharmaceutical drug development (Table 88.3), with much of the impetus coming from oncology. The field is so vast that it is impractical to comprehensively review all this information in one chapter; therefore we will focus both on important historical precedents in the field and then discuss drugs and targets that are most immunologically relevant. We will start by briefly reviewing some of the basics of kinase biochemistry.

STRUCTURE AND FUNCTION OF PROTEIN KINASES

KEY CONCEPTS

Kinase Families

- 518 kinases in the genome
- 90 protein tyrosine kinases (*e.g.*, Janus kinases [JAKs])
- 400 protein serine/ threonine kinases:
  - AGC kinase family (*e.g.*, protein kinase B/AKT)
  - CAMK kinase family (*e.g.*, calmodulin-dependent kinase)
  - CMGC kinase family (*e.g.*, MAP kinases: ERK, JNK, p38)
  - STE kinase family (*e.g.*, MAPK kinases, MAPKK kinases)
  - TKL kinase family (*e.g.*, IRAK)
  - Others: casein kinase family, GYC kinase family, IKK family

Protein kinases or phosphotransferases catalyze the transfer of the γ-phosphate from a purine nucleotide triphosphate (*i.e.*, adenosine triphosphate [ATP] and guanosine triphosphate [GTP]) to the hydroxyl groups of their protein substrates. They generate phosphate monoesters by using protein alcohol groups (on serine and threonine residues) and/or protein phenolic groups (on tyrosine residues) as phosphate acceptors. Thus protein kinases can be classified by the amino acid substrate preference: serine/threonine kinases, tyrosine kinases, and dual kinases (*i.e.*, both serine/threonine and tyrosine residues can be phosphorylated).

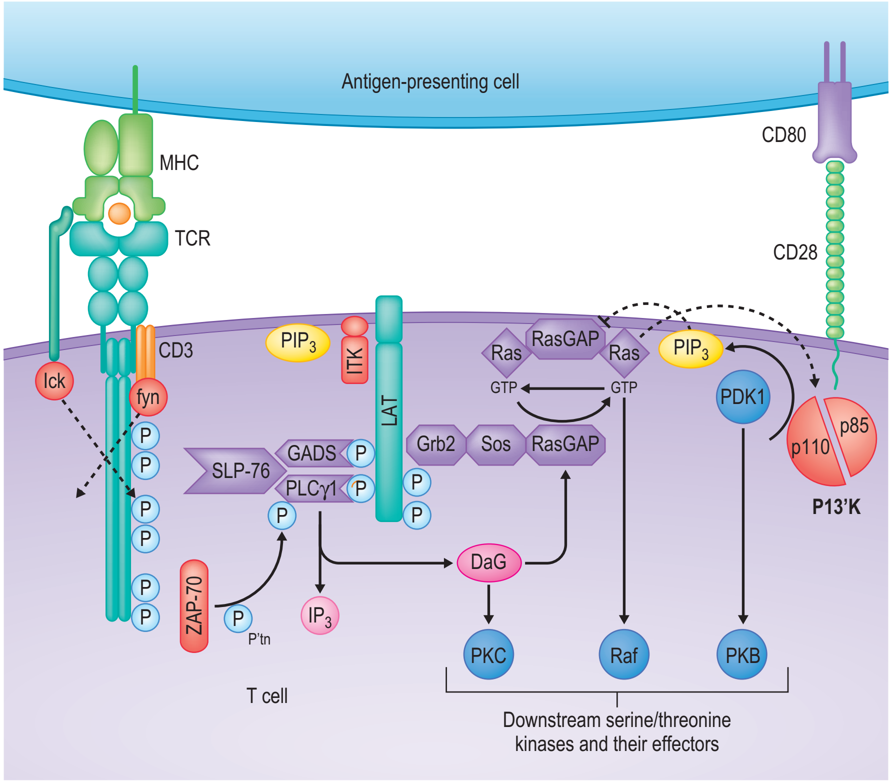

Antigen-presenting cell
MHC
TCR
CD80
CD28
CD3
PIP$_3$
ITK
LAT
RasGAP
Ras
Ras
PIP$_3$
lck
fyn
PDK1
GTP
GTP
P
GADS
SLP-76
PLCγ1
Grb2
Sos
RasGAP
p110
p85
P13'K
DaG
ZAP-70
P'tn
IP$_3$
PKC
Raf
PKB
T cell
Downstream serine/threonine kinases and their effectors

FIG 88.1 The Proximal Signaling Events in Response to T-Cell Receptor (TCR) Activation by an Antigen-Presenting Cell. Tyrosine and lipid kinases are indicated in red, and serine/ threonine kinases are indicated in blue.

TABLE 88.1 Functions and Signaling of Cytokines Sharing the Common γ Chain

Cytokine Receptor	IL-2R	IL-4R	IL-7R	IL-9R	IL-15R	IL-21R
Functions	Control of peripheral self-tolerance (mice)	Regulation of B-cell function (in concert with interleukin [IL]-21 and IL-25)	Thymocyte survival and development factor	Goblet cell hyperplasia	CD8 memory T-cell survival and proliferation factor	T- and B-cell proliferation factor
	Development and maintenance of regulatory T cells	Immunoglobulin class switching	Peripheral T-cell survival factor Mediates homeostatic reconstitution of lymphopenic animals	Mucus production	Peripheral T cell survival factor Mediates homeostatic reconstitution of lymphopenic animals	Regulates immunoglobulin production (with IL-4)
	Differentiation of helper and cytotoxic T cells	Differentiation of T helper cells (Th2 lineage)	B-cell progenitor survival factor (mice)		Natural killer (NK) cell development, differentiation and survival factor	NK cell proliferation and activation factor
	In vitro expansion and differentiation of antigen-selected T and NK cells	Costimulant for growth in T, B, and mast cells				
		Inhibition of Th1 differentiation and macrophage activation				
Downstream signaling pathways	PI3K, Ras MAPK, STAT1,3,5	PI3K, Ras MAPK, STAT6,5	PI3K, STAT(1), (3), 5	PI3K, Ras MAPK, STAT1,3,(5)	PI3K, Ras MAPK, STAT1,3,5	PI3K, Ras MAPK, STAT1,3,(5)

TABLE 88.2 The JAK Family of Tyrosine Kinases

Gene	Murine Phenotype Associated With Gene Deletion	Associated Receptor
Jak1	Perinatal lethality, block in thymocyte development	Many including IL-7R and IFNRs
Jak2	Embryonic lethality due to anemia	Erythropoietin receptor
Jak3	T-, B-, and NK-cell lymphopenia, SCID phenotype	Common gamma chain receptor
Tyk2	Failure to clear *Toxoplasma,* reduced arthritis	IL-12/23 receptor

JAK, Janus kinase; IL, interleukin; IFNR, interferon receptor; NK, natural killer; SCID, severe combined immune deficiency.

Almost all protein kinases have catalytic domains that belong to a single eukaryotic protein kinase (ePK) superfamily. The common evolutionary ancestry of the kinase domain (also known as the *catalytic domain*), which consists of 250–300 amino acid residues, manifests as a highly conserved three-dimensional structure. In the human genome, there are 518 kinases, which are divided into eight major groups, in totality, representing 1.7% of the human genome. The PTK family has 90 members, one-third of which are RTKs, and the remainder are cytoplasmic proteins that typically function in close proximity to, and downstream of, receptor–ligand complexes.

In terms of its catalytic role, the kinase domain has three functions: the binding of the ATP (or GTP) phosphate donor as a complex with a divalent cation (usually Mg^{2+} or Mn^{2+}), the

TABLE 88.3 Selected Kinase Inhibitors and Related Drugs

Mechanism	Compound	Kinase Inhibited	Comments
Direct binding to the kinase	Imatinib	Kit, platelet-derived growth factor (PDGFR), and Abl	Approved for treatment of chronic myeloid leukemia (CML), eosinophilic leukemia, and gastric stromal tumors Used in the treatment of chronic sclerodermatous graft-versus-host disease (GvHD) Under evaluation in multiple types of cancer and in combination with other cancer drugs
	Ponatinib		US Food and Drug Administration (FDA) approved for the treatment of CML and PH + acute lymphoblastic leukemia (ALL) Phase II trial for acute myeloid leukemia (AML) and a number of solid cancers
	Nilotinib		Approved for Philadelphia chromosome-positive CML; in Phase II trials for AML and ALL
	Radotinib	PDGFR and Abl	FDA approved for the treatment of CML
	Dasatinib	Multiple Src family tyrosine kinases, including Abl	Approved for treatment of adult CML and Philadelphia chromosome positive ALL Phase III trial for prostate cancer Phase II trial for CLL, Sarcomas
	Bosutinib		Approved for the treatment of imatinib-resistant CML Phase II trial for glioblastoma
	Ibrutinib	Btk, Itk	FDA approved for the treatment of CLL, mantle-cell lymphoma (MCL), lymphoplasmacytic lymphoma (LPL) Phase III trial for the treatment of diffuse large B-cell lymphoma (DLBCL), posttransplantation lymphoproliferative disease (PTLD), pancreatic cancer Phase II trial for GvHD
	Spebrutinib		Phase II trial for RA, Phase I trial for CLL, non-Hodgkin lymphoma (NHL)
	Acalabrutinib		Phase III trial for CLL, phase II trial for rheumatoid arthritis (RA) FDA approved for the treatment of mantle cell lymphoma (MCL)
	Lapatinib	ErbB-2 and epidermal growth factor receptor (EGFR) kinases	Approved for treatment of breast cancer Minor survival improvement in head and neck cancer Phase III trials for brain metastasis
	Afatinib		FDA approved for treatment of non–small cell lung carcinoma (NSCLC) Phase III trial for breast cancer and head and neck cancer
	Neratinib		Phase III trial for treatment of breast cancer Phase II trial for NSCLC
	Gefitinib		Approved for use in NSCLC as a first-line therapy by FDA Successful phase III study in esophageal cancer
	Erlotinib	HER1/EGFR, JAK2	Approved for treatment of pancreatic cancer and certain categories of NSCLC In phase II and III trials for various other cancers, including myeloproliferative disease
	Vandetanib	VEGFR-2, EGFR, RET, and ErbB-1 tyrosine kinases	Approved in 2011 for certain types of progressive medullary thyroid cancer Phase II trial for small cell lung cancer
	Tivozanib	Vascular endothelial growth factor receptor (VEGFR) kinases	Phase I trial for solid tumors
	Lenvatinib		FDA approved for thyroid cancer and under phase III trials for hepatocellular cancer
	Regorafenib		FDA approved for the treatment of Colorectal cancer and gastrointestinal stromal tumor (GIST) Phase III trials for hepatocellular cancer
	Brivanib	VEGFR-2, FGFR	Orphan drug status for the treatment of hepatocellular cancer
	(N)intedanib	VEGFR, PDGFR, FGFR	Approved for the treatment of idiopathic pulmonary fibrosis and NSLCLC adenocarcinoma In phase III trials for the treatment of ovarian and colorectal carcinoma

Continued

TABLE 88.3 Selected Kinase Inhibitors and Related Drugs—cont'd

Mechanism	Compound	Kinase Inhibited	Comments
	Cabozantinib	VEGFR, MET, RET	FDA approved for treatment of thyroid cancer Phase III registration for hepatocellular cancer
	Axitinib	VEGFR1, 2 and 3, Abl (T315I), PDGFR-β and Kit	Approved for treatment of renal cancer; has orphan drug status for follicular, medullary, and anaplastic thyroid carcinoma and metastatic or locally advanced papillary thyroid cancer Effective in imatinib resistant CML associated with the T315I mutation
	Sunitinib	Inhibits multiple receptor tyrosine kinases (RTKs), including PDGFR3-alpha, PDGFR-β, VEGFR1, VEGFR2, VEGFR3, Kit, Flt3, CSF-1R, and RET	For treatment of gastrointestinal stromal tumor after disease progression on or intolerance of imatinib and for treatment of advanced renal cell carcinoma and pancreatic neuroendocrine tumors Completed phase II trials in NSCLC and breast cancer Potential benefit in Flt3 + AML
	Sorafenib	Dual-specific inhibitor blocking both tyrosine and serine/threonine kinases, including RAF kinase, VEGFR-2, VEGFR-3, PDGFR-B, KIT, Flt3, and RET	FDA approved for treatment of advanced renal cell carcinoma, thyroid cancer, and hepatocellular cancer In phase II trials for AML, ALL, CML, myelodysplastic syndrome (MDS), neurofibromatosis, portal hypertension
	Pazopanib	Inhibits multiple RTKs, including PDGFR3-α, PDGFR-β, VEGFR1 VEGFR2, VEGFR3, cKIT, Lck, c-FMS, FGFR-1, and FGFR-3	Approved for treatment of advanced renal cell cancer and soft tissue sarcoma Mixed results as a treatment for age related macular degeneration when used topically
	Entrectinib	ALK, TrkA, TrkB, TrkC, ROS1	Phase II trial for colorectal cancer, NSCLC Phase I trial for neuroblastoma
	Crizotinib	ALK, c-Met	Approved for treatment of ALK-positive NSCLC Phase II trial for anaplastic lymphoma kinase (ALK)–positive lymphoma
	Filgotinib	JAK 1	Completed phase IIb studies in RA Completed phase II studies in Crohn disease
	ABT-494		In phase III trial for RA and phase II trial for Crohn disease
	Ruxolitinib	JAK1 and JAK 2	FDA approved for myeloproliferative diseases In trials for treatment of GvHD
	Baricitinib		Completed phase III trials for rheumatoid arthritis Completed phase IIb trials in psoriasis
	Gandotinib (LY2784544)		In phase II clinical trials for myeloproliferative disorders
	Oclacitinib		FDA approved for the treatment of canine allergic dermatitis
	AC-410		In phase I clinical trials for autoimmune and myeloproliferative syndrome
	Tofacitinib	JAK3, JAK1, (JAK2)	FDA approved for the treatment of RA Completed phase III trials for the treatment of ulcerative colitis (UC)and psoriasis
	Peficitinib		Phase IIb for RA and UC
	Decernotinib	JAK 3	Completed phase III studies in the treatment of RA
	Momelotinib	JAK1, JAK2, IKKe, and TBK1	In phase III trials for pancreatic adenocarcinoma
	TG02/ SB-1317	JAK2, Flt3, CDK1, -2, -7, -9	In phase I clinical trials for ALL, AML, CLL
	Fedratinib	JAK2, Flt3, RET	In phase II trial for chronic beryllium disease
	AT-9283	JAK2, JAK3, aurora A/B kinase, highly active against the Gleevec-resistant T315I Abl mutation	Completed phase I trial for solid tumors in children
	Lestaurtinib	Flt3, TrkA, and JAK2	In phase III clinical trials for treatment of patients with AML who have an Flt3-activating mutation at first relapse from standard induction chemotherapy In phase II trials for psoriasis and pancreatic cancer

TABLE 88.3 Selected Kinase Inhibitors and Related Drugs—cont'd

Mechanism	Compound	Kinase Inhibited	Comments
	Quizartinib	Class III receptor tyrosine kinases (Flt3, Kit, CSF1)	Completed phase II clinical trial for drug-resistant AML
	Fostamatinib	Syk, Flt3	Phase III for idiopathic thrombocytopenic purpura (ITP) Phase II for immunoglobulin A (IgA) nephropathy Failed phase III for RA, phase I for GvHD
	PRT-062607/ BIIB057/ P505-15	Syk	In phase I clinical trials for RA, CLL, NHL; preclinical for systemic lupus erythematosus (SLE)
	R-348	JAK/Syk	In phase II trials for ocular GvHD
	Vemurafenib	BRAF	FDA approved for the treatment of BRAF mutated melanoma; phase II trial for colorectal cancer and myeloma
	Dabrafenib		FDA approved for the treatment of BRAF mutated melanoma
	Selumetinib	MKK family	Phase III trial for NSCLC and thyroid cancer Phase II for colorectal cancer and biliary cancer
	Cobimetinib		FDA approved for the treatment of melanoma Phase II trial for the treatment of breast cancer and Langerhans cell histiocytosis
	Trametinib		FDA approved for the treatment of BRAF mutated melanoma in combination with dabrafenib
	Binimetinib		Phase III trial for the treatment of BRAF mutated melanoma
	PD-325901		Phase II trial for neurofibromatosis
	CC 90003	ERK	Phase I trial for solid tumors
	XG102	JNK	Phase III trial for ocular inflammation and sensorineural hearing loss
	RES-529	AKT, MTOR, VEGFR	Phase I trial for age-related macular degeneration
	Uprosertib	AKT	Phase II trial for AML, solid tumors
	MK-2206		Phase II trial for solid and hematological malignancies
	AZD-5363		Phase II trial for breast and gastric cancer
	Pexidartinib	CSF-1R, Kit	Phase II trial for pigmented villonodular synovitis
	AZD 2014	MTORC1 and C2	Phase II trial for DLBCL and solid tumors
	CC 223		Phase II trial for NHL and solid tumors
	Sotrastaurin	PKC	Phase II trial for DLBCL
	Ribociclib	CDK4,6	Phase III trial for breast cancer Phase II trial for gastrointestinal cancer
	Palbociclib		FDA approved for the treatment of breast cancer Phase III trial for NSCLC
	Buparlisib	PI3K	Phase III trial for breast cancer
	Bardoxolone	IKbK, COX2, NOSII, STAT3	Phase II trial for diabetic neuropathy and pulmonary hypertension
	Fasudil	ROCK1/ ROCK2	Phase III for Raynaud phenomenon and scleroderma completed
	KD-025	ROCK2	Phase II for psoriasis vulgaris Phase II for idiopathic pulmonary fibrosis
Indirect binding to kinase	Sirolimus (Rapamycin)	mTORC1	FDA licensed for use in solid organ and bone marrow transplantation
	Everolimus		FDA licensed for use with cyclosporine in cardiac and renal transplantation FDA licensed for use in neuroendocrine tumors and renal cell carcinoma
	Temsirolimus		FDA licensed for use with MCL, renal cancer
	Ridaforolimus		Phase II trial for coronary artery restenosis
Monoclonal antibodies binding to receptor tyrosine kinases	Trastuzumab	EGFR	FDA licensed for use in HER2/neu positive breast carcinoma and metastatic gastric cancer
	Cetuximab		FDA licensed for use in relapsed colorectal and head/neck cancers
	Bevacizumab	VEGF—prevents binding to its receptor (Flt-1)	FDA licensed for use in metastatic breast and colorectal cancers, glioblastoma, and NSCLC (non-squamous cell histology) in combination with traditional chemotherapy
Other mechanisms of action	Barasertib	Aurora kinase B (functions in the attachment of the mitotic spindle to the centromere)	Disrupts mitosis and cellular division in tumor cells; in phase III clinical trials for AML

binding of the protein substrate, and the transfer of the γ-phosphate from ATP or GTP to the protein substrate. Despite the huge number of serine/threonine and tyrosine kinases, there is evidence of a common ancestor, and this is reflected in structural similarities, particularly in the active (ATP bound) confirmation. The major kinase domains of all typical protein kinases consists of two lobes (N-lobe and C-lobe) that surround the nucleotide binding site[1] (Fig. 88.2). The smaller N-lobe consists of a cluster of β-pleated sheets with a single α helix. The larger C-lobe is made up of α helices. Within the C-lobe lies the substrate-binding site, typically a groove on the surface. A hinge region connects the two lobes. The hinge together with two loops emerging from each lobe form the ATP-binding pocket: the primary target for most kinase inhibitors. In many protein kinases, a loop emerging from the C-lobe must be phosphorylated for the kinase to be fully active (see Fig. 88.2). This is known as the *activation loop.* Substrates of PTKs often include the activation loop of downstream kinases, creating signaling cascades of proteins, which, in turn, phosphorylate each other; examples include the MAPKs (Fig. 88.3).

THE DISCOVERY OF KINASE INHIBITORS

Given that protein kinases bind ATP, the notion that therapeutically useful kinase inhibitors could be generated was initially met with some skepticism. First, as there are more than 500 human kinases, many of which serve critical cellular functions, would it really be possible to attain the specificity needed? Second, protein kinases are not the only kinases—there are lipid kinases and nucleotide kinases, as well as many other ATP-binding

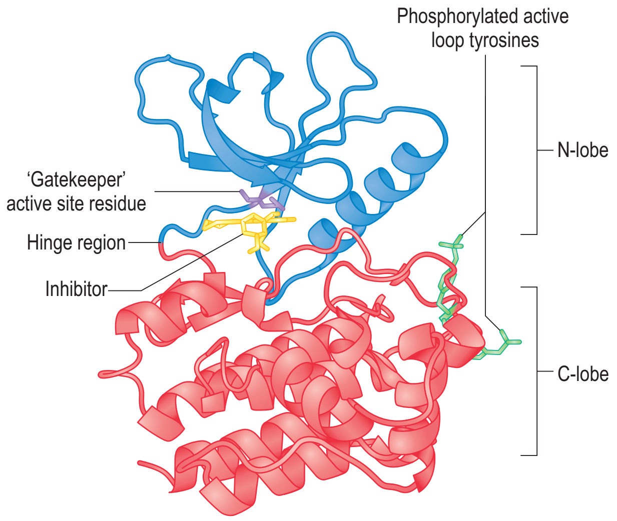

Phosphorylated active loop tyrosines
N-lobe
'Gatekeeper' active site residue
Hinge region
Inhibitor
C-lobe

FIG 88.2 Crystal Structure of the Janus Kinase 3 (JAK3) Domain Complexed With Staurosporine (pdb Accession Code 1YVJ). This structure captures the active conformation of JAK3 with both active loop tyrosine residues phosphorylated (green). The molecule can be described as halves, with the N-terminal lobe presented in blue and the C-terminal domain in red. These are linked by a hinge region that forms part of the active site. Highlighted in purple within the active site is the gatekeeper residue. Bound within this site is an analogue of the inhibitor staurosporine, and its proximity to the "gatekeeper" residue highlights why this residue and this region are critical for the specificity of inhibitors for individual protein kinases.

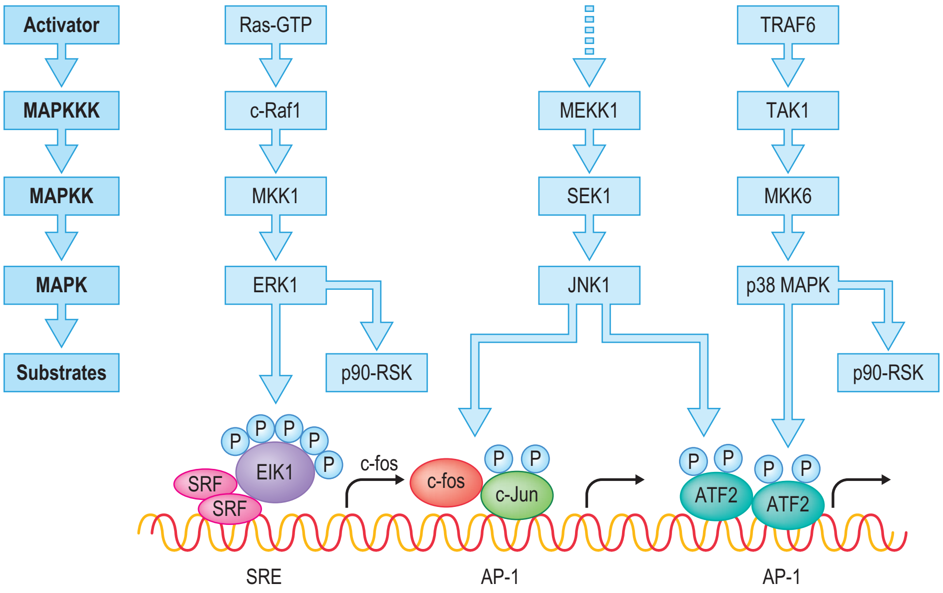


FIG 88.3 A Summary of the Mitogen-Activated Protein Kinase (MAPK) Signal Transduction Pathways. Examples of receptors that activate Ras include the interleukin (IL)-2 and T-cell receptor (TCR). Examples of receptors that activate tumor necrosis factor receptor–associated factor 6 (TRAF6) include the IL-1 receptor.

proteins, all of which share structural similarities with PTKs. Third, despite the many potential ways of designing a small-molecule kinase inhibitor, in practice, the majority work by sitting within the ATP-binding pocket.[2] *A priori* then, one might conclude that it would be impossible to generate an antagonist that did not target some other essential ATP-dependent process. Fortunately though, this skeptical view does not reflect reality.

THERAPEUTIC PRINCIPLES

Protein Kinase Inhibitors As Drugs

- Drugs can inhibit protein kinases with a high degree of specificity.
- Multi-kinase inhibitors can be well tolerated and be more efficacious than single kinase inhibitors (*e.g.*, second-generation Abl kinase inhibitors).
- Conversely, unacceptable toxicity has limited the creation of clinical inhibitors to particular kinase members (*e.g.*, p38 MAPK inhibitors).

Imatinib and Other First-Generation Protein Tyrosine Kinases Inhibitors

The first protein kinase inhibitor approved by the US Food and Drug Administration (FDA) is imatinib (see Table 88.3). The mutated form of the Abl tyrosine kinase BCR-Abl represents a fusion protein that is the result of a chromosomal translocation (Philadelphia chromosome) observed in patients suffering from chronic myeloid leukemia (CML). The pathognomonic presence of BCR-Abl in CML has led to it becoming one of the most intensively studied PTKs. The fusion protein consists of an oligomerization domain, a PH domain, and a Dbl/cdc24 guanine nucleotide exchange factor homology domain that contains the N-terminal breakpoint cluster region (BCR) of the protein. The Abl half of the fusion protein contains a tyrosine kinase domain, a Src homology 2 (SH2) domain, and a DNA-binding domain, together with nuclear localization and nuclear export motifs. The Abl kinase is constitutively active within the fusion protein and has been implicated in initiating numerous signaling pathways that mediate cell survival and proliferation. In view of this, and the essential requirement for BCR-Abl kinase activity in CML, it was thought to be an ideal target despite the aforementioned caveats with targeting protein kinases. As predicted, imatinib has revolutionized the treatment of CML. This inhibitor has been remarkably successful in arresting the progression of the CML but is also well tolerated with acceptable side effects.[3]

Although conservation of the kinase ATP-binding pocket has posed a potential problem for designing kinase inhibitors, in practice, this has not happened for several reasons. Although different kinases may be structurally similar in an active ATP-bound conformation, the inactive conformation is substantially less frequent and can be used to generate selective inhibitors.[4] The ATP-binding region is made up of six polar amino acid residues that are invariant across whole families of kinases; similarly, there are numerous lipophilic residues that are highly conserved. In addition, this critical region contains an amino acid whose amide carbonyl binds to N-6 of adenine in the active conformation. The side chain of this amino acid sticks into the reaction pocket in the inactive state and, for this reason, is referred to as "the gatekeeper residue."[5] As the side chain is not involved in direct ATP binding, it varies across kinases, and variation of this gatekeeper residue is exploited by a number of inhibitors that are able to bind the inactive conformation of specific kinases. In the case of Abl kinase, the gatekeeper residue is threonine, which binds directly to a methyl group of the phenyl ring of the Abl kinase inhibitor imatinib.[4] Across the collective kinase superfamily, almost any amino acid can appear as the gatekeeper, although in practice, it is typically a bulky nonpolar residue (methionine, tyrosine, phenylalanine, lysine).[5] In principle, as more detailed structural information emerges from the many protein kinases, capitalizing on subtle differences in structure is expected to lead to the discovery of novel agents with improved potency and specificity. For instance, cyclin-dependent kinase 2 (CDK2) contains an additional pocket on its C-lobe next to the ATP-binding pocket.[6] Several CDK2-specific inhibitors exploit this by binding to both pockets.

Of further structural significance is the emergence of tumor drug resistance in response to the chronic use of protein kinase inhibitors. Mutant forms of BCR-Abl, Kit, and epidermal growth factor receptor (EGFR) have been associated with loss of drug activity and disease relapse. Interestingly, one of the most common sites of mutation is the otherwise conserved "gatekeeper residue." The appearance of such mutations that cause resistance to imatinib have led to the need for other drugs with broader activity against a number of kinases. Thus in the setting of oncology, "multikinase" inhibitors, including dasatinib and sunitinib, have now entered clinical practice and have been approved by the FDA. Although a major problem in the treatment of malignancy, this is less likely to be an issue in the treatment of autoimmune disease; nonetheless, drugs used for oncological indications often end up being quite useful in the treatment of autoimmunity. Such precedents include cyclophosphamide, azathioprine, and methotrexate. Therefore it is not unreasonable to speculate that a number of the kinase inhibitors developed as anticancer agents may ultimately be used to treat inflammatory or immunological diseases.

An unexpected finding for a clinically well-tolerated inhibitor, which was originally thought to be a highly specific inhibitor for Abl kinase, was the realization that imatinib has activity against several other PTKs.[3] Consequently, it has been found to be useful in the treatment of a number of cancers that do not have abnormal Abl kinase activity. Imatinib has been used to treat gastrointestinal (GI) stromal tumor and hypereosinophilic syndrome through its effects on Kit and PDGFR-FIPIL1 kinases, respectively. In spite of efforts to develop highly specific kinase inhibitors, there is increasing evidence that a partial inhibition of multiple kinases is potentially less toxic than originally feared and may be important for the efficacy of many inhibitors. Second-generation Abl inhibitors, including dasatinib and bosutinib, are less selective than imatinib. This lack of specificity may contribute to their improved response rates in the treatment of CML.[7] With respect to immune-mediated diseases, imatinib is used in the clinic for the treatment of fibrotic diseases, including skin fibrosis associated with chronic graft-versus-host disease (GvHD).[8] At the time of writing, there are >20FDA-approved small molecule kinase inhibitors, most which are approved for oncological indications. In addition, there are numerous other kinase inhibitors in clinical trial or development (summarized in Table 88.3). Besides small-molecule direct inhibitors of kinases, which typically block the ATP-binding pocket, there are several alternative strategies. These include small inhibiting RNA to block protein expression of the kinase Syk (Excellair; ZaBeCor Pharmaceutical Co., Bala Cynwyd, PA), inhibition of associated proteins required to activate the target kinase as in the rapamycin derivatives and the use of monoclonal antibodies (mAbs) that inhibit ligand-dependent activation of transmembrane receptor kinases. These mAbs include and target the designated RTK: bevacizumab (vascular endothelial growth factor receptor [VEGFR]), ranibizumab (VEGFR), cetuximab (epidermal growth factor receptor [EGFR]), pertuzumab (human epidermal growth factor receptor [HER]), and trastuzumab (HER2/neu).

TARGETING CYTOKINE SIGNALING BY INHIBITING JANUS KINASES: TOFACITINIB, RUXOLITINIB, AND RELATED COMPOUNDS

Cytokines regulate growth, survival, development, and differentiation of immune cells. Their importance in driving inflammatory and immunological responses has already made them attractive targets as antiinflammatory and immunosuppressive agents. As indicated above, a large subset of cytokines (roughly 60) signal through Janus kinases (JAKs) (see Table 88.2). The essential function of JAKs was documented by knock-out mice and by identifying mutations in patients with immunodeficiencies. Specifically, mutation of *JAK3* results in a severe combined immune deficiency (SCID), characterized by an almost complete absence of T cells and NK cells with defective B cells. This phenocopies deficiency of the cognate receptor that associates with JAK3, the IL-2 receptor common γ chain, cγc (encoded by *IL2RG*), mutation of which underlies X-SCID (see Table 88.1; Fig. 88.4). The profound, but selective, phenotype associated with JAK3 deficiency led to the suggestion that targeting JAKs might be a strategy for the development of a new class of immunomodulatory drugs. There are now several FDA-licensed JAK inhibitors for the treatment of immune and neoplastic diseases, and more are in clinical trials.

Tofacitinib, formerly designated CP-690,550, was one of the first JAK inhibitors to enter the clinic. It inhibits JAK3 and JAK1 and to a lesser extent JAK2, but has little effect on TYK2.[9] Consequently, tofacitinib potently inhibits cγc cytokines but also blocks interferon (IFN)-γ, interleukin (IL)-6, and to a lesser extent

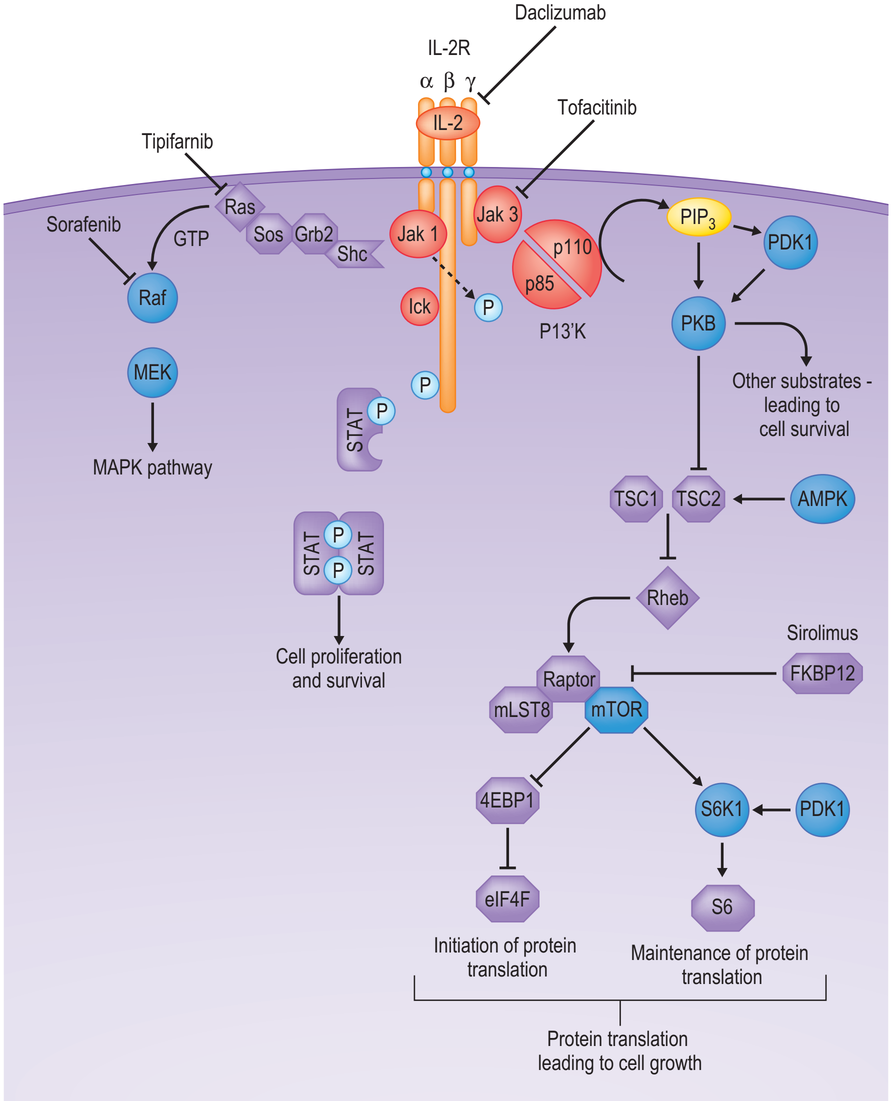

Daclizumab
IL-2R
α β γ
IL-2
Tofacitinib
Tipifarnib
Ras
Sorafenib
GTP
Sos
Grb2
Shc
Jak 1
Jak 3
p110
p85
P13'K
PIP$_3$
PDK1
Raf
lck
P
PKB
MEK
P
STAT
P
Other substrates - leading to cell survival
MAPK pathway
TSC1
TSC2
AMPK
STAT
P
P
STAT
Rheb
Cell proliferation and survival
Sirolimus
Raptor
FKBP12
mLST8
mTOR
4EBP1
S6K1
PDK1
eIF4F
S6
Initiation of protein translation
Maintenance of protein translation
Protein translation leading to cell growth

FIG 88.4 **Signal Transduction Pathways Stemming From the Interleukin (IL)-2 Receptor in T Cells Culminating in the Activation of the Mammalian Target of Rapamycin (mTOR) Serine/Threonine Kinase.** Tyrosine kinases are indicated in red, and serine/ threonine kinases are indicated in blue.

IL-12 and IL-23. Functionally, tofacitinib inhibits T-helper 1 (Th1) cells and Th2 differentiation, as well as pathogenic Th17 cells.[9] In addition to inhibiting adaptive immune responses, tofacitinib appears to inhibit innate immune responses as well.[9] Tofacitinib passed phase III trials for the treatment of methotrexate-resistant rheumatoid arthritis (RA) and was FDA approved for this indication in 2012.[10] It completed phase III trials in both ulcerative colitis and psoriasis with successful results.[11] Other diseases in which tofacitinib is being studied include ankylosing spondylitis and juvenile idiopathic arthritis. Tofacitinib has been used in the treatment of alopecia areata, vitiligo, and atopic dermatitis.[11] A variety of additional inhibitors that target JAK3, including decernotinib, peficitinib, and R-348 (Rigel), are under development or in clinical trials (see Table 88.3).

As gene targeting of *Jak2* in mice was embryonically lethal, it was initially thought that inhibition of JAK2 should be avoided. However, the discovery that gain-of-function mutations of *JAK2* underlie primary polycythemia and myelofibrosis led to the idea that pharmacologically targeting JAK2 could be useful. The JAK1/JAK2 inhibitor ruxolitinib was the first JAK inhibitor for the treatment of myelofibrosis to be approved by the FDA in 2011 and was subsequently approved for the treatment of primary polycythemia in 2014.[11] As cγc cytokines employ both JAK1 and JAK3 for signaling, it might be expected that ruxolitinib and tofacitinib might block some of the same cytokines. It is therefore of interest to note that in a phase II study in RA, ruxolitinib was as efficacious as tofacitinib.[12] A number of novel JAK1/JAK2 inhibitors under development both for the treatment of myeloproliferative and inflammatory diseases. The JAK1/JAK2 inhibitors baricitinib, filgotinib, and AC-410 (Daiichi Sankyo, Tokyo, Japan) all are in trials for the treatment of RA (see Table 88.3).

The side effects of JAK inhibitors include infection, including serious infections, anemia, and leukopenia, presumably related to JAK2 inhibition and interference with cytokines, such erythropoietin, IL-11, and colony-stimulating factors.[11] Little reduction in CD4 T cells has been seen, but significant reduction in NK cells and CD8 T cells does occur; just how significant this will be in terms of infection risk remains to be determined.[11]

TARGETING ANTIGEN RECEPTOR SIGNALING

The first event in TCR signaling is the activation of the Src family kinase Lck (see Fig. 88.1), making it an attractive target as a therapy for autoimmune diseases and transplant rejection. Several Lck inhibitors have been developed and showed promise in preclinical models of allograft rejection and autoimmunity, although at the cost of inducing progressive lymphopenia. This is consistent with the finding that induced deficiency of Lck in mice leads to a progressive lymphopenia.[13] Despite considerable effort, the discovery of a selective Lck inhibitor suitable for use as an immunosuppressive agent remains elusive. Lck activation leads to the recruitment of a second round of tyrosine kinases to the TCR complex; these include Zap70 or Syk. Deficiency of Zap70 causes SCID and preferential loss of CD8 T cells, but curiously, a successful Zap70 inhibitor has yet to be obtained. In contrast, Syk inhibitors have been generated; BIIB057 (Biogen, Cambridge, MA) is being investigated for the treatment of inflammatory diseases, and a second, fostamatinib, is being investigated for the treatment of immune thrombocytopenia and immunoglobulin A (IgA) nephropathy. The recruitment of Zap70/Syk results in activation of another class of PTKs, the Tec family kinases, Rlk and Itk in T cells and Btk in B cells. The importance of this class of kinases is exemplified by the fact that the mutations of *BTK* underlie Bruton agammaglobulinemia a condition characterized by the absence of all B cells (Chapter 34). Ibrutinib is the first BTK inhibitor approved by the FDA in 2014 for the treatment of mantle cell lymphoma and chronic lymphocytic leukemia (CLL).[14] Recent work has identified ibrutinib as an inhibitor of Itk, which suggests it may have value as therapy in autoimmune disease. Antigen receptor activation of phospholipase C (PLC)γ1 appears to require Lck, Zap70, and Tec kinases working in concert. The action of PLCγ1 leads to elevation in intracellular calcium, which in turn activates the phosphatase calcineurin. Calcineurin dephosphorylates and activates nuclear factor of activated T cells (NFAT), which translocates to the nucleus and in cooperation with activator protein-1 (AP-1) transcription factors induces transcription of IL-2 and other key lymphocyte activation genes. A number of potent clinically successful immunosuppressive drugs inhibit calcineurin, including cyclosporine and tacrolimus, drugs that have revolutionized organ transplantation.[15] Despite their success, long-term use of these drugs is limited, as they may cause renal toxicity

CLINICAL PEARLS

Mutations Reveal Key Functions of Kinases in Patients With Primary Immunodeficiencies

- Common γ-chain deficiency, JAK3 deficiency: severe combined immunodeficiency
- ZAP70 deficiency: severe combined immunodeficiency
- TYK2 deficiency: rare cause of hyper-IgE syndrome

Protein Kinase C Family and NF-κB

TCR signaling also leads to the activation of members of the PKC family, which, in turn, activate the transcription factor complex nuclear factor (NF)-κB and the Ras guanyl nucleotide-releasing proteins (RasGRPs). Genetic studies in mice have identified the importance of one member, PKC-θ in TCR signaling. Recent work has identified PKC-θ as an inhibitor of regulatory T cell function.[16] Regulatory T cells (Tregs), unlike conventional effector T cells, act to constrain the immune system. Enhancing Treg numbers or function has been suggested as the treatment of a number of autoimmune diseases.[17] For this reason, preserving or enhancing Treg function through the inhibition of PKC-θ is an attractive strategy. Along with many other receptors, TCR signaling leads to the action of NF-κB family transcription factors, which control the genes involved in cellular activation and resistance to apoptosis.[18] In T cells, PKC-θ is the main isoform responsible for NF-κB activation. The novel PKC inhibitor Sotrastaurin, a potent inhibitor of the PKC-θ, β and α isoforms, is currently undergoing phase II trials for the treatment of diffuse large B-cell lymphoma. Early clinical testing on patients with psoriasis demonstrated that the drug is well tolerated and was able to inhibit T-cell proliferation, IL-2, and IFN-γ secretion. However, despite initial positive trial results, it was not effective in the treatment of psoriasis, ulcerative colitis, or liver transplantation. Downstream of several inflammatory signaling pathways, including PKC, lies the nuclear factor kappa-light-chain-enhancer of activated B cells (NF-κB).[18] NF-κB has been implicated as a major mediator governing both cancer cell proliferation and survival and governing chronic inflammation. It is held in an inactive state by the binding of the inhibitor of κB (IκB). A cascade of protein kinases that include IκB kinases (IKK) and

IKK kinases (IKKKs), of which PKC and PKB are members, regulate IκB.[18] Several IKK inhibitors have been in development; however, all of them have failed clinical trials.

Lipid Kinases and Downstream Signaling

Kinases are also important in phosphorylating lipids, and these modifications are very relevant in signal transduction through both the TCR and cytokine receptors. In addition to the production of inositol triphosphate by the action of PLCγ1, there is a second pathway of inositol lipid metabolism regulated by the TCR that is a response shared by costimulatory molecules, such as CD28, cytokines, and chemokines. This response is mediated by the class I group of phosphatidyl-inositol 3 kinases (PI3Ks), which is composed of four isoforms (PI3Kα to -δ), which phosphorylate the 3'-OH position of the inositol ring of phosphatidyl-inositol (4,5) bisphosphate (PI(4,5)P_2) to produce PI(3,4,5)P_3.[19] This lipid and its metabolite PI(3,4)P_2 bind to the pleckstrin homology (PH) domains of proteins and either induce localization of the protein to defined areas of the plasma membrane, where activation can occur, or induce conformational changes that allow for allosteric modifications of activity. Targets for D-3 phosphoinositides in T cells include a number of downstream protein serine/threonine kinases and the Rac-1 and RhoA guanine nucleotide exchange proteins. Small-molecule inhibitors of PI3'K, Wortmannin and LY294002, are both potent inhibitors of T-cell activation, although toxicity prevents either from being clinically useful. In contrast to the more widely expressed PI3'Kα and β, PI3'Kγ and -δ are only expressed in hematopoietic tissue and deletion of PI3'Kγ results in defective migration of neutrophils and macrophages to sites of inflammation without other pathology. This limited expression makes PI3'Kγ a potentially useful target, and selective PI3'Kγ inhibitors have shown to be effective in mouse models of collagen-induced arthritis.[20] For similar reasons, PI3'Kδ is another attractive target for which the inhibitor idelalisib was FDA-approved for the treatment of lymphoma.[21] Over 25 other inhibitors of PI3'K are currently being studied in preclinical and clinical trials and include buparlisib, which is in a phase III trial for the treatment of breast cancer.

One PI(3,4,5)P_3-regulated kinase activated by the TCR is protein kinase B (PKB/AKT).[22] T cells stimulated in the presence of the PI3'K inhibitors Wortmannin or LY-294002 fail to activate and proliferate. The ability of cells to take up nutrients and switch on glycolysis is essential for lymphocyte activation; PKB is proposed to be the main effector that mediates the action of PI3'K on these fundamentally important metabolic pathways. There are three PKB isoforms (PKB-α, -β, -γ, or AKT1, -2, -3, respectively), and T cells lacking both *Akt1* and *Akt2* are greatly impaired.[22] Specific inhibitors of PKB are in clinical trials with mixed results; MK-2206 (Merck, Kenilworth, NJ), AZD-5363 (Astra Zeneca, Cambridge, UK), and uprosertib are all in phase II trials for the treatment of several malignancies (see Table 88.3).

In addition to glucose metabolism, the PI3'K-regulated serine/threonine kinase mammalian target of rapamycin (mTOR) regulates protein synthesis in response to cellular nutrient and energy levels.[23] It is activated not only by a number of growth factor receptors but also by cytokines, including IL-2. Many signaling pathways link growth factor receptors with activation of mTOR, including the adenosine monophosphate (AMP)–dependent kinase (AMPK) and phosphatidyl inositol 3' kinase (PI3'K) (see Fig. 88.4). mTOR promotes cell growth by activation of p70 S6K1 and inactivation of 4E-BP1,[23] which are critical for translation of new proteins.

As its name suggests, mTOR is inhibited by the macrolide rapamycin, now licensed as the drug sirolimus for the treatment of graft rejection. Sirolimus does not inhibit mTOR by directly binding to the ATP-binding pocket but acts indirectly, associating with FK506-binding protein 12 (FKBP12). This, in turn, inhibits the kinase complex made up of mTOR, mLST8, and raptor (mTORC1).[23] It was hoped that sirolimus would be a potent anticancer drug, but it has met limited success in this regard. In contrast, it has been successfully used as an immunosuppressant, typically as part of a combination regimen for allograft rejection prophylaxis.[24] In view of the ubiquitous expression of mTOR and its role in protein translation, it is not surprising that sirolimus would be associated with a number of side effects, including hyperlipidemia, hypertriglyceridemia, myelosuppression, and delayed wound healing. There is some evidence of renal toxicity, but this is minor compared with that caused by the calcineurin inhibitors cyclosporine and tacrolimus. Myelosuppression associated with sirolimus is typically dose related and rapidly reversible, even in patients receiving the drug for chronic GvHD.

There are currently three rapamycin derivatives, namely, temsirolimus, everolimus, and AP23573, undergoing clinical trials. Everolimus is licensed for use in the management of heart, liver, and kidney transplantation in conjunction with cyclosporine. It is under investigation for the treatment of arthritis and a number of solid tumors. Temsirolimus is licensed for the treatment of renal cell carcinoma and mantle cell lymphoma. The introduction of novel rapamycin derivatives has been joined by the introduction of pan-mTOR direct inhibitors AZD2014 (AstraZeneca, Cambridge, UK) and CC-223 (Celgene, Summit, NJ), both of which are in phase II trials for the treatment of non-Hodgkin lymphoma and solid tumors.

MAPK Pathways

The MAPK family constitute a complex phospho-relay system of signal transduction, composed of three sequentially activated kinases that are themselves modulated by phosphorylation.[25] Substrates of MAPK pathways include transcription factors, phospholipases, cytoskeletal proteins, and other protein kinases. Three main MAPK cascades have been identified in mammalian cells: the extracellular signal regulated kinase (ERK) cascade, the Jun kinase (JNK) cascade, and the p38 MAPK cascade. All start with a membrane-localized activator followed by three MAPKs that sequentially phosphorylate each other (see Fig. 88.3). The top level of kinases is termed MAPK kinase kinases (MAPKKKs, MKKKs, or Map3Ks). The middle level MAPK kinases (MAPKKs, MKKs, or Map2Ks) phosphorylate a common Thr-Xaa-Tyr motif, where Xaa is any amino acid. The lowest tier consists of the MAPKs that phosphorylate Ser/Thr-Pro motifs. Each pathway terminates in the phosphorylation of substrate proteins.

The ERK Cascade

ERK1 and 2 were identified as kinases that were activated in response to growth factor stimulation, which are mimicked by expression of constitutively active Ras. The link between active Ras and subsequent phosphorylation of the ERKs was made both by the discovery of the MAPK kinase MEK1 and its phosphorylation by the known Ras effector RAF1, now known as a MAPK kinase kinase (M3K). The ERK cascade is ubiquitous in mammalian cells and is generally considered one of the main effector pathways regulated by the GTPase p21ras. Activation of the TCR or the IL-2 receptor is able to trigger ERK signaling in T cells.[25] Two small-molecule inhibitors of this serine/threonine kinase pathway are undergoing clinical trials: the farnesyl inhibitor of Ras tipifarnib, which is in phase III trials for the treatment of acute myeloid leukemia; and the multikinase inhibitor sorafenib, which is FDA approved for the treatment of renal cell carcinoma.[26] Sorafenib is relatively well tolerated despite its ability to inhibit numerous kinases, including the M3K RAF, as well as receptor tyrosine kinases, including platelet-derived growth factor receptor (PDGFR), VEGFR, Kit, and FLT-3. Its role as an immunosuppressant has yet to be explored. There are three RAF isoforms: A-RAF, B-RAF, and C-RAF. A number of cancers are associated with B-RAF mutations resulting in a constitutively active kinase, the best known of which is the V600E mutation. Two kinase inhibitors, vemurafenib and dabrafenib, which are highly active against the B-RAF V600E kinase, are FDA licensed for the treatment of melanoma. Their early use was associated with a partial and short-lived inhibition of the MAPK pathway as C-RAF signaling was preserved, which is often hyperactivated by upstream mutations in Ras signaling or by the presence of B-RAF V600E, which is able to dimerize and activate C-RAF even in the presence of inhibitors.[27] Moreover, clinical trials are ongoing for the treatment of the histiocytosis known as *Erdheim-Chester disease,* for which B-RAF mutations have shown to be present in the vast majority of the case.

The M2K MEK1 and the MAPK ERK1 lie downstream of RAF. Cobimetinib and trametinib are selective MEK inhibitors that are FDA licensed for the treatment of melanoma both as a single agent and in combination with B-RAF V600E inhibitors as they are able to block residual C-RAF signaling.[28] Although all of these agents are seen as anticancer drugs rather than immunosuppressants, it is of note that cobimetinib is currently being assessed in the treatment of Langerhans cell histiocytosis (LCH). LCH is a clonal disorder characterized both by the B-RAF V600E mutation in histiocytes and by the presence of inflammatory lesions that often respond to simple corticosteroid therapy. Selective inhibitors of ERK have been developed: FR180204 has been shown to inhibit the development of collagen-induced arthritis in mice and is being considered as an agent in the treatment of RA. CC90003 (Celgene, Summit, NJ) is currently being assessed in a phase I trial in patients with solid tumors.

The JNK Cascade

Another limb of the MAPK pathway is the JNK pathway. Many inflammatory agents, including lipopolysaccharide (LPS), TNF-α, and IL-1 are able to activate the JNK pathway.[25] In synoviocytes, this results in secretion of proteases implicated in joint destruction seen in RA.[29] A number of small-molecule inhibitors of JNKs have been identified and are currently being investigated in the treatment of inflammation, cancer, and neurological diseases. XG102 (Xigen Ltd., Bedford, UK) is being assessed in phase III trials for the treatment of ocular inflammation and sensorineural hearing loss.

The p38 MAPK Cascade

This cascade was originally identified as part of a drug screen looking for inhibitors of TNF-α–mediated inflammatory responses.[30] TLR-dependent production of IL-1 and TNF-α is p38 MAPK dependent. The success of TNF-α–blocking antibodies in the treatment of RA has led to much interest in the development of p38 MAPK inhibitors. However, although many p38 inhibitors have been reported, their development into therapeutic drugs has been frustrated by either unacceptable toxicity or poor efficacy. Currently, a new generation of p38 MAPK inhibitors are being assessed in phase II trials for the treatment of chronic obstructive pulmonary disease (COPD); losmapimod has recently failed phase II testing because of lack of efficacy, and AZD7624 (AstraZeneca, Cambridge, UK) and acumapimod are still being investigated.

CONCLUSIONS

ON THE HORIZON

- Many kinase inhibitors, initially designed as antitumor agents, have been repurposed as immunosuppressants (*e.g.,* ruxolitinib).
- Further understanding of the pathophysiology of autoimmune disease will lead to the increased use and creation of kinase inhibitors.
- Increasing use of specific kinase inhibitors may improve our understanding of the pathophysiology of autoimmune disease.

Scientific advances in the 1990s have led to the discovery of many novel intracellular signaling pathways that link membrane-bound receptor and cytokine signaling with alteration of gene expression and cellular activation necessary to trigger an immune cell response. Many of these pathways are interlinked to make up a complex array of networks composed of enzymes, adaptor proteins, and transcription factors, all of which are potential targets for drug discovery in the quest to make a therapy that will be able to treat a specific autoimmune disease without an unacceptable degree of immunosuppression. Now more than a decade since the identification of many new targets, the first generation of drugs designed to interfere with specific immune cell signals are being brought to the clinic. The success of the anticancer BCR-Abl inhibitor imatinib and the immunosuppressive mTOR inhibitor sirolimus has placed the protein kinases center stage as targets of future drug discovery. As many of the key steps in the activation of an immune cell are often shared with those that allow a cancer cell to proliferate, many of these agents are being tested as anticancer drugs rather than as immunosuppressants. Despite this, some agents, such as the mTOR inhibitors, originally intended for the treatment of cancer have been far more successful in the field of immunology, and this may continue to be true for future modifiers of cell signaling. Conversely, JAK inhibitors could potentially be used in the treatment of lymphoma. Either way, we are likely to see a large number of novel immunosuppressants appear both serendipitously and intentionally as new protein kinase inhibitors are licensed for a wide range of debilitating illnesses.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Hanks SK, Quinn AM, Hunter T. The protein kinase family: conserved features and deduced phylogeny of the catalytic domains. Science 1988;241:42–52.
2. Noble ME, Endicott JA, Johnson LN. Protein kinase inhibitors: insights into drug design from structure. Science 2004;303:1800–5.
3. Druker BJ, Lydon NB. Lessons learned from the development of an Abl tyrosine kinase inhibitor for chronic myelogenous leukemia. J Clin Invest 2000;105:3–7.

4. Schindler T, Bornmann W, Pellicena P, et al. Structural mechanism for STI-571 inhibition of Abelson tyrosine kinase. Science 2000;289:1938–42.
5. Adams J, Huang P, Patrick D. A strategy for the design of multiplex inhibitors for kinase-mediated signalling in angiogenesis. Curr Opin Chem Biol 2002;6:486–92.
6. Davies TG, Pratt DJ, Endicott JA, et al. Structure-based design of cyclin-dependent kinase inhibitors. Pharmacol Ther 2002;93:125–33.
7. Kantarjian H, et al. Dasatinib versus imatinib in newly diagnosed chronic-phase chronic myeloid leukemia. N Engl J Med 2010;362:2260–70.
8. Arai S, et al. A Randomized Phase II Crossover Study of Imatinib or Rituximab for Cutaneous Sclerosis after Hematopoietic Cell Transplantation. Clin Cancer Res 2016;22:319–27.
9. Ghoreschi K, et al. Modulation of Innate and Adaptive Immune Responses by Tofacitinib (CP-690,550). J Immunol 2011;186:4234–43.
10. Fleischmann R, et al. Placebo-controlled trial of tofacitinib monotherapy in rheumatoid arthritis. N Engl J Med 2012;367:495–507.
11. Schwartz DM, Bonelli M, Gadina M, et al. Type I/II cytokines, JAKs, and new strategies for treating autoimmune diseases. Nat Rev Rheumatol 2016;12:25–36.
12. Mesa RA. Ruxolitinib, a selective JAK1 and JAK2 inhibitor for the treatment of myeloproliferative neoplasms and psoriasis. IDrugs 2010;13:394–403.
13. Zamoyska R, Basson A, Filby A, et al. The influence of the Src-family kinases, Lck and Fyn, on T cell differentiation, survival and activation. Immunol Rev 2003;191:107–18.
14. Wang ML, Rule S, Martin P, et al. Targeting BTK with ibrutinib in relapsed or refractory mantle-cell lymphoma. N Engl J Med 2013;369:507–16.
15. Calne RY. Organ transplantation has come of age. Sci Prog 2010;93:141–50.
16. Zanin-Zhorov A, Ding Y, Kumari S, et al. Protein kinase C-theta mediates negative feedback on regulatory T cell function. Science 2010;328:372–6.
17. Riley JL, June CH, Blazar BR. Human T regulatory cell therapy: take a billion or so and call me in the morning. Immunity 2009;30:656–65.
18. DiDonato JA, Mercurio F, Karin M. NF-κB and the link between inflammation and cancer. Immunol Rev 2012;246:379–400.
19. Cantrell DA. Phosphoinositide 3-kinase signalling pathways. J Cell Sci 2001;114:1439–45.
20. Camps M, et al. Blockade of PI3Kgamma suppresses joint inflammation and damage in mouse models of rheumatoid arthritis. Nat Med 2005;11:936–43.
21. Gopal AK, Kahl BS, de Vos S, et al. PI3Kdelta inhibition by idelalisib in patients with relapsed indolent lymphoma. N Engl J Med 2014;370:1008–18.
22. Juntilla MM, Wofford JA, Birnbaum MJ, et al. Akt1 and Akt2 are required for αβ thymocyte survival and differentiation. Proc Natl Acad Sci USA 2007;104:12105–10.
23. McManus EJ, Alessi DR. TSC1-TSC2: a complex tale of PKB-mediated S6K regulation. Nat Cell Biol 2002;4:E214–16.
24. Kreis H, Cisterne JM, Land W, et al. Sirolimus in association with mycophenolate mofetil induction for the prevention of acute graft rejection in renal allograft recipients. Transplantation 2000;69:1252–60.
25. Dong C, Davis RJ, Flavell RA. MAP kinases in the immune response. Annu Rev Immunol 2002;20:55–72.
26. Josephs DH, Ross PJ. Sorafenib in hepatocellular carcinoma. Br J Hosp Med (Lond) 2010;71:451–6.
27. Park EJ, Lee JH, Yu GY, et al. Dietary and genetic obesity promote liver inflammation and tumorigenesis by enhancing IL-6 and TNF expression. Cell 2010;140:197–208.
28. Larkin J, Ascierto PA, Dréno B, et al. Combined vemurafenib and cobimetinib in BRAF-mutated melanoma. N Engl J Med 2014;371:1867–76.
29. Gum R, Wang H, Lengyel E, et al. Regulation of 92 kDa type IV collagenase expression by the Jun aminoterminal kinase—and the extracellular signal-regulated kinase-dependent signaling cascades. Oncogene 1997;14:1481–93.
30. Lee JC, Laydon JT, McDonnell PC, et al. A protein kinase involved in the regulation of inflammatory cytokine biosynthesis. Nature 1994;372:739–46.

MULTIPLE-CHOICE QUESTIONS

1. Specific inhibitors of individual protein kinases are possible because:
   - **A.** Most kinase inhibitors are noncompetitive covalent inhibitors.
   - **B.** Protein kinases can be divided into families of serine, threonine, and tyrosine kinases that are structurally very dissimilar.
   - **C.** The adenosine triphosphate (ATP)–binding pocket of kinases are sufficiently distinct to allow generation of selective inhibitor.
   - **D.** Most kinase inhibitors avoid the structurally similar ATP-binding pocket.
   - **E.** Distinct kinases occupy different parts of the cell and do not interact with each other.

2. The first clinically successful direct tyrosine kinase inhibitor was:
   - **A.** Ruxolitinib
   - **B.** Infliximab
   - **C.** Imatinib
   - **D.** ZAP70
   - **E.** Methotrexate

3. Patients deficient in the following protein kinases suffer from a severe combined immunodeficiency (SCID):
   - **A.** BCR-Abl
   - **B.** JAK3
   - **C.** PTEN
   - **D.** TYK2
   - **E.** BTK

Biological Modifiers of Inflammatory Diseases

W. Winn Chatham

During recent decades, the elucidation of specific inflammatory mediators and identification of lymphocyte lineages that contribute to the pathology underlying chronic inflammatory disorders have identified immunological targets amenable to modification by use of recombinant DNA technologies. The resulting biological modifiers of inflammatory diseases have revolutionized the treatment of rheumatoid arthritis (RA), inflammatory bowel disease (IBD), psoriasis, spondyloarthropathies, antineutrophil cytoplasmic antibody (ANCA)–associated vasculitis syndromes, systemic lupus erythematosus (SLE), and other autoimmune diseases, as well as allergic disorders. These targeted therapies, although costly, offer individuals affected by these disorders options for better disease control and less of the morbidity associated with greater exposure to broader immune suppression and/or corticosteroid use. This chapter summarizes the biological modifiers of inflammatory diseases that are currently available or in development classed by mechanism of activity and molecular nature.

IMMUNOMODULATORY CYTOKINES

Recombinant Interleukin-2 (Aldesleukin)

First identified as a T-cell growth factor, recombinant interleukin-2 (rIL-2) has potent immunomodulatory and antitumor activity, promoting the proliferation, differentiation, and recruitment of T cells, B cells, and natural killer (NK) cells (Table 89.1). rIL-2 has primarily been used therapeutically in patients with advanced melanoma and in patients with advanced renal cell carcinoma. IL-2 is a potent inducer of proinflammatory T-cell cytokines, such as IL-1, tumor necrosis factor-α (TNF-α), and interferon-γ (IFN-γ), all of which likely play a major role in dose-related IL-2 toxicity, which may include hypotension, cardiac arrhythmias, increased capillary permeability with pulmonary edema, fever, and rarely death. The use of aldesleukin in patients with cancer has been increasingly supplanted by more targeted immune-enhancing "checkpoint" therapies targeting T-cell death receptors (programmed death 1 [PD-1]) and cytotoxic T lymphocyte antigen-4 (CTLA-4), which have comparable antitumor effects with much less acute toxicity.

More recently, the administration of low-dose rIL-2 has been shown to promote preferential growth and proliferation of regulatory T cells (Tregs) without significantly affecting the release of inflammatory mediators by CD4 T cells.[1] Such low-dose regimens of rIL-2 have been used with success in suppressing manifestations of graft-versus-host disease (GvHD) and may also favorably impact the course of type 1 diabetes.[1] The use of rIL-2 and/or other cytokine therapies promoting enhanced Treg differentiation and proliferation may therefore provide novel interventions for managing a variety of inflammatory disorders associated with autoimmunity, including SLE.[2]

KEY CONCEPTS

Interferons: Therapeutic Potential

- Interferon-α (IFN-α) has been used with success to enhance resolution of chronic viruses, notably hepatitis C virus (HCV); newer antivirals are supplanting the need for IFN therapy in HCV.
- IFN-β has immunomodulatory effects that are of benefit in the treatment of certain subsets of patients with multiple sclerosis (MS).
- Type 1 IFNs (including IFN-α and IFN-β) may unmask or trigger flares of autoimmune disease, most notably systemic lupus erythematosus (SLE).
- IFN-γ may be of some benefit in the management of patients with chronic granulomatous disease (CGD) not responding to antifungal and antibacterial prophylactic regimens.

Interferons

IFNs, released primarily by T lymphocytes as well as dendritic cells (DCs) in the context of stimulation of Toll-like receptors (TLRs) have a wide array of immunomodulatory effects that include upregulation of genes governing angiogenesis, cell differentiation, expression of human leukocyte antigen (HLA) molecules, and production of inflammatory cytokines. IFN-α and IFN-β bind to the same cell surface receptor (IFN-1R) and are designated type-1 IFNs, whereas IFN-γ binds to a different receptor (IFN-2R) and is designated as a type 2 IFN. Recombinant preparations of all three IFNs have been used for the management of inflammatory disorders associated with chronic viral infection, primary immunodeficiency (PID), or select autoimmune disorders (most notably multiple sclerosis [MS]). Although there are immunomodulatory attributes that render IFNs useful in certain circumstances, constitutional symptoms and upregulation of a wide variety of genes that promote inflammation frequently limit the use of both type 1 and type 2 IFNs.

Interferon-α

Recombinant IFN-α2b (rIFN-α2b) has most commonly been used in combination with ribavirin for treatment of hepatitis C, including hepatitis C virus (HCV)–associated cryoglobulin syndromes.[3] Other disorders linked to viral infections, including lymphomatoid granulomatosis (caused by Epstein-Barr virus [EBV]) and polyarteritis nodosa (caused by hepatitis B virus [HBV]), have been successfully treated with regimens incorporating IFN-α2b as part of the treatment regimen.[4] Refractory retinal

TABLE 89.1 Recombinant Immunomodulatory Cytokines

Molecule	Construct	Half-Life	Dosing
aldesleukin	rIL-2	1.5 hours	0.3–3 × 10^6 IU/m^2 subcutaneously (SC) daily
IFN-α2b	rIFN-α2b	1.7 hours	3–10 × 10^6 IU 3x/week
IFN-α2b pegylated	rIFN-α2b pegylated	17 hours	50–150 μg/week
IFN-β1a	rIFN-β1a	69 hours	8.8–44 μg SC 3x/week
INF-β1b	rIFN-β1b	0.2–4.3 hours	2–8 × 10^6 IU every other day
IFN-γ	rIFN-γ	6 hours	50 μg (1 × 10^6 IU)/m2 SC 3x/week

vasculitis in the context of Behçet disease, severe flare-ups of familial Mediterranean fever not responding to treatment with colchicine, and eosinophilic granulomatosis with polyangiitis (EGPA) have also been reported to respond favorably to treatment with rIFN-α2b.[5,6]

Common side effects of rIFN-α2b include flu-like syndrome, fatigue, anorexia, nausea, weight loss, hair loss, emotional lability and depression, cytopenias, and induction of autoantibodies with enhancement of autoimmune disease. Because of these consequences of treatment, the inconstant efficacy against certain HCV genotypes, and the availability of more highly effective antiviral agents targeting HCV replication, use of rIFN-α2b in the treatment of HCV has decreased.

Interferon-β

Recombinant IFN-β1a or IFN-β1b (rIFN-β) has been shown to decrease relapse rates, disease severity, and central nervous system (CNS) magnetic resonance imaging (MRI) lesions in patients with relapsing MS. However, there are conflicting data with regard to the efficacy of IFN-β1a or IFN-β1b preparations for patients with secondary progressive variants of MS.[7] Immunomodulatory effects attributed to the beneficial impact of IFN-β1b on MS include enhancement of suppressor T-cell activity, reduction of proinflammatory cytokines, downregulation of antigen presentation, and reduced trafficking of lymphocytes into the CNS.[7] Side effects observed during treatment with both rIFN-β preparations include injection site reactions as well as most of the effects observed during treatment with rIFN-α2b. Reported autoimmune complications include SLE-related complications as well as features often associated with SLE, including immune complex glomerulonephritis, cutaneous vasculitis, and panniculitis.[8]

Interferon-γ

Recombinant IFN-γ (rIFN-γ) is effective in decreasing the frequency and severity of infections in patients with chronic granulomatous disease, an inherited disorder associated with any number of genetic defects that impair assembly of the nicotinamide adenine dinucleotide phosphate (NADPH) oxidase and production of superoxide anion metabolites in neutrophils.[9] In permissive X-linked variants of chronic granulomatous disease (CGD), treatment with rIFN-γ appears to increase superoxide production. However, infection outcomes in murine models of CGD void of any capability of superoxide generation are still improved, so it is unclear whether the noted favorable impact on infection frequency and severity observed with rIFN-γ use is attributable to its effects on superoxide production.[10]

Fever, myalgias, rash, fatigue, and diarrhea are the most common adverse events associated with IFN-γ treatment. During intercurrent infections, treatment with rIFN-γ often enhances infection-related constitutional symptoms, potentially obscuring responses to antimicrobial therapy. With greater adherence to antimicrobial prophylactic regimens incorporating combination therapy with itraconazole and trimethoprim-sulfamethoxazole (TMP-SMX), the added benefit of treatment with rIFN-γ in patients with CGD may be marginal.

KEY CONCEPTS

Tumor Necrosis Factor-α (GR) Inhibitors

- Tumor necrosis factor (TNF)-α inhibitor treatment significantly improves patient symptoms and function and inhibits structural damage in patients with rheumatoid arthritis (RA).
- TNF-α inhibitors are very effective in suppressing skin lesions caused by psoriasis as well as entheseal inflammation in patients with seronegative spondyloarthropathies; they do not appear to have any inhibitory effects on the development of bony ankyloses.
- Vigilance for underlying fungal disease or mycobacterial disease is prudent prior to commencing with and during treatment courses with TNF-α inhibitors.
- TNF-α inhibitors may trigger or exacerbate certain autoimmune syndromes and should be used with caution or avoided in patients with autoantibody-associated immune disorders, such as systemic lupus erythematosus (SLE).

INHIBITORS OF INFLAMMATORY CYTOKINES

Tumor Necrosis Factor-α Inhibitors

TNF-α is a significant mediator of inflammation associated with psoriasis, RA, spondyloarthropathies, and chronic inflammatory bowel diseases (IBDs; Crohn disease and ulcerative colitis [UC]). TNF-α promotes the ingress of immunocompetent cells at sites of inflammation through activation of the vascular endothelium and upregulation of adhesion molecules. TNF-α also stimulates synthesis of other pro-inflammatory cytokines (IL-1β, IL-6, granulocyte macrophage–colony-stimulating factor [GM-CSF]), chemokines (IL-8), and inflammatory eicosanoids (prostaglandin E$_1$ [PGE$_1$], leukotriene B$_4$ [LTB$_4$]).[11] In RA, TNF-α stimulates macrophages and monocyte-derived osteoclasts to release mediators destructive to bone and cartilage, including matrix metalloproteinases (MMPs), such as collagenase and stromelysin. These noted effects of TNF-α have, therefore, prompted the development of both soluble receptor and monoclonal reagent strategies to inhibit the effects of TNF-α in disorders, in which it appears to have a significant role in promoting inflammation and tissue injury (Table 89.2).

Etanercept is a recombinant soluble p75 TNF receptor (TNFR; CD120b)-IgG Fc fusion protein that binds soluble TNF-α as well as lymphotoxin-β. The avidity of the p75 TNFR for soluble TNF-α is comparable with that of the p55 TNFR and therefore effectively decreases net TNF signaling through either of these receptors. The clinical impact of binding lymphotoxin-β is not entirely known, but recent work elucidating the role of lymphotoxin-β in maintaining the physiological regulatory function of macrophages around lymphoid germinal centers suggests that inhibition of this cytokine may enhance the delivery of apoptotic body–derived self-antigens to germinal centers, a

TABLE 89.2 Recombinant Inhibitors of Pro-Inflammatory Cytokines

Molecule	Construct	Half-Life	Dosing (Maintenance)
Etanercept	sTNFR:Fc	3–4 days	50 mg subcutaneously (SC) once a week or 25 mg SC 2x/week
Infliximab	aTNF-α (chimeric IgG1κ)	7–12 days	3–10 mg/kg intravenously (IV) q4–8 weeks
Adalimumab	aTNF-α (human IgG1κ)	10–20 days	10–40 mg SC q1–2 weeks
Certolizumab	aTNF-α (humanized Fab')-PEG	14 days	200-400 mg SC q2–4 weeks
Golimumab	aTNF-α (human IgG1κ)	14 days	50-100 mg SC every 4 weeks 2 mg/kg IV every 8 weeks
Anakinra	sIL-1Ra	4–6 hours	1–8 mg/kg SC daily (max 200 mg/day)
Rilonacept	IL-1R1:Fc(IgG1):IL-1RAcP	8–9 days	2.2 mg/kg (max 160 mg) SC per week
Canakinumab	aIL-1β (human IgG1κ)	26 days	2–8 mg/kg SC (max 600 mg) SC every 8 weeks
Tocilizumab	aIL-6R (humanized IgG1κ)	13 days	4-12 mg/kg (max 800 mg) IV every 2–4 weeks 160 mg SC every 1–2 weeks
Ustekinumab	aIL-12/23 p40 (human IgG1κ)	15–32 days	0.75 mg/kg (pediatric) or 45–90 mg SC every 12 weeks
Secukinumab	aIL-17A (human IgG1κ)	22–31 days	150-300 mg SC every 4 weeks
Ixekizumab	aIL-17A (humanized IgG4κ)	13 days	80 mg SC every 2–4 weeks

consideration of potential clinical relevance in patients at risk for SLE or related autoimmune syndromes (see below).

Infliximab, adalimumab, and *golimumab* are monoclonal antibody (mAb) reagents with comparable avidity for both soluble and cell-bound TNF-α but differ with regard to the derivation of the antigen-binding sequence and half-life. Infliximab is a chimeric antibody with a murine-derived variable-region TNF-binding domain. Adalimumab and golimumab are both human sequence–derived mAb reagents. *Certolizumab pegol* is a construct consisting of recombinant human sequence–derived F(ab)' anti–TNF-α covalently liked to 40-kilodalton (kDa) polyethylene glycol. The pegylated construct enhances the half-life of the reagent, and absence of the Fc- and complement-binding domains renders certolizumab less likely to engender local injection site reactions and precludes placental crossing of certolizumab into the fetal circulation when administered to pregnant women.

In cytokine profile studies of joint tissues and synovial fluids derived from patients with RA, TNF-α is consistently among the highest expressed inflammatory cytokines.[11] All five TNF-α inhibitors have been shown to decrease the signs and symptoms of disease activity and to inhibit the progression of structural damage in RA, with greatest benefits and probability of disease remission seen among patients with early RA. Inhibition of radiographic progression of disease, as measured by joint space narrowing and osseous erosions at joint margins, can be dissociated from clinical efficacy as measured by composite scoring measures, such as the American College of Rheumatology 20%/50%/70% improvement criteria or the Disease Activity Score (DAS)-28.[12]

TNF-α levels are also notably increased in biopsy samples of skin and synovial tissues from patients with psoriasis and psoriatic arthritis, and the TNF-α inhibitors have been shown to be highly effective in suppressing manifestations of plaque psoriasis as well as articular disease in psoriatic arthritis.[13,14] Improvements in joint symptoms (axial and peripheral) and uveitis manifestations are observed in patients with other spondyloarthropathies, including ankylosing spondylitis, during treatment with TNF-α inhibitors; however, use of these agents has not been shown to delay fusion of axial joints in patients with spondylitis. Insights from rodent models and the known inhibitory effect of TNF-α on Smad pathway–mediated bone formation by osteoblasts suggest that TNF inhibition may, in fact, promote ossification at sights of inflammation and increased bone turnover.[15]

mAb inhibitors of TNF-α as well as certolizumab (but not etanercept) have been shown beneficial in the management of intestinal inflammation in patients with Crohn disease or UC and are approved for treatment of these disorders.[16] Pyoderma syndromes and the oral and genital lesions of Behçet disease are reported to respond favorably to TNF-α inhibition therapy.[17,18]

Although very well tolerated as a class, TNF-α inhibitors may impair innate host defenses, resulting in delay of resolution of intercurrent infections. TNF-α inhibitors are associated with reactivation of tuberculosis and fungal infections, including histoplasmosis and coccidioidomycosis, and viral infections such as hepatitis B.[19,20] In all circumstances of intercurrent infection, the standard recommendation is for treatment with TNF-α inhibitors to be withheld until the infection has resolved. No complications have been reported regarding their use in patients with intercurrent hepatitis C infection or in patients with human immunodeficiency virus (HIV) infection that is well controlled with highly active antiretroviral therapy (HAART).

In the context of its role in promoting cell apoptosis, TNF-α constitutes a component of the host defense against tumor cell survival and growth. As such, the incidence and prevalence of cancer in populations treated with TNF-α inhibitors has been emphasized in pharmacovigilance programs surrounding the use of this class of therapeutic agents. Lymphoma has been reported in patients treated with TNF-α inhibitors, with most reported cases occurring in the context of treatment for RA. However, the relative risk for lymphoma in patients with RA in the pre–biologicals era approximates 3.0, and the prevalence of lymphoma in patients with RA exposed to TNF-α inhibitors has not been shown to exceed the otherwise expected prevalence among patients with RA. Among patients with RA, IBD, or psoriasis, the risk of cancer has not been shown to be increased in the context of treatment with TNF-α inhibitors.[21] Although it is often recommended that TNF-α inhibitor therapy be discontinued in the setting of a newly diagnosed cancer and that initiation of such therapy be avoided in patients with known malignancy, there are no clinical data to confirm whether continued treatment with TNF-α inhibitors impairs therapeutic response to cancer therapy.

The role of TNF-α inhibition in facilitating apoptosis is also of relevance with regard to the development or potentiation of autoimmunity. Although rare, the development of psoriasiform skin eruptions, demyelinating syndromes, and drug-induced SLE are well documented occurrences in patients treated with TNF-α inhibitors.[22-24]All of the anti–TNF-α reagents may attenuate TNF-α–mediated apoptosis of autoreactive T cells. Infliximab, adalimumab, and golimumab may induce the release of apoptotic products in the context of antibody-dependent cellular cytotoxicity (ADCC) of cells bearing surface TNF-α, potentially driving production of antibodies to nucleoproteins.[25] Binding of lymphotoxin-β by etanercept may alter the clearance of apoptotic products by germinal center adjacent follicular DCs and macrophages.[26] Given these collective considerations, it is generally recommended that TNF-α inhibitors be discontinued in patients who develop autoimmune complications during treatment with this class of biological therapy. Moreover, anti–TNF-α therapy is not recommended for patients with established SLE or other overlap syndromes with lupus-like features associated with autoantibody production.

Interleukin-1β Inhibitors

IL-1β is synthesized as an inactive precursor, with activation of IL-1β occurring following engagement of nucleotide-binding oligomerization domain (NOD)–like receptors (NLRs) by a variety of exogenous or endogenous danger signals, which then trigger formation of molecular platforms (NLR family domain containing protein 3 [NALP] inflammasomes) that facilitate cleavage of the IL-1β precursor by IL-1–converting enzyme (ICE).[27] IL-1β stimulates proliferation of lymphocytes, upregulates the expression of adhesion molecules, and triggers the release of a variety of inflammatory mediators from leukocytes, including chemotactic factors, prostaglandins, proteases, and procoagulants. In patients with RA, IL-1β triggers the release of proteases from phagocytic cells and macrophages that are destructive to bone and cartilage. Inflammasome activation has also been shown to mediate acute flare-ups of arthritis in patients with gout or calcium pyrophosphate deposition. Dysregulation of NLRP3 as a result of gain-of-function mutations in inflammasome-related genes has been implicated in the pathogenesis of familial cryopyrin-associated periodic syndromes (CAPS, cryopyrinopathies), familial Mediterranean fever (FMF), and the pyogenic arthritis, pyoderma gangrenosum, and acne syndrome (PAPA). Increased expression and levels of IL-1β are also noted during flare-ups of inflammation in patients with adult-onset Still disease (AoSD), systemic-onset juvenile idiopathic arthritis (soJIA), and macrophage activation syndromes (MAS) associated with these and other autoimmune or infectious disorders.[27] The naturally occurring IL-1 receptor antagonist (IL-1Ra) prevents the binding of IL-1β and IL-1α to the IL-1 receptor (IL-1R), but tissue levels of IL-1Ra in in the above disorders may be insufficient to counteract the effects of IL-1β.

KEY CONCEPTS

Interleukin Inhibitors

- Interleukin (IL)-1β inhibitors are very effective in the treatment of disorders associated with constitutional or secondary activation of the IL-1 inflammasome, including cryopyrin syndromes, familial Mediterranean fever flare-ups, and flare-ups of crystal induced arthropathy (gout and calcium pyrophosphate dihydrate [CPPD] crystal deposition disease/pseudogout).
- Use of IL-1β inhibitors with a short half-life may be of benefit in managing severe flare-ups of inflammation associated with macrophage activation, such as systemic-onset juvenile idiopathic arthritis (soJIA), adult-onset Still disease, and severe sepsis syndromes.
- IL-6 signaling is complex in the context of signaling that can occur through cell membrane bound IL-6R, which mediates primarily IL-6 homeostatic effects, or through many other cell receptors that engage IL-6/sIL-6R complexes (trans-signaling), which mediate IL-6 inflammatory effects.
- Inhibition of IL-6 binding to IL-6R mediated by monoclonal antibody (mAb) reagents, such as tocilizumab, is very effective in suppressing clinical manifestations of rheumatoid arthritis (RA).

Anakinra is a nonglycosylated recombinant IL-1 receptor antagonist (rIL-1Ra) that differs from the endogenous IL-1Ra by a single amino acid addition at the amino terminus. Administered subcutaneously at daily (or for some indications more frequent) intervals because of its very short serum half-life, anakinra functions as a competitive inhibitor of IL-1α and IL-1β binding to IL-1 receptors.

Anakinra, presently approved for the treatment of RA, inhibits joint erosions. However, compared with the clinical responses observed with use of TNF inhibitors and IL-6 inhibitors (tocilizumab) with regard to tender and swollen joints, such responses with anakinra are much more modest. Given these observations with biologicals that require much less frequent dosing, anakinra is used infrequently in the treatment of RA. With the increased understanding of the role of IL-1 inflammasome activation in crystal-induced arthropathies, such as gout and systemic inflammatory disorders, such as soJIA and AoSD, anakinra has been gaining renewed interest as a therapeutic option in the treatment of these disorders. Case reports support use of anakinra in managing acute flare-ups of gout or flare-ups of calcium pyrophosphate–associated pseudogout when the use of corticosteroids or nonsteroidal antiinflammatory drugs (NSAIDs) is not favored because of comorbid conditions, such as uncompensated heart failure or diabetes with significant renal impairment.[28] Prompt resolution of inflammatory markers and clinical manifestations have been reported in patients with severe flare-ups of AoSD and soJIA treated with anakinra, including those associated with MAS.[29,30] The relatively short half-life of anakinra may render its use feasible when short-term blockade of IL-1 may be of benefit in managing severe flare-ups of IL-1–driven inflammation when there is a concern for intercurrent infection. The potential utility of anakinra in this context is supported by a recent study reexamining the outcomes from a previous sepsis trial, where survival was shown to have improved in patients with clinical features of MAS who were randomized to anakinra treatment.[31]

Rilonacept (also referred to as *IL-1-Trap*) is a recombinant fusion protein that comprises the extracellular domain of the IL-1 accessory protein and IL-1 receptor type 1 attached to the Fc portion of immunoglobulin G1 (IgG1). Rilonacept binds to IL-1α and IL-1β with high affinity and is approved for the treatment of CAPS. Similar to anakinra, rilonacept is dosed subcutaneously, albeit at weekly intervals because of its longer half-life. It is generally well tolerated, with injection site reactions being the most common adverse events.

Canakinumab is a human genome sequence–derived mAb with specificity for IL-1β that also binds to IL-1α. Approved for the treatment of CAPS as well as soJIA, canakinumab has the longest half-life of the currently approved therapies that target IL-1. In a randomized clinical trial comparing the use of canakinumab with parenteral triamcinolone administration for acute gout flare-ups, canakinumab was shown to have greater reductions in joint pain and swelling, decreased needs for rescue medications, and a greater time to subsequent flare-ups.[32] However, given that the observed rates of serious infections were twice as high in the canakinumab arm, a half-life that well exceeds the duration of typical flare-ups, and the questionable cost-effectiveness of using an intervention 5000 times the cost of current therapies that are, by and large, effective for managing acute gout, the use of canakinumab has not found favor in treatment paradigms for managing gout.

Interleukin-6 Inhibitors

IL-6 mediates the activation of macrophages and osteoclasts, terminal proliferation and differentiation of B cells, differentiation of T-helper 17 (Th17) cells, and production of liver acute-phase proteins.[33] However, IL-6 also governs homeostatic processes, including granulopoiesis, enteric epithelial proliferation, and antiinflammatory responses involved in resolution of inflammation, including production of the soluble TNFR p55 and the IL-1Ra.[34]

The biology of IL-6 signaling is complex in that signaling may occur directly through the cell membrane–bound IL-6 receptor (IL-6R)/gp130 protein complex (classic signaling) or through binding of IL-6 to soluble IL-6R (sIL-6R) with the resulting heterodimer and then ligating a variety of other gp130-containing membrane receptors (trans-signaling) that mediate signaling from cytokines other than IL-6. IL-6R is expressed primarily on leukocytes, hepatocytes, and megakaryocytes, whereas gp130-containing receptor complexes are expressed in almost all organs, including the heart, kidneys, spleen, liver, lungs, placenta, and brain. In murine models of inflammation employing sgp130:Fc constructs that selectively bind and neutralize IL-6/sIL-6R, trans-signaling appears to mediate many of the observed inflammatory consequences of the upregulation of IL-6, whereas classic signaling through membrane IL-6R mediates homeostatic processes, such as granulopoiesis, thrombopoiesis, and epithelial cell proliferation.[35] With greater understanding of the different IL-6 signaling pathways in humans, implementation of therapeutic strategies targeting IL-6 will require consideration of the impact of blocking classic membrane-bound IL-6R signaling and IL-6/sIL-6R trans-signaling.

Tocilizumab is a humanized mAb with specificity for the human IL-6 receptor. Tocilizumab blocks the binding of IL-6 to IL-6R, inhibiting both classic signaling by IL-6 through membrane IL-6R and formation of IL-6/sIL-6R heterodimer ligands that engender trans-signaling.[36] Treatment with tocilizumab improves tender and swollen joint counts and slows the development of joint erosions in patients with RA.[37] Treatment with tocilizumab is also reported to be of benefit in patients with soJIA, AoSD, and Castleman disease.[38,39]

Presumably related to the role of IL-6 in promoting granulopoiesis, neutropenia may be observed in some patients following treatment with tocilizumab, but this is uncommon. Thrombocytopenia and elevated serum aminotransferase levels have also been observed on rare occasions. A predictable mild increase in serum lipid levels is observed following the initiation of treatment with tocilizumab; the clinical significance of this remains indeterminate, but monitoring of lipid levels is recommended. Reactivation of tuberculosis and invasive fungal infections can occur, and rare instances of gastric and intestinal rupture have been reported in patients treated with tocilizumab. These observed effects on hematopoiesis and loss of intestinal integrity reflect the potential downsides of complete blockade of IL-6–mediated functions that is inherent in the use of mAb reagents that block the interaction of IL-6 with the IL-6R and emphasize the potential advantages of more selective IL-6 pathway inhibition (see below, sgp130:Fc).

On the Horizon: Other Anti–IL-6R and Anti–IL-6 mAbs and Sgp130:Fc Reagents

Several other recombinant mAb-based reagents that inhibit both classic and trans-signaling mediated by IL-6 are in advanced clinical development. *Sarilumab* (human anti–IL-6R) has been shown in phase III trials to have demonstrated efficacy in the treatment of RA. *Vobarilizumab* is a nanobody construct consisting of an IL-6R–specific non-Fc IgG heavy chain covalently linked to serum albumin that blocks binding of IL-6 to both membrane-bound and soluble IL-6R, currently being evaluated for treatment of SLE. *Sirukumab* and *Olokizumab* are humanized anti–IL-6 constructs that have also been shown in phase II trials to have efficacy in RA. *Clazakizumab* (humanized anti–IL-6) has shown promising results for the treatment of RA as well as musculoskeletal symptoms in patients with psoriatic arthritis. By virtue of engineered aglycosylation, clazakizumab has a longer half-life than other mAb reagents targeting IL-6 or IL-6R, thereby requiring less frequent dosing; however, with higher dose ranges, the frequency of serious adverse events and infections appears to be increased.

Of potentially greater interest is the development of soluble gp130 covalently linked to immunoglobulin Fc (*sgp130:Fc*), a construct with specificity for the IL-6/sIL-6R heterodimer that mediates IL-6 trans-signaling. Selective inhibition of the trans-signaling pathway linked to inflammatory responses may be superior to complete IL-6 blockade because important physiological functions of IL-6 mediated through membrane-bound IL-6R are left intact. These differences have been emphasized in a murine model of peritonitis, where selective inhibition of IL-6 trans-signaling by sgp130:Fc as opposed to complete blockade of IL-6 signaling by anti–IL-6 mAb was shown to significantly improve survival and completely prevent intestinal epithelial cell apoptosis.[40] Observations in murine models of tuberculosis have further confirmed the absence of any impaired response to disease resolution in the setting of either sgp130:Fc transfection or administration of sgp130:Fc constructs.[41]

Interleukin-12/-23 Inhibitors

IL-12 and IL-23 are heterodimeric molecules that share a common p40 subunit but have different immunological effects on T-cell lineage development. IL-12 comprises p35/p40 subunits and promotes the development and maturation of Th1 lineage T cells, a significant source of IFN-γ. IL-23 comprises p19/p40 subunits and promotes the maturation and survival of Th17 lineage T cells. Initial strategies targeting the shared p40 subunit of IL-12/IL-23 in murine models of autoimmune inflammation yielded promising results and prompted human trials in psoriasis, RA, and IBD. Subsequent comparative IL12/IL-23 p40, IL-12 p35, and IL-23 p19 knock-out studies in the murine models suggest that the observed ameliorative effects of anti-p40 on inflammation in the models studied were by virtue of IL-23 inhibition.[42] As a consequence, selectively targeting IL-23 and IL-17 (the primary effector cytokine of Th17 cells) has evolved as a strategy to treat autoimmune inflammation.

Ustekinumab is a humanized IgG1κ mAb with specificity for the p40 subunit component of IL-12 and IL-23. Approved for the treatment of psoriasis and psoriatic arthritis, ustekinumab significantly improves skin lesions as measured by the psoriasis area and severity index (PASI), and has been shown to decrease both peripheral and axial joint/enthesis manifestations in psoriatic arthritis.[43] Although not yet approved for patients with IBD, significant improvements in Crohn disease activity have been reported in patients treated with ustekinumab, most notably among patients who failed to respond adequately to anti–TNF-α therapies.[44]

In the reported clinical trials for psoriasis, there were no increased occurrences of serious infections or other serious adverse events in the ustekinumab-treatment arms relative to the placebo-treatment arms. Higher than expected rates of infection or malignancy have also not been observed in other controlled studies of ustekinumab in spondylitis or IBD. Nonetheless, vigilance for mycobacterial, fungal, and *Salmonella* infections is recommended, given the role of IL-12 and IL-23 in host defense against these pathogens.

On the Horizon: IL-23 Specific Inhibitors

Tildrakizumab and *Guselkumab* are fully human IL-23 p19–specific mAbs currently being evaluated in clinical trials for treatment of psoriasis.

Interleukin-17 Inhibitors

IL-17 is a cytokine elaborated primarily by the Th17 lineage of effector T cells, which differentiate from Foxp3$^-$CD4$^+$ thymocytes under the influence of IL-6 and transforming growth factor-β (TGF-β) and proliferate/survive under the influence of IL-23. Th17 cells are prevalent in the inflammatory lesions of a variety of inflammatory disorders, including RA, IBD, psoriasis, and spondyloarthropathies. IL-17 is produced by a variety of innate immune cells, including neutrophils, mast cells, keratinocytes, macrophages, NK cells, NKT cells, and innate lymphoid cells (ILCs; LLC/LTi), triggering the induction and release of IL-6, TNF-α, CCL2, CCL3, MMPs from macrophages, inducing activation of osteoclasts at sites of bone resorption, and promoting the proliferation, maturation, and chemotaxis of neutrophils.[45,46] As such, IL-17 is important both in the initial innate host defense to infection and also inflammation driven by acquired immune responses. Given the multitude of effects mediated by IL-17 in perpetuating inflammation in a number of autoimmune disorders, mAb reagent–based strategies have been developed targeting IL-17 and its receptor (Table 89.2).

KEY CONCEPTS

Approaches to IL-17 Inhibition

- Selective inhibition of pathways promoting differentiation and activation products of T-helper 17 (Th17) cells is of significant benefit in the treatment of psoriasis and the seronegative spondyloarthropathies.
- Selective inhibition of interleukin (IL)-17A appears to be of marginal benefit in the treatment of inflammatory bowel disease (IBD), whereas inhibition of both IL-12 and IL-23 by targeting their shared p40 subunit appears to be of benefit in both psoriasis and IBD.

Secukinumab

Secukinumab is a fully human IgG1 mAb that selectively binds to and neutralizes IL-17A. In clinical trials, secukinumab has been shown to significantly decrease the activity of skin lesions in patients with psoriasis, decrease tender and swollen joints in patients with psoriatic arthritis, and decrease axial pain and limitation in patients with ankylosing spondylitis.[47,48] Clinical responses to secukinumab in patients with RA have been less robust. Despite the apparent role of Th17 cells in patients with IBD, no significant improvement and was observed in patients with Crohn disease, and in some of these patients, the disease worsened during treatment with secukinumab.[49]

In clinical trials of secukinumab in patients with psoriasis, rates of infection are increased over that observed in patients randomized to placebo, but not in excess of the infection rates noted with anti-TNF therapy. Increased rates of serious infection have not been observed in relatively short-term RA clinical trials with secukinumab. However, given the pivotal role of Th17 in host responses to chronic and acute infections by bacteria, parasites, and fungi, it is possible that alterations in the containment of microbial pathogens in the enteric mucosa may account for the failure of targeting IL-17 in IBD.

Ixekizumab

Ixekizumab is a humanized IgG4 mAb targeting IL-17A that has been shown to be effective and approved for the treatment of plaque psoriasis. Similar to what has been observed in clinical trials with secukinumab, rates of infection are slightly increased in patients treated with ixekizumab versus placebo, but in short-term studies conducted so far, the rate of serious infections has not been observed to be increased relative to that observed in patients randomized to placebo or a comparative arm of patients randomized to anti-TNF treatment with etanercept.[50]

On the Horizon: Il17 Receptor Antagonists

Brodalumab is an anti–IL-17Ra monoclonal antibody. Data from two identically designed phase III randomized trials support the efficacy of brodalumab for treatment of moderate to severe plaque psoriasis. Efficacy with regard to clearing skin lesions was noted to be significantly greater than that observed in a comparator group receiving ustekinumab, but *Candida* infections and neutropenia were more frequent in the brodalumab groups than in the ustekinumab and placebo groups.[51]

Other Immunomodulatory Cytokines and Cytokine Inhibitors

On the Horizon: Inhibitors of the Type 1 Interferon Pathway

With increasing recognition of the role of type 1 IFNs in the immunopathogenesis of SLE, strategies to target IFN-α for treatment of SLE have emerged. Initial trials with mAb reagents, such as *rontalizumab* or *sifalimumab* with specificity to one or more of the known IFN-α subtypes (14 known presently), have yielded modest results, emphasizing the complexity of the IFN system in SLE. This strategy is further complicated by the recognition that autoantibodies to IFN-α are often prevalent in SLE sera. The alternative approach of targeting the type 1 IFN receptor (IFN-1R) in a recent phase II trial with *anifrolumab* has yielded more encouraging results.[52] Targeting the IFN-1R may also prove efficacious in other autoimmune disorders, such as SLE and Sjögren syndrome, which are associated with an enhanced type 1 IFN signature. Other approaches targeting the production of type 1 IFNs currently under investigation include the use of mAb reagents to deplete plasmacytoid dendritic cells (pDCs), perceived to be the major source of IFN-1 in SLE.

In the relatively short-term clinical trials conducted thus far with antibodies targeting IFN-α or IFN-1R, increased susceptibility to viral pathogens has not been observed. Whether depleting pDCs, blocking one or more TLRs linked to upregulation of type 1 IFNs, blocking IFN-1R, or using mAb reagents capable of neutralizing multiple IFN subtypes critically important in SLE and related autoimmune disorders will prove efficacious but not at the expense of susceptibility to viral disease remains to be determined in phase III trials.

On the Horizon: Recombinant Promoters of Treg Function

IL-27 may modulate autoimmune inflammation via promotion of Treg lineage and function; it often opposes the action of IL-6 and is the only member of the gp130-related cytokine family to predominantly signal via the latent transcription factor signal transducer and activator of transcription 1 (STAT1) instead of STAT3.[53] In this context, IL-27 acts as a negative regulator of Th17 commitment, whereas the p28 subunit of IL-27 antagonizes IL-6–STAT3-mediated T-cell responses. Treg function may also be enhanced by recombinant *IL-35.*[54]

Complement Pathway Inhibitors

In addition to opsonization of microbial pathogens and products of apoptosis, products of complement activation play a significant role in recruitment and activation of phagocytic cells (C3a, C5a), cell membrane damage (C5b-9), and have potent procoagulant activity (C5a, C5b-9) (Table 89.3). Congenital deficiency of C3 or functional C3 impairment is typically associated with severe, if not fatal, clinical phenotypes, rendering C3 or its cleavage products unattractive targets for immunomodulation. However, inhibition of C5 appears to be well tolerated and may have beneficial effects on disorders in which enhanced cleavage of C5 contributes to disease manifestations (Table 89.3).

KEY CONCEPTS

Inhibitors of Complement Activation and Cell Migration

- Inhibition by eculizumab of C5 cleavage to its active products and resulting assembly of the membrane attack complex is of significant benefit to patients who have complement-mediated thrombotic microangiopathy syndromes occurring with atypical hemolytic–uremic syndrome, catastrophic antiphospholipid syndrome, or severe flare-ups of systemic lupus erythematosus (SLE).
- Immunization against *Neisseria* species is recommended for patients who require dosing with eculizumab.
- Antibody targeting the lymphocyte integrin addressin $\alpha_4\beta_1$ (natalizumab) has been shown to be of benefit in managing multiple sclerosis (MS) and inflammatory bowel disease (IBD) but is associated with a risk of central nervous system (CNS) JC virus activation (progressive multifocal leukoencephalopathy [PML]).
- Antibody targeting the $\alpha_4\beta_7$ integrin (vedolizumab) selectively inhibits trafficking of lymphocytes into the intestinal lamina propria and is of benefit in the management of IBD.

Eculizumab, a humanized IgG2/4κ mAb with specificity for C5, blocks the generation of C5a and C5b by the C5 convertase. The resulting failure to generate C5b impairs assembly of the C5b-9 terminal attack complex, and failure to generate C5a diminishes complement pathway mediated recruitment of phagocytic cells. Since both C5a and C5b-9 have potent procoagulant activity at sites of inflammation, complement activation is hypothesized to be critical to the generation of intravascular thrombi in thrombotic microangiopathy syndromes, including atypical hemolytic–uremic syndrome (aHUS), catastrophic anti-phospholipid syndrome, and thrombotic microangiopathy that may be seen in SLE. Eculizumab is currently approved for use in patients with aHUS but has also been reported to be of benefit in patients with these other complement activation–mediated thrombotic syndromes.[55-57]

Eculizumab inhibits terminal complement activation and therefore renders patients vulnerable to infection with encapsulated organisms. Life-threatening and fatal meningococcal infections have occurred in patients who received eculizumab, and meningococcal vaccination is recommended at least 2 weeks prior to receiving eculizumab. In cases where the risks of delaying eculizumab therapy outweigh the risk of developing a meningococcal infection, meningococcal vaccine should be administered as soon as possible. However, current meningococcal vaccines do not protect against serogroup B strains of *Meningococcus,* and thus immunized patients are still at risk of infection by this strain.

TABLE 89.3 Recombinant Inhibitors of Complement Activation and Molecules Mediating Cell Migration

Molecule	Construct	Half-Life	Dosing (Maintenance)
Eculizumab	anti-C5 (humanized IgG2/4κ)	8–15 days	300–1200 mg intravenously (IV) every 1–2 weeks
Natalizumab	anti-$\alpha_4\beta_1$ (humanized IgG4κ)	7–15 days	300 mg IV every 4 weeks
Vedolizumab	anti-$\alpha_4\beta_7$ (humanized IgG1κ)	25 days	300 mg IV every 8 weeks

Adhesion Molecule Inhibitors

Trafficking of phagocytic cells and lymphocytes across vascular endothelium is critical to the development of inflammatory lesions and associated tissue injury. The transmigration of cells across the endothelium is dependent on the upregulation of integrin/adhesion molecules, some of which have binding specificity for ligands (addressins) in specific organs, such as the CNS or the enteric mucosa. Strategies to decrease the influx of monocytes and/or T and B lymphocytes using biologicals that target adhesion molecules have, therefore, been deployed as treatments for several inflammatory disorders, including MS and IBD.

Natalizumab is a humanized IgG4κ mAb with specificity for the α_4 subunit of the very late activation antigen-4 (VLA-4)–integrin molecule expressed on lymphocytes and monocytes. Natalizumab blocks association of the VLA-4 $\alpha_4\beta_1$ and $\alpha_4\beta_7$ integrins with their respective vascular receptors, thereby limiting cell transmigration into tissue sites of inflammation in the CNS and in the intestinal mucosa. On the basis of the premise that administration of natalizumab blocks the interaction of T-cell $\alpha_4\beta_1$ with its addressin ligand on venules in the CNS, this therapeutic approach was validated in murine models of encephalomyelitis, and subsequent clinical trials confirmed natalizumab to be efficacious in decreasing the frequency of relapses of MS.[58] Natalizumab also inhibits the interaction of the $\alpha_4\beta_7$ integrin with mucosal addressin cell adhesion molecule-1 (MAdCAM-1) expressed on venules in the enteric mucosa and has been shown to induce remissions and prevent flare-ups of Crohn disease in patients who have failed to have an adequate response to anti–TNF-α therapy.[59]

Recent additional work in experimental allergic encephalomyelitis (EAE) models in different strains of mice suggests that natalizumab affects primarily the transmigration of Th1 cells, with much less of an inhibitory effect on the transmigration of Th17 cells.[60] As such, there may be differential efficacy of natalizumab in the treatment of Th1- versus Th17-driven subtypes of MS and/or IBD.

The major concern with regard to the use of natalizumab is the occurrence of reactivation of John Cunningham (JC) polyoma virus in patients who are carriers, resulting in progressive multifocal leukoencephalopathy (PML). This complication has a very low rate of occurrence, but it is quite debilitating and often fatal, resulting in dampened enthusiasm for the use of natalizumab. However, the risk of PML can be mitigated by limiting the duration of natalizumab treatment to 1 year in known carriers of JC virus (as defined by seroconversion) and not administering natalizumab concurrent with other immunosuppressive therapy.

Vedolizumab is a humanized mAb that specifically binds to $\alpha_4\beta_7$ integrin, blocking its interaction with MAdCAM-1 on intestinal endothelial cells. Since vedolizumab does not bind to or block the interaction of the $\alpha_4\beta_1$ integrin to its addressin ligand in the CNS, the risk of developing PML during treatment is substantially lower than that associated with the use of natalizumab. In clinical trials, vedolizumab has been shown to significantly improve disease activity and reduce flare-ups in patients with Crohn disease or UC.[61] This strategy may prove to be particularly useful in subsets of patients with IBD as well as clinical features of SLE; in these patients, use of anti–TNF-α agents may carry an increased risk of SLE exacerbation.

Inhibitors of B-Cell Activation

Identified targets for regulation of B-cell activation include growth and survival factors, such as B-lymphocyte stimulator (BLyS; BAFF) and its respective receptors (BAFF-R, transmembrane activator and CAML interactor [TACI], and B-cell maturation antigen [BCMA]); costimulatory receptors and their ligands, such as CD40/CD40-ligand; and cell surface receptors, such as CD22 or FcγRIIb, which engender inhibitory signaling when ligated. Given the prominent role of multiple autoantibodies and generalized B-cell activation in SLE, most of the clinical trials, to date, employing strategies targeting B-cell activation have been in patients with this disorder.

Belimumab is a recombinant human genome–derived IgG1 mAb with specificity for soluble (non–membrane-bound) BLyS (BAFF). Through ligation of the BAFF-R and TACI receptors on B lymphocytes, BLyS promotes the maturation of B cells into antibody-secreting plasmablasts. On the basis of the hypothesis that autoreactive B lymphocytes may have greater dependency on BLyS to survive and proliferate, a number of preclinical studies were undertaken to target BLyS using gene knock-out strategies or soluble receptor (TACI) transfections in murine models of SLE. Following encouraging results in these models, human studies using belimumab were undertaken, and improvement in SLE disease activity was demonstrated by using validated disease activity measures (SLE Disease Activity Index [SLEDAI] and British Isles Lupus Assessment Group [BILAG]) as well as additional secondary endpoints related to disease flare-ups and sparing of corticosteroid use.[62] Consistent with its mechanism of action of targeting a growth and survival factor, noted clinical improvements do not become manifest until after 6 months of treatment with belimumab. The majority of observed clinical improvements have been observed in the musculoskeletal, mucocutaneous, and serological domains of disease activity.[62] The potential effects on more severe neurological or renal domains of disease activity have not been assessed in controlled trials.

Significant decreases in the measured numbers of circulating activated B cells and plasmacytoid B lymphocytes are observed following 6 months' treatment with belimumab.[63] Transient increase in the number of circulating memory B cells is observed immediately after administration of belimumab, with numbers gradually returning to the pretreatment baseline level over the course of several months of treatment. Total numbers of circulating B cells are decreased by 20–25% following 1-year treatment periods, with no observed decreases in CD4 and CD8 T lymphocytes.[63] Levels of autoantibodies, including anti-dsDNA, anti-Smith, anti-SSA, and anti-cardiolipin, are decreased by 40–50% percent after the first year of treatment and continue to decrease over time with long-term treatment.[63] In contrast, total antibody levels decrease, on average, by only 15%, with no significant decreases in measured preexisting antibody titers to influenza, tetanus toxoid, or pneumococcal serotypes; furthermore, treatment with belimumab does not appear to have any significant effects on the primary immune responses to pneumococcal bacterial antigens.[64,65] In two major randomized trials, the frequency of infections was not shown to be increased in the belimumab treatment arms.

Receptors for BLyS (BAFF-R) and BCMA have also recently been identified on murine T-follicular helper (Tfh) cells, with ligation of BAFF-R promoting activation of Tfh and ligation of BCMA appearing to play a role in downregulation of Tfh responses. The significance of these findings in human SLE and what impact treatment with belimumab has on Tfh responses remain to be determined.[66]

On the Horizon: Other Inhibitors of B-Cell Activation

Atacicept is a recombinant human fusion protein containing the extracellular, ligand-binding portion of the receptor TACI and a modified Fc portion of human IgG. Atacicept binds both BLyS (BAFF) and a proliferation-inducing ligand (APRIL), thereby functioning as an antagonist to the ability of these two ligands to stimulate B lymphocytes. Amelioration of disease manifestations in murine SLE models employing TACI-Ig transfections provided a mechanistic rationale for pursuing human SLE studies using TACI-Ig as a therapeutic intervention; such studies are still ongoing. However, the therapeutic window with this approach may be relatively narrow, as membrane-bound TACI is the major receptor mediating Ig class switching during B-cell maturation, and excess of the soluble receptor administered over time may result in significant humoral immune deficiency as a result of decreases in IgG-secreting B cells and plasma cells.[67]

Blisibimod is a recombinant fusion polypeptide heterodimer consisting of a BlyS (BAFF)–binding domain covalently linked to the N-terminus of the Fc region of human IgG1. The tetravalent peptibody construct binds to both soluble and cell membrane–bound BlyS, with reported higher binding affinities compared with conventional mAbs, inhibiting the interaction of BlyS with BAFF-R and TACI. Blisibimod and has been shown to have biological effects in early-phase human trials, but the clinical significance of the membrane-binding attribute of blisibimod has not been determined.[68]

Anti-CD40-ligand mAb reagents interfere with the interaction of T-cell CD40L with B-cell CD40, thereby blocking costimulatory signals required for cognate T-cell help that promotes antigen-specific B-cell proliferative responses. Following favorable results with targeting of CD40L in murine SLE models, human trials with anti-CD40L were undertaken but halted in the context of observed thrombotic complications. Subsequent studies demonstrated upregulation of CD40L on platelets, with platelet aggregation occurring in the context of complement fixation to platelet membrane–bound anti-CD40L. Newer anti-CD40L

TABLE 89.4 **Recombinant Inhibitors of Lymphocyte Proliferation, Survival and Activation**

Molecule	Construct	Half-Life	Dosing (Maintenance)
Belimumab	aBLyS/BAFF (human IgG1λ)	19 days	10 mg/kg intravenously (IV) every 4 weeks
Abatacept	CTLA-4:IgG1Fc	13 days	10 mg/kg ; 500–1000 mg subcutaneously (SC) or IV every 4 weeks
Belatacept	CTLA-4:IgG1Fc	10 days	10 mg/kg IV every 4 weeks
Basiliximab	aCD25/IL-2Ra (chimeric IgG1κ)	7–9 days	20 mg IV (repeat ×1 at 4 days)

constructs employing PEGylated anti-CD40L F(ab)' that appear to be free of the effect on platelet activation/aggregation are being evaluated in human SLE trials.[69]

KEY CONCEPTS

Inhibitors of B-Cell and T-Cell Activation

- Inhibition of BAFF/BLyS (B lymphocyte stimulator) over a period of 6–12 months decreases the survival and maturation of autoreactive B cells, decreases autoantibody titers, and decreases lupus disease activity, with minimal impact on preexisting antibody titers to microbial pathogens.
- Inhibition of T-cell costimulation targeting CD80/86 and CD28 via exogenous cytotoxic T lymphocyte antigen-4 (CTLA-4):IgFc constructs is effective in suppressing disease activity in rheumatoid arthritis (RA).

Inhibitors of T-Cell Activation

Given the role of T-cell lymphocytes in orchestrating adaptive immune responses, selective inhibition of T-cell activation is an attractive target for modulating inflammatory disorders associated with immune responses to autoantigens or allografts. Blocking cell receptors for T-cell growth factors, such as IL-2, has been employed to prevent allograft rejection and therefore may be potentially of use in the management of autoimmune disorders. Since productive immune responses are not generated in the absence of effective costimulatory signals, blocking T-cell costimulation has also been an attractive target for treatment of inflammatory diseases driven by autoreactive T cells (Table 89.4).

Basiliximab

Basiliximab is a chimeric (murine/human) mAb that blocks the α chain of the IL-2R complex expressed on activated T lymphocytes, inhibiting the binding of IL-2 to IL-2R (CD25). Basiliximab is approved for and used primarily in induction regimens for prevention of rejection of transplanted allograft.[70] Uncontrolled small case series have reported manifestations of systemic sclerosis and pulmonary complications associated with amyopathic dermatomyositis responding well to (off-label) treatment with basiliximab.[71,72]

Abatacept

Abatacept (CTLA-4 Ig) is a recombinant human protein consisting of the extracellular domain of CTLA-4 linked to the Fc portion of IgG1. Prominent among the T-cell costimulatory molecules is CD28, which binds CD80/CD86 on antigen-presenting cells (APCs). In the context of T-cell activation, CTLA-4 expression is upregulated on the surface of T cells. The binding avidity of CD80/CD86 for CTLA-4 is considerably higher than that for CD28, resulting in preferential ligation of CTLA-4, thereby disrupting further costimulation as well as engendering inhibitory signaling in the T cell as a consequence of CTLA-4 ligation. Soluble CTLA-4 Ig is not bound to the surface of the T cell and therefore does not engender inhibitory signaling; rather, the much higher avidity of the construct for CD80/86 inhibits T cell costimulation through CD28 (Table 89.4).

Although the time to maximum clinical response is somewhat longer than that observed with TNF-α inhibitors, abatacept is effective in reducing disease activity and inhibiting progression of structural damage in RA that is unresponsive to treatment with methotrexate.[73] Abatacept has no direct impact on phagocytic cell responses, and its use in patients with RA may be associated with fewer bacterial infection complications relative to patients on anti-TNF therapy. It is, nonetheless, recommended that treatment with abatacept be withheld in the context of serious intercurrent microbial infections and not be used in conjunction with other biological therapies targeting inflammation. Abatacept may be associated with an increased risk of lung cancer; however, the occurrence of lymphoma in patients with RA treated with abatacept has not been shown to exceed the expected occurrence in patients with RA. In contrast to TNF-α inhibitors, use of abatacept does not appear to promote autoimmune complications and may therefore be a preferred option for treatment of patients with RA who need biological therapy and have overlapping features of SLE or other autoimmune disorders.

Belatacept is a second-generation CTLA-4 Ig, which, compared with abatacept, has higher binding to both CD80 and CD86. Currently used primarily in organ transplantation, belatacept is associated with improved patient and renal allograft survival compared with cyclosporine.[74] Although perceived to be of potential use in autoimmune disorders, studies of belatacept use in disorders, such as SLE and RA, have not yet been reported.

Inhibitors of Mast Cell Activation

Omalizumab

Omalizumab is a recombinant humanized mAb that binds to the Cε3 domain of IgE. The binding domain is the same site at which IgE normally binds to both high- and low-affinity FcεRI on mast cells and basophils; as a consequence, free IgE is prevented from binding to the mast cell FcεRI receptor. Omalizumab is specific to IgE and does not bind to IgG or IgA. Omalizumab also cannot bind to FcεRI or to IgE already attached to FcεRI and therefore does not interact with cell-bound IgE or activate mast cells or basophils. Omalizumab is most useful, and presently approved for use in, the treatment of poorly controlled asthma despite inhaled corticosteroid use and in the setting of documented sensitization to a perennial allergen in the setting of serum IgE levels ≥30 IU/mL.[75] It is also approved for use in adults and adolescents with chronic idiopathic urticaria who remain symptomatic despite H1 antihistamine treatment. Omalizumab may also decrease the severity of asthma in patients with nonatopic (intrinsic) asthma, occupational asthma, virus-induced asthma, and eosinophilic granulomatosis with polyangiitis, but its use has not been fully studied in these populations to merit approved labeling.[76] Controlled and long-term use studies of omalizumab have shown the incidence of adverse events is not significantly increased (Table 89.5).

TABLE 89.5 Recombinant Inhibitors of Mast Cell and Eosinophil Activation

Molecule	Construct	Half-Life	Dosing (Maintenance)
Omalizumab	aIgE (humanized IgG1κ)	24–26 days	150–300 mg subcutaneously (SC) every 2–4 weeks
Mepolizumab	aIL-5 (humanized IgG1κ)	16–22 days	100 mg SC every 4 weeks
Reslizumab	aIL-5 (humanized IgG4κ)	24 days	3 mg/kg intravenously (IV) every 4 weeks

Observed clinicopathological responses to treatment with omalizumab include a marked downregulation of the surface expression of FcεRI on basophils and mast cells and reductions in FcεR1-mediated production of Th2 cytokines by basophils. Markers of airway inflammation are significantly reduced, with demonstrated reductions in sputum eosinophil counts and reduced numbers of eosinophils, CD3, CD4, and CD8 T lymphocytes, B lymphocytes, IL-4–positive cells, and IgE-positive cells in the bronchial mucosa.[77]

KEY CONCEPTS

Inhibitors of IgE and IL-5R as Treatment for Severe Allergies

- Antibodies targeting the Fc binding domain of immunoglobulin E (IgE; omalizumab) are effective in the treatment of severe asthma and recurrent mast-cell mediated chronic urticaria.
- Antibodies targeting the interleukin (IL)-5R binding domains of IL-5 are effective in the treatment of resistant asthma associated with eosinophilia as well as syndromes associated eosinophil-mediated tissue injury.

Inhibitors of Eosinophil Activation

Eosinophils mediate airway inflammation in patients with asthma, contribute to vascular and organ inflammation in patients with EGPA, and may mediate organ injury in other hypereosinophilic syndromes. IL-5 is a potent cytokine mediator of eosinophil hematopoiesis and has been shown to mediate eosinophilic inflammation in the airways. Although corticosteroids are potent suppressors of eosinophil survival, proliferation, and function, biologicals targeting IL-5 or the IL-5 receptor may prove useful as steroid-sparing therapy in managing chronic disorders mediated predominantly by eosinophils.

Mepolizumab is an mAb that binds to IL-5, thereby inhibiting its binding to the alpha chain of the IL-5R complex expressed on the surface of eosinophils. Mepolizumab has demonstrated efficacy in patients with severe asthma who have blood eosinophil counts of ≥150/µL and has been approved by the US Food and Drug Administration (FDA) for add-on, maintenance treatment of severe asthma in patients who are age ≥12 years and have demonstrated hypereosinophilia.[78] However, given the impact of corticosteroids on peripheral eosinophil counts, limiting selection of patients for treatment based on eosinophil count thresholds should not be absolute in the setting of required intercurrent steroid use. Adverse events associated with use of mepolizumab include hypersensitivity reactions and possible increased occurrence of herpes zoster.

TABLE 89.6 Recombinant T Cell– and B Cell–Depleting Agents

Molecule	Construct	Half-Life	Dosing (Maintenance)
Rituximab	aCD20 (chimeric IgG1κ)	18–23 days	375 mg/m^2 intravenously (IV) per week ×4 (q4–6 months) 1000 mg IV q2 week ×2 (q4–6 months)
Ofatumumab	aCD20 (human IgG1κ)	17 days	1000 mg IV every 4–8 weeks
Obinutuzumab	aCD20 (humanized IgG1κ)	24–36 days	1000 mg IV every 2 weeks ×2
Alemtuzumab	aCD52 (humanized IgG1κ)	1–14 days	12 mg IV daily ×5 days 10–30 mg subcutaneously (SC) 3 times per week
Brentuximab-vedotin	aCD30 (chimeric IgG1κ):MMAE	4–6 days	0.3–1.8 mg/kg IV every 3 weeks

Reslizumab is an mAb that targets IL-5R binding domains of IL-5. Administered intravenously, reslizumab has been shown effective in decreasing the frequency and severity of asthma exacerbations and is approved by the FDA for use in adult patients with demonstrated eosinophilia and otherwise treatment-resistant asthma.[79] Adverse events include occasional episodes of severe hypersensitivity reactions during infusions.

On the Horizon: Anti-IL-5 Receptor Reagents

The anti–IL-5R antibody *benralizumab* blocks activation of eosinophils by IL-5 but has an additional property of depleting IL-5R–bearing eosinophils and basophils via enhanced ADCC.[80] The recombinant antibody is afucosylated, rendering a higher affinity for FcγRIII and enhanced ADCC of IL-5R–bearing cells. Studies are currently underway to confirm the clinical efficacy of benralizumab in managing asthma and to determine its utility in managing other disorders mediated by tissue infiltration of eosinophils.

B Cell– and T Cell–Depleting Agents

There are several mAb-based constructs that have been developed primarily for the treatment of lymphoid malignancies and/or prevention of transplant rejection. Since B cells have a demonstrated role in the generation of autoantibodies as well as antigen presentation to and costimulatory support for autoreactive T cells, strategies to deplete B lymphocytes have been successfully employed in the treatment of autoimmune and inflammatory diseases. In subsequent randomized trials, some of these agents (rituximab) have been shown to have efficacy and are approved for use in the treatment of both RA and ANCA-associated vasculitis syndromes. Case reports have also documented the successful use of mAb-mediated B cell– and also T cell–depleting strategies in the management of patients with severe, refractory flare-ups of SLE, autoimmune inflammatory myopathies, and GvHD (Table 89.6).

Rituximab

Rituximab is a chimeric IgG1 mAb with specificity for the B-lymphocyte surface antigen CD20, a cell-surface molecule expressed on the surface of pre-B through activated mature B cells. Rituximab induces lysis of $CD20^+$ B cells by several mechanisms, including complement activation, ADCC, and induction of apoptosis. Induction regimens using four weekly doses or two larger doses administered 2 weeks apart appear to be equally effective at effecting the depletion of circulating $CD20^+$ B cells, which can last up to ≥9 months after a single course of therapy. Treatment with rituximab has been shown to cause significant transient decreases in inflammatory CD4 T cells that express IL-17; the extent to which this is caused by an indirect consequence of $CD20^+$ B-cell depletion or depletion of identified subsets of $CD4^+CD20^+$ T cells is uncertain.[81,82]

Approved for use in patients with RA failing to respond to initial disease-modifying antirheumatic drug (DMARD) therapy, rituximab has been shown to improve the signs and symptoms of disease, functional status, and quality of life and to slow radiographic progression of disease in patients with RA.[83] Greater clinical responses are observed among patients with RA who are seropositive for rheumatoid factor (RF) and/or anticitrullinated peptide (CCP) autoantibody compared with patients seronegative for RF or CCP.[84] The optimal treatment schedule for use of rituximab in RA remains to be determined, but retreatment is generally recommended no sooner than 6 months following initial treatment for patients in whom the disease flares up in association with recovery of B-cell counts and for those who had initially responded to rituximab therapy.

Rituximab has also been used with considerable success and is approved for use in the management of ANCA-associated and other vasculitis syndromes. Rituximab plus glucocorticoid therapy was shown to be as effective as cyclophosphamide in the treatment of patients with granulomatosis with polyangiitis (GPA) or microscopic polyangiitis (MPA).[85] In a subgroup of patients with relapsing ANCA vasculitis, rituximab was, in fact, superior to cyclophosphamide in inducing disease remissions.[85] In patients with cryoglobulin syndromes occurring in the context of either HCV infection or lymphoma, use of rituximab has been shown to decrease cryoglobulin and constituent Ig titers and hasten resolution of cryoglobulin manifestations, including skin ulcers, glomerulonephritis, peripheral neuropathy, arthritis, and/or hyperviscosity complications.[86]

Although not yet formally approved, rituximab has been used with reported success in the management of other autoantibody-mediated disorders, including SLE, primary Sjögren syndrome, inflammatory myopathy, chronic inflammatory demyelinating polyneuropathy (CIDP), MS, and pemphigus.[87-92]

Despite the potential for immunodeficiency-related to depletion of $CD20^+$ B cells, only minimal increases in serious or opportunistic infections have been reported in patients with RA or ANCA vasculitis treated with repeated cycles of rituximab. Mild decreases in the overall levels of serum Igs may be observed during treatment, but Ig levels are rarely depleted, likely due to the preservation of more mature B cells and plasma cells that have lost surface expression of CD20. However, if used concomitantly with other immunosuppressive agents impacting lymphocyte proliferation, significant hypogammaglobulinemia may ensue over time as a result of the inability to replenish the plasma cell compartment.[93] Use of rituximab carries a safety warning related to an observed increased risk of viral infections, including CMV, herpes simplex virus (HSV), varicella-zoster virus (VZV), HBC, and JC virus. Reactivation of JC virus, latent in more than 80% of the general population, can occur in immunosuppressed patients to cause PML, a fatal demyelinating disease of the CNS. Cases of PML have been reported among patients with hematological malignancies, SLE, and RA who were treated with rituximab, but PML usually occurs in the context of other therapies impacting lymphocyte survival and proliferation.[94]

Ofatumumab

Ofatumumab is a fully human mAb that binds to an epitope encompassing both small and large loops of the extracellular domain of the CD20 cell surface antigen on B lymphocytes. The binding epitope of ofatumumab is distinct from that of rituximab, residing more proximal to the cell membrane. In comparative studies with rituximab using chronic lymphocytic leukemia (CLL) cells, ofatumumab elicits similar ADCC but elicits greater complement-dependent cytotoxicity (CDC) because of the greater proximity of the binding site to the cell membrane and/or binding affinity to CD20 epitopes.[95]

Obinutuzumab

Obinutuzumab (GA101) is a recombinant, humanized, and glycoengineered type II CD20 mAb of the IgG1 isotype that targets the extracellular loop of the CD20 transmembrane expressed on the surface of pre-B and mature B lymphocytes. Relative to rituximab and ofatumumab, glycoengineering of the Fc portion of obinutuzumab results in a higher affinity for FcγRIII receptors on immune effector cells, such as NK cells and phagocytic cells. *In vitro* studies with B-cell lymphoma–derived cell lines show that compared with rituximab and ofatumumab, obinutuzumab mediates greater induction of direct cell death and effector cell–mediated ADCC and cellular phagocytosis. However, compared with rituximab and ofatumumab, CDC is significantly reduced with obinutuzumab.[96]

Alemtuzumab

Alemtuzumab is an mAb with specificity for CD52, an antigen present on the surface of B and T lymphocytes as well as the majority of monocytes, macrophages, NK cells, and a subpopulation of neutrophils. Approved for use in the treatment of B-cell CLL and relapsing-remitting MS, alemtuzumab has also been used with success for the treatment of T-cell prolymphocytic leukemia, prevention and treatment of acute GvHD, and prevention of allograft rejection. Alemtuzumab has been used off-label with reported success in the treatment of patients with severe SLE and Behçet disease refractory to other treatments.[97,98] Despite the depletion of T lymphocyte, B lymphocyte, NK cell, and monocyte populations following treatment, reported rates of serious infections following treatment with alemtuzumab are not significantly increased compared with other immunosuppressive regimens employed to manage the disorders for which it is used. However, there is a significant occurrence of secondary autoimmunity following use of alemtuzumab, most often manifest as autoimmune thyroid disease and immune thrombocytopenia. The observed autoimmune complications may be caused by homeostatic proliferation of self-reactive memory T cells in the absence of an effective T regulatory response during immune reconstitution.[99]

Brentuximab Vedotin

Brentuximab vedotin (BTX-v) is an mAb with specificity for CD30 that is covalently linked to the antitubulin agent monomethyl auristatin E. Expressed on subpopulations of T lymphocytes, some populations of B cells, and most notably Hodgkin

KEY CONCEPTS

B-Cell and T-Cell Depletion Strategies

- Monoclonal antibody (mAb)–mediated depletion of CD20$^+$ B lymphocytes is effective in the treatment of antineutrophil cytoplasmic antibody (ANCA)–associated vasculitis, cryoglobulin syndromes, and rheumatoid arthritis (RA); the best responses in patients with RA are those with significant elevations in rheumatoid factor (RF) and anti–citrullinated peptide (CCP) antibodies.
- There are subsets of T cells that also express CD20, and part of the clinical efficacy of anti-CD20 mAb reagents may be attributed, in part, to depletion of these subsets of T cells.
- mAb reagents with greater binding affinity and binding sites more proximate to the cell membrane appear to have greater efficacy depleting CD20$^+$ lymphocytes.
- Depletion of CD20$^+$ lymphocytes has been associated with reactivation of hepatitis B virus (HBV), and vigilance for latent HBV reactivation is prudent when using anti-CD20 mAb reagents.
- mAb reagents targeting CD52 deplete both T-helper cells as well as regulatory T cell (Treg) subsets; imbalances in relative homeostatic T-cell reconstitution following treatment may result in autoimmune complications.

disease Reed-Sternberg cells, BTX-v is presently approved for use in the treatment of patients with Hodgkin lymphoma as well as of those with refractory non-Hodgkin lymphomas expressing CD30. Enhanced expression of CD30 on CD8 T cells as well increased levels of sCD30 have been observed in patients with active SLE, and BTX-v is currently being evaluated for potential clinical efficacy in the management of SLE.[100]

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Rosenzwajg M, Churlaud G, Mallone R, et al. Low-dose interleukin-2 fosters a dose-dependent regulatory T cell tuned milieu in T1D patients. J Autoimmun 2015;58:48–58.
2. von Spee-Mayer C, Siegert E, Abdirama D, et al. Low-dose interleukin-2 selectively corrects regulatory T cell defects in patients with systemic lupus erythematosus. Ann Rheum Dis 2016;75:1407–15.
3. Saadoun D, Resche Rigon M, Thibault V, et al. Peg-IFNα/ribavirin/protease inhibitor combination in hepatitis C virus associated mixed cryoglobulinemia vasculitis: results at week 24. Ann Rheum Dis 2014;73(5):831–7.
4. Wilson WH, Kingma DW, Raffeld M, et al. Association of lymphomatoid granulomatosis with Epstein-Barr viral infection of B lymphocytes and response to interferon-alpha 2b. Blood 1996;87(11):4531–7.
5. Park JY, Chung YR, Lee K, et al. Clinical experience of interferon alfa-2a treatment for refractory uveitis in Behçet's disease. Yonsei Med J 2015;56(4):1158–62.
6. Kötter I, Hamuryudan V, Oztürk ZE, et al. Interferon therapy in rheumatic diseases: state-of-the-art 2010. Curr Opin Rheumatol 2010;22(3):278–83.
7. Kasper LH, Reder AT. Immunomodulatory activity of interferon-beta. Ann Clin Transl Neurol 2014;1(8):622–31.
8. Bonaci-Nikolic B, Jeremic I, Andrejevic S, et al. Anti-double stranded DNA and lupus syndrome induced by interferon-beta therapy in a patient with multiple sclerosis. Lupus 2009;18(1):78–80.
9. The International Chronic Granulomatous Disease Cooperative Study Group. A controlled trial of interferon gamma to prevent infection in chronic granulomatous disease. N Engl J Med 1991;324(8):509.
10. Jackson SH, Miller GF, Segal BH, et al. IFN-gamma is effective in reducing infections in the mouse model of chronic granulomatous disease (CGD). J Interferon Cytokine Res 2001;21(8):567–73.
11. Ridderstad A, Abedi-Valugerdi M, Möller E. Cytokines in rheumatoid arthritis. Ann Med 1991;23(3):219–23.
12. Aaltonen KJ, Virkki LM, Malmivaara A, et al. Systematic review and meta-analysis of the efficacy and safety of existing TNF blocking agents in treatment of rheumatoid arthritis. PLoS ONE 2012;7(1):e30275.
13. Anandarajah AP, Ritchlin CT. Pathogenesis of psoriatic arthritis. Curr Opin Rheumatol 2004;16(4):338–43.
14. Boehncke WH, Prinz J, Gottlieb AB. Biologic therapies for psoriasis. A systematic review. J Rheumatol 2006;33(7):1447.
15. Li Y, Li A, Strait K, et al. Endogenous TNF-alpha lowers maximum peak bone mass and inhibits osteoblastic Smad activation through NF-κB. J Bone Miner Res 2007;22(5):646–55.
16. Peyrin-Biroulet L, Deltenre P, de Suray N, et al. Efficacy and safety of tumor necrosis factor antagonists in Crohn's disease: meta-analysis of placebo-controlled trials. Clin Gastroenterol Hepatol 2008;6(6):644.
17. Hubbard VG, Friedmann AC, Goldsmith P. Systemic pyoderma gangrenosum responding to infliximab and adalimumab. Br J Dermatol 2005;152(5):1059–61.
18. Vallet H, Riviere S, Sanna A, et al; French Behçet Network. Efficacy of anti-TNF alpha in severe and/or refractory Behçet's disease: multicenter study of 124 patients. J Autoimmun 2015;62:67–74.
19. Kisacik B, Pamuk ON, Onat AM, et al. Characteristics Predicting Tuberculosis Risk under Tumor Necrosis Factor-α Inhibitors: report from a Large Multicenter Cohort with High Background Prevalence. J Rheumatol 2016;43(3):524–9.
20. Tsiodras S, Samonis G, Boumpas DT, et al. Fungal infections complicating tumor necrosis factor alpha blockade therapy. Mayo Clin Proc 2008;83(2):181–94.
21. Hellgren K, Dreyer L, Arkema EV, et al; ARTIS Study Group, For the DANBIO Study Group. Cancer risk in patients with spondyloarthritis treated with TNF inhibitors: a collaborative study from the ARTIS and DANBIO registers. Ann Rheum Dis 2016 May 4.
22. Ko JM, Gottlieb AB. Kerbleski JF. Induction and exacerbation of psoriasis with TNF-blockade therapy: a review and analysis of 127 cases. J Dermatol Treat 2009;20(2):100–8.
23. Magnano MD, Robinson WH, Genovese MC. Demyelination and inhibition of tumor necrosis factor (TNF). Clin Exp Rheumatol 2004;S22(5 Suppl. 35):S134–40.
24. De Bandt M, Sibilia J, Le Loët X, et al; Club Rhumatismes et Inflammation. Systemic lupus erythematosus induced by anti-tumour necrosis factor alpha therapy: a French national survey. Arthritis Res Ther 2005;7(3):R545–51.
25. Cantaert T, De Rycke L, Mavragani CP, et al. Exposure to nuclear antigens contributes to the induction of humoral autoimmunity during tumour necrosis factor αblockade. Ann Rheum Dis 2009;68(6):1022–9.
26. Kranich J, Krautler NJ, Heinen E, et al. Follicular dendritic cells control engulfment of apoptotic bodies by secreting Mfge8. J Exp Med 2008;205(6):1293–302.
27. Church LD, Cook GP, McDermott MF. Primer: inflammasomes and interleukin 1beta in inflammatory disorders. Nat Clin Pract Rheumatol 2008;4(1):34–42.
28. Thueringer JT, Doll NK, Gertner E. Anakinra for the treatment of acute severe gout in critically ill patients. Semin Arthritis Rheum 2015;45(1):81–5.
29. Bruck N, Suttorp M, Kabus M, et al. Rapid and sustained remission of systemic juvenile idiopathic arthritis-associated macrophage activation syndrome through treatment with anakinra and corticosteroids. J Clin Rheumatol 2011;17(1):23–7.
30. Ortiz-Sanjuán F, Blanco R, Riancho-Zarrabeitia L, et al. Efficacy of Anakinra in Refractory Adult-Onset Still's Disease: Multicenter Study of 41 Patients and Literature Review. Medicine (Baltimore) 2015;94(39):e1554.
31. Shakoory B, Carcillo JA, Chatham WW, et al. Interleukin-1 Receptor Blockade Is Associated With Reduced Mortality in Sepsis Patients With

Features of Macrophage Activation Syndrome: Reanalysis of a Prior Phase III Trial. Crit Care Med 2016;44(2):275–81.

32. Schlesinger N, Alten RE, Bardin T, et al. Canakinumab for acute gouty arthritis in patients with limited treatment options: results from two randomised, multicentre, active-controlled, double-blind trials and their initial extensions. Ann Rheum Dis 2012;71:1839–48.
33. Veldhoen M, Hocking RJ, Atkins CJ, et al. TGF-beta in the context of an inflammatory cytokine milieu supports de novo differentiation of IL-17-producing T cells. Immunity 2006;24(2):179–89.
34. Tilg H, Trehu E, Atkins MB, et al. Interleukin-6 (IL-6) as an anti-inflammatory cytokine: induction of circulating IL-1 receptor antagonist and soluble tumor necrosis factor receptor p55. Blood 1994;83(1):113–18.
35. Grivennikov S, et al. IL-6 and STAT3 signaling is required for survival of intestinal epithelial cells and colitis associated cancer. Cancer Cell 2009;16(2):103–13.
36. Genovese MC, McKay JD, Nasonov EL, et al. Interleukin-6 receptor inhibition with tocilizumab reduces disease activity in rheumatoid arthritis with inadequate response to disease-modifying antirheumatic drugs. Arthritis Rheum 2008;58:2968–80.
37. Yokota S, Imagawa T, Mon M. Efficacy and safety of tocilizumab in patients with systemic-onset juvenile idiopathic arthritis: a randomized, double-blind, placebo controlled, withdrawal phase III trial. Lancet 2008;371:998–1006.
38. Ortiz-Sanjuán F, Blanco R, Calvo-Rio V, et al. Efficacy of tocilizumab in conventional treatment-refractory adult-onset Still's disease: multicenter retrospective open-label study of thirty-four patients. Arthritis Rheumatol 2014;66(6):1659–65.
39. Kawabata H, Kotani S, Matsumura Y, et al. Successful treatment of a patient with multicentric Castleman's disease who presented with thrombocytopenia, ascites, renal failure and myelofibrosis using tocilizumab, an anti-interleukin-6 receptor antibody. Intern Med 2013;52(13):1503–7.
40. Barkhausen T, Tschernig T, Rosenstiel P, et al. Selective blockade of interleukin-6 trans-signaling improves survival in a murine polymicrobial sepsis model. Crit Care Med 2011;39(6):1407–13.
41. Sodenkamp J, Waetzig GH, Scheller J, et al. Therapeutic targeting of interleukin-6 trans-signaling does not affect the outcome of experimental tuberculosis. Immunobiology 2012;217(10):996–1004.
42. Murphy CA, Langrish CL, Chen Y, et al. Divergent pro- and antiinflammatory roles for IL-23 and IL-12 in joint autoimmune inflammation. J Exp Med 2003;198(12):1951–7.
43. Ritchlin C, Rahman P, Kavanaugh A, et al; PSUMMIT 2 Study Group. Efficacy and safety of the anti-IL-12/23 p40 monoclonal antibody, ustekinumab, in patients with active psoriatic arthritis despite conventional non-biological and biological anti-tumour necrosis factor therapy: 6-month and 1-year results of the phase 3, multicentre, double-blind, placebo-controlled, randomised PSUMMIT 2 trial. Ann Rheum Dis 2014;73(6):990–9.
44. Wils P, Bouhnik Y, Michetti P, et al; Groupe d'Etude Thérapeutique des Affections Inflammatoires du Tube Digestif (GETAID). Subcutaneous Ustekinumab Provides Clinical Benefit for Two-Thirds of Patients With Crohn's Disease Refractory to Anti-Tumor Necrosis Factor Agents. Clin Gastroenterol Hepatol 2016;14(2):242–50.
45. Isailovica N, Daigob K, Mantovanib A, et al. Interleukin-17 and innate immunity in infections and chronic inflammation. J Autoimmunity 2015;60:1–11.
46. Menon B, Gullick NJ, Walter GJ, et al. Interleukin-17+CD8+ T cells are enriched in the joints of patients with psoriatic arthritis and correlate with disease activity and joint damage progression. Arthritis Rheumatol 2014;66(5):1272–81.
47. Mease PJ, McInnes IB, Kirkham B, et al; FUTURE 1 Study Group. Secukinumab Inhibition of Interleukin-17A in Patients with Psoriatic Arthritis. N Engl J Med 2015;373(14):1329–39.
48. Baeten D, Sieper J, Braun J, et al; MEASURE 1 Study Group; MEASURE 2 Study Group. Secukinumab, an Interleukin-17A Inhibitor, in Ankylosing Spondylitis. N Engl J Med 2015;373(26):2534–48.
49. Hueber W, Sands B, Lewitzky S, et al. Secukinumab, a human anti-IL-17A monoclonal antibody, for moderate to severe Crohn's disease: unexpected results of a randomised, double-blind placebo-controlled trial. Gut 2012;61:1693–700.
50. Blauvelt A, Papp KA, Griffiths CEM, et al. Efficacy and Safety of Switching to Ixekizumab in Etanercept Non-Responders: A Subanalysis from Two Phase III Randomized Clinical Trials in Moderate-to-Severe Plaque Psoriasis (UNCOVER-2 and -3). Am J Clin Dermatol 2017;18(2):273–80.
51. Lebwohl M, Strober B, Menter A, et al. Phase 3 Studies Comparing Brodalumab with Ustekinumab in Psoriasis. N Engl J Med 2015;373(14):1318–28.
52. Mathian A, Hie M, Cohen-Aubart F, et al. Targeting interferons in systemic lupus erythematosus: current and future prospects. Drugs 2015;75(8):835–46.
53. Meka RR, Venkatesha SH, Dudics S, et al. IL-27-induced modulation of autoimmunity and its therapeutic potential. Autoimmun Rev 2015;14(12):1131–41.
54. Cai Z, Wong CK, Dong J, et al. Remission of systemic lupus erythematosus disease activity with regulatory cytokine interleukin (IL)-35 in Murphy Roths Large (MRL)/lpr mice. Clin Exp Immunol 2015;181(2):253–66.
55. Cofiell R, Kukreja A, Bedard K, et al. Eculizumab reduces complement activation, inflammation, endothelial damage, thrombosis, and renal injury markers in aHUS. Blood 2015;125(21):3253–62.
56. Zikos TA, Sokolove J, Ahuja N, et al. Eculizumab Induces Sustained Remission in a Patient With Refractory Primary Catastrophic Antiphospholipid Syndrome. J Clin Rheumatol 2015;21(6):311–13.
57. El-Husseini A, Hannan S, Awad A, et al. Thrombotic microangiopathy in systemic lupus erythematosus: efficacy of eculizumab. Am J Kidney Dis 2015;65(1):127–30.
58. Miller DH, Khan OA, Sheremata WA, et al. International Natalizumab Multiple Sclerosis Trial Group. A controlled trial of natalizumab for relapsing multiple sclerosis. N Engl J Med 2003;348(1):15.
59. Ghosh S, Goldin E, Gordon F, et al. Natalizumab for active Crohn's disease. N Engl J Med 2003;348(1):24–32.
60. Glatigny S, Duhen R, Oukka M, et al. Cutting edge: loss of α4 integrin expression differentially affects the homing of Th1 and Th17 cells. J Immunol 2011;187(12):6176–9.
61. Sands BE, Feagan BG, Rutgeerts P, et al. Effects of vedolizumab induction therapy for patients with Crohn's disease in whom tumor necrosis factor antagonist treatment failed. Gastroenterology 2014;147(3):618–27.
62. Manzi S, Sánchez-Guerrero J, Merrill JT, et al; BLISS-52 and BLISS-76 Study Groups. Effects of belimumab, a B lymphocyte stimulator-specific inhibitor, on disease activity across multiple organ domains in patients with systemic lupus erythematosus: combined results from two phase III trials. Ann Rheum Dis 2012;71(11):1833–8.
63. Stohl W, Hiepe F, Latinis KM, et al. Belimumab reduces autoantibodies, normalizes low complement, and reduces select B-cell populations in patients with systemic lupus erythematosus. Arthritis Rheum 2012;64(7):2328–37.
64. Chatham WW, Wallace DJ, Stohl W, et al; on behalf of the BLISS-76 Study Group. Effect of Belimumab on Vaccine Antigen Antibodies to Influenza, Pneumococcal, and Tetanus Vaccines in Patients with Systemic Lupus Erythematosus in the BLISS-76 Trial. J Rheumatol 2012;39(8):1632–40.
65. Chatham WW, Chadha A, Fettiplace J, et al. Arthritis Rheumatol 2016;68:748.
66. Coquery CM, Loo WM, Wade NS, et al. BAFF regulates follicular helper T cells and affects their accumulation and interferon-γ production in autoimmunity. Arthritis Rheumatol 2015;67(3):773–84.
67. He B, Santamaria R, Xu W, et al. The transmembrane activator TACI triggers immunoglobulin class switching by activating B cells through the adaptor MyD88. Nat Immunol 2010;11(9):836–45.
68. Stohl W, Merrill JT, Looney RJ, et al. Treatment of systemic lupus erythematosus patients with the BAFF antagonist "peptibody"

blisibimod (AMG 623/A-623): results from randomized, double-blind phase 1a and phase 1b trials. Arthritis Res Ther 2015;17:215.
69. Tocoian A, Buchan P, Kirby H, et al. First-in-human trial of the safety, pharmacokinetics and immunogenicity of a PEGylated anti-CD40L antibody fragment (CDP7657) in healthy individuals and patients with systemic lupus erythematosus. Lupus 2015;24(10):1045–56.
70. Brennan DC, Daller JA, Lake KD, et al. Thymoglobulin Induction Study Group Rabbit antithymocyte globulin versus basiliximab in renal transplantation. N Engl J Med 2006;355(19):1967.
71. Zou J, Li T, Huang X, et al. Basiliximab may improve the survival rate of rapidly progressive interstitial pneumonia in patients with clinically amyopathic dermatomyositis with anti-MDA5 antibody. Ann Rheum Dis 2014;73(8):1591–3.
72. Becker MO, Brückner C, Scherer HU, et al. The monoclonal anti-CD25 antibody basiliximab for the treatment of progressive systemic sclerosis: an open-label study. Ann Rheum Dis 2011;70(7):1340–1.
73. Maxwell L, Singh JA. Abatacept for rheumatoid arthritis. Cochrane Database Syst Rev 2009;(4):CD007277.
74. Vincenti F, Rostaing L, Grinyo J, et al. Belatacept and long-term outcomes in kidney transplantation. N Engl J Med 2016;374(4):333–43.
75. Easthope S, Jarvis B. Omalizumab. Drugs 2001;61(2):253.
76. Jachiet M, Samson M, Cottin V, et al; French Vasculitis Study Group (FVSG). Anti-IgE monoclonal antibody (omalizumab) in refractory and relapsing eosinophilic granulomatosis with polyangiitis (Churg-Strauss): data from 17 patients. Arthritis Rheumatol 2016;68(9):2274–82.
77. Oliver JM, Tarleton CA, Gilmartin L, et al. Reduced FcεRI-mediated release of asthma-promoting cytokines and chemokines from human basophils during omalizumab therapy. Int Arch Allergy Immunol 2010;151:275.
78. Ortega HG, Liu MC, Pavord ID, et al; MENSA Investigators. Mepolizumab treatment in patients with severe eosinophilic asthma. N Engl J Med 2014;371(13):1198–207.
79. Castro M, Zangrilli J, Wechsler ME, et al. Reslizumab for inadequately controlled asthma with elevated blood eosinophil counts: results from two multicentre, parallel, double-blind, randomised, placebo-controlled, phase 3 trials. Lancet Respir Med 2015;3(5):355–66.
80. Laviolette M, Gossage DL, Gauvreau G, et al. Effects of benralizumab on airway eosinophils in asthmatic patients with sputum eosinophilia. J Allergy Clin Immunol 2013;132(5):1086–96.
81. Mélet J, Mulleman D, Goupille P, et al. Rituximab-induced T cell depletion in patients with rheumatoid arthritis: association with clinical response. Arthritis Rheum 2013;65(11):2783–90.
82. Alunno A, Carubbi F, Bistoni O, et al. Interleukin (IL)-17-producing pathogenic T lymphocytes co-express CD20 and are depleted by rituximab in primary Sjögren's syndrome: a pilot study. Clin Exp Immunol 2016;184(3):284–92.
83. Lopez-Olivo MA, Amezaga Urruela M, McGahan L, et al. Rituximab for rheumatoid arthritis. Cochrane Database Syst Rev 2015;(1):CD007356.
84. Sellam J, Hendel-Chavez H, Rouanet S, et al. B cell activation biomarkers as predictive factors for the response to rituximab in rheumatoid arthritis: a six-month, national, multicenter, open-label study. Arthritis Rheum 2011;63(4):933–8.
85. Stone JH, Merkel PA, Spiera R, et al; RAVE-ITN Research Group. Rituximab versus cyclophosphamide for ANCA-associated vasculitis. N Engl J Med 2010;363(3):221.
86. De Vita S, Quartuccio L, Isola M, et al. A randomized controlled trial of rituximab for the treatment of severe cryoglobulinemic vasculitis. Arthritis Rheum 2012;64(3):843–53.
87. Iaccarino L, Bartoloni E, Carli L, et al. Efficacy and safety of off-label use of rituximab in refractory lupus: data from the Italian Multicentre Registry. Clin Exp Rheumatol 2015;33(4):449–56.
88. Isaksen K, Jonsson R, Omdal R. Anti-CD20 treatment in primary Sjögren's syndrome. Scand J Immunol 2008;68(6):554–64.
89. Reed AM, Crowson CS, Hein M, et al; RIM Study Group. Biologic predictors of clinical improvement in rituximab-treated refractory myositis. BMC Musculoskelet Disord 2015;16:257.
90. Querol L, Rojas-García R, Diaz-Manera J, et al. Rituximab in treatment-resistant CIDP with antibodies against paranodal proteins. Neurol Neuroimmunol Neuroinflamm 2015;2(5):e149.
91. Milo R. Therapeutic strategies targeting B-cells in multiple sclerosis. Autoimmun Rev 2016;15(7):714–18.
92. Wang HH, Liu CW, Li YC, et al. Efficacy of rituximab for pemphigus: a systematic review and meta-analysis of different regimens. Acta Derm Venereol 2015;95(8):928–32.
93. Kado R, Sanders G, McCune WJ. Suppression of normal immune responses after treatment with rituximab. Curr Opin Rheumatol 2016;28(3):251–8.
94. Carson KR, Focosi D, Major EO, et al. Monoclonal antibody-associated progressive multifocal leucoencephalopathy in patients treated with rituximab, natalizumab, and efalizumab: a Review from the Research on Adverse Drug Events and Reports (RADAR) Project. Lancet Oncol 2009;10(8):816–24.
95. Pawluczkowycz AW, et al. Binding of submaximal C1q promotes complement-dependent cytotoxicity (CDC) of B cells opsonized with anti-CD20 mAbs ofatumumab (OFA) or rituximab (RTX): considerably higher levels of CDC are induced by OFA than by RTX. J Immunol 2009;183:749–58.
96. Bologna L, Gotti E, Manganini M, et al. Mechanism of action of type II, glycoengineered, anti-CD20 monoclonal antibody GA101 in B-chronic lymphocytic leukemia whole blood assays in comparison with rituximab and alemtuzumab. J Immunol 2011;186(6):3762–9.
97. Keith MP, Pitchford C, Bernstein WB. Treatment of hemophagocytic lymphohistiocytosis with alemtuzumab in systemic lupus erythematosus. J Clin Rheumatol 2012;18(3):134–7.
98. Mohammad AJ, Smith RM, Chow YW, et al. Alemtuzumab as Remission Induction Therapy in Behçet Disease: A 20-year Experience. J Rheumatol 2015;42(10):1906–13.
99. Jones JL, Coles AJ. Mode of action and clinical studies with alemtuzumab. Exp Neurol 2014;262(Pt A):37–43.
100. Cabrera CM, Urra JM, Carreño A, et al. Differential expression of CD30 on CD3 T lymphocytes in patients with systemic lupus erythematosus. Scand J Immunol 2013;78(3):306–12.

MULTIPLE-CHOICE QUESTIONS

1. A patient has symmetrical inflammatory arthritis and photosensitive skin rashes. The serological profile is notable for positive rheumatoid factor (RF), positive anti–citrullinated peptide (CCP) antibody, and positive antinuclear antibody (ANA) (titer 1:640). She is diagnosed with rheumatoid arthritis (RA) and systemic lupus erythematosus (SLE) overlap. Treatment with weekly dosing with methotrexate and treatment with an antimalarial (hydroxychloroquine) fails to achieve resolution of the joint inflammation. Which of the following biological treatments would be most appropriate to treat the ongoing joint inflammation?
A. Etanercept
B. Adalimumab
C. Certolizumab
D. Rituximab
E. Belimumab

2. A patient with congestive heart failure, chronic kidney disease, and diabetes is hospitalized with manifestation of volume overload and fever. *Escherichia coli* bacteremia from urinary sepsis is confirmed. Appropriate parenteral antimicrobial therapy and diuresis are initiated, resulting in clinical improvement. But on the third hospital day, the patient develops pain, swelling, and tenderness in the left great toe, midfoot, and ankle. An ankle arthrocentesis is performed, and it demonstrates of urate crystals in the joint fluid; gram staining is negative for bacterial organisms. Given the reluctance to prescribe nonsteroidal antiinflammatory drugs (NSAIDs) or corticosteroids to manage the acute polyarticular gout because of this patient's comorbidities, which of the following biologicals would be MOST appropriate to use in this setting?
A. Canakinumab
B. Anakinra
C. Infliximab
D. Abatacept
E. Mepolizumab

3. A patient with inflammatory bowel disease (IBD) has increased weight loss in association with ongoing abdominal pain and diarrhea despite use of adalimumab. Abdominal imaging confirms evidence of significant bowel inflammation involving multiple segments of the ileum. Which of the following of the following biologicals would be MOST appropriate to prescribe?
A. Secukinumab
B. Tocilizumab
C. Etanercept
D. Natalizumab
E. Vedolizumab

4. A 23-year-old patient has a history of asthma, for which an inhaled corticosteroid preparation and a beta-agonist have been prescribed, and recurrent urticaria, for which daily doses of an antihistamine (loratadine) and a leukotriene receptor antagonist (montelukast) have been prescribed. Although occasional use of a rescue inhaler is required, the asthma is felt to be reasonably well controlled; however, there are recurring episodes of urticaria and angioedema. Laboratory studies are notable for a peripheral eosinophil count of 210/mm^3 and serum immunoglobulin E (IgE) of 110 IU/mL. Complement level C4 is 12 mg/dL (normal >14), complement C3 is normal, and CH50 is normal. Tests for antinuclear antibodies (ANAs) and antineutrophil cytoplasmic antibodies (ANCAs) are negative. Addition of which of the following would be MOST useful in managing this disorder?
A. Omalizumab
B. Rituximab
C. Mepolizumab
D. Reslizumab
E. Eculizumab

5. A patient with an undifferentiated autoimmune connective tissue disease has ongoing severe inflammatory manifestations affecting the central nervous system, lungs, joints, and kidneys despite use of standard therapy, prompting consideration of more aggressive cell-depleting immunosuppressive therapy. Use of which of the following is MOST likely to be associated with the reemergence of autoimmune manifestations following immune reconstitution?
A. Rituximab followed by belimumab
B. Obinutuzumab followed by belimumab
C. Alemtuzumab
D. Brentuximab vedotin
E. Cyclophosphamide followed by low-dose IL-2

90

Vaccines

Sarah Kabbani, Mark J. Mulligan

It might be fair to imagine that the field of vaccinology began with the first recognition that a person who survived an infection was much less likely to suffer from that disease in the future in comparison with a person who had not previously had the disease. That is, the infectious disease, if it did not kill the person, resulted in an experienced, protected state. Today we would say that the protected survivor had immunity.

Vaccines are clinically simple but immunologically complex preventive or therapeutic interventions that can dramatically reduce morbidity and mortality caused by infectious diseases in children, adolescents, adults, and seniors. In recent years, vaccinations against other categories of disease have become an increasingly active area of research (*e.g.,* cancer and dementia). In addition, vaccines against microbes are increasingly appreciated for their potential role in the medical field's heightening battle against antimicrobial-resistant bacteria. For example, the protein conjugate pneumococcal vaccine was shown to reduce the incidence of antibiotic-resistant pneumonia.[1] Preventing an illness through vaccination obviates the need to treat a bacterial infection with antibiotics, thereby avoiding potential induction of antibiotic resistance in either the targeted pathogenic bacterium or the patients' healthy microbiota.

Another important area of recent emphasis has been the double power of vaccinations for pregnant women, protecting two—mother and infant—against the targeted pathogen (e.g., pertussis[2] or influenza[3,4]). In addition, a critical need for pre-pregnancy vaccination to protect pregnant women and their fetuses ahead of and during the highly vulnerable first trimester of pregnancy is currently underscored by acquisition of Zika virus infections during pregnancy with resulting microcephaly and other birth and developmental anomalies.[5-8]

Although senior citizens (>65 years of age) experience high proportions of the total morbidity and mortality for several vaccine preventable diseases (*e.g.,* seasonal influenza, pneumococcal disease, herpes zoster) as a result of immunosenescence, they are least able to mount their own protective immune responses after vaccination. Community protection (also termed *herd immunity*) of seniors by vaccination of children, who are the primary spreaders of many vaccine-preventable infectious diseases, and younger adults can provide dramatic reductions in infectious disease incidence in seniors—for example, the reduction in pneumococcal disease in seniors after the introduction of pneumococcal conjugate vaccine for children.[9,10] Despite suboptimal vaccine responses with aging, several vaccines are specifically recommended for seniors (see Fig. 90.3) and some of the vaccines for seniors now employ novel strategies to enhance the immunosenescent response, such as high doses of influenza antigen[11,12] or coformulation of virus antigen with an adjuvant.[13,14]

That vaccines have a societal benefit in addition to the individual protection provided to the vaccine recipient is a special and gratifying aspect of this highly valuable biomedical intervention. Vulnerable community members whose immune systems are unable to respond well to vaccines (newborns, immunocompromised persons, older adults) or are unable to receive vaccines (because of allergy or a medical contraindication) depend on immunization of surrounding community members for protection against vaccine preventable diseases. The vaccinated community becomes an immunological cocoon or wall of protection for these vulnerable members of society.

This chapter first reviews selected events in the history of vaccination, both distant and recent. The remarkable accomplishments that have resulted from programs of vaccination to date are then highlighted. We also describe important recent milestones and changes in vaccine development strategies that have the potential to revolutionize the field and offer great hope for providing solutions for unmet vaccine needs. Current vaccine recommendations are then summarized. Finally several present and future challenges for the field of vaccinology are discussed.

HISTORY OF VACCINES

The earliest known vaccines for which records have been identified were against smallpox and were used in Asia early in the second millennium.[15] Of course, they were not called "vaccines." The practice was called *variolation* and involved exposing, usually through the intranasal route, a smallpox-susceptible (not previously infected) person to material from the dried scabs of a person who had had smallpox. If the recipient survived, he or she was protected against future smallpox disease. Since natural smallpox had a 30% mortality rate and variolation had a lower (≈1%) mortality rate, this ancient practice was an early example of weighing the risk-to-benefit ratio for a human health intervention. By 1700, variolation was employed in a number of societies in Africa, India, and the Ottoman empire, and it was in use in England and France in the 1700s.[16] The practice of variolation was not without risk and sometimes caused outbreaks of a mild form of the disease. One of the many who died as a result of variolation was a son of England's King George III.

In 1796, an English physician was searching for a safer alternative to variolation. This physician was to become known as the father of vaccinology—Edward Jenner. He performed a smallpox vaccination experiment on James Phipps on May 14, 1796, using as vaccine cowpox pus from the lesions on the hands of Sarah Nelmes, a milkmaid.[17] Dr. Jenner then collected lesion material from a patient with smallpox and used this as the viral challenge

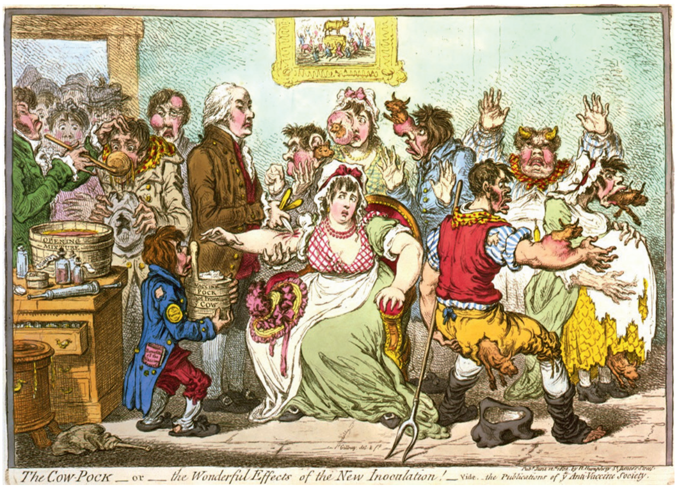

The Cow-Pock — or — the Wonderful Effects of the New Inoculation! — Vide. the Publications of ye Anti-Vaccine Society.

FIG 90.1 **Antivaccination Caricature by James Gillray From 1802 Entitled "The Cow-Pock—or—The Wonderful Effects of the New Inoculation!" Directed Against Jenner and His Work to Promote Smallpox Vaccination.** (Originally by James Gillray, The Cow-Pock—or—the Wonderful Effects of the New Inoculation! (1802). vide. the Publications of the Anti-Vaccine Society. Print (color engraving) published June 12, 1802 by H. Humphrey, St. James's Street. Library of Congress, Prints & Photographs Division.)

stock. Young Phipps survived both the vaccination and the challenge, and Jenner first reported his experiment in 1797 in a short communication to the Royal Society. After Jenner's initial report submitted for publication was rejected, he added a few more cases of his vaccine experiment (including one on his own son), and in 1798, he reported his findings in a self-published booklet entitled "An inquiry into the causes and effects of the variolae vaccinae, a disease discovered in some of the western counties of England, particularly Gloucestershire, and known by the name of the cow pox."[18]

CLINICAL RELEVANCE

Why vaccines are critical interventions for every medical practice

- Vaccines are highly effective interventions for preventing infectious diseases and have public health importance.
- Both individual protection and community (herd) immunity result from vaccination programs.
- The reductions in disease burden (morbidity and mortality) achieved through implementation of childhood vaccination programs are extraordinary.
- In the last several years, new adolescent and adult vaccines have become available and are now recommended. Vaccination is not just for children anymore.
- Clinicians of all specialties should take vaccine histories and provide access to vaccines relevant to their patients' ages and medical conditions. Access can be provided through referral or by stocking and administering the indicated vaccines.

This self-published report and Jenner's subsequent use of cowpox vaccination were greeted with mixed reaction. Jenner's work was considered controversial by some because of the introduction of a cow virus into humans. A famous satirical cartoon by James Gillray depicts cow parts emerging from vaccinated persons (Fig. 90.1). Other opponents of vaccination were those with financial interests in lucrative variolation practices. When vaccination in England was made compulsory by the Vaccination Act of 1853, an organized antivaccination movement arose almost immediately.[19] Incredibly, even in the present day, despite all the evidence supporting the safety and effectiveness of licensed vaccines, organized antivaccination movements continue to challenge contemporary clinicians and public health officials. The explosive growth of the Internet in the last 2 decades has made it relatively easy to self-publish anything and everything, and this has resulted in rapid and wide dissemination of misinformation about vaccines, antivaccination propaganda, and pseudoscience that obstructs traditional scientific evaluation and further feeds the general public's fears and misunderstandings.

It is worth noting that Jenner's work of 220 years ago encapsulates many of the elements of today's translational vaccinology.

KEY CONCEPTS

Edward Jenner's Work Relevant to Translational Vaccinology Today

- **Disease burden, surveillance, epidemiology:** A significant and unacceptable burden of smallpox disease drove development of a safer intervention to improve health.
- **Innovation:** Jenner's innovation resulted from the need for an improved biomedical intervention given the significant risk of harm associated with the centuries old variolation practice.
- **Clinical insight:** An observation that dairymaids who had recovered from an occupational illness (cowpox) were seldom affected by smallpox led to Jenner's promotion of smallpox vaccination. The observation of the protected state (immunity) in dairymaids led to a concept that was tested and promoted by Jenner.
- **Postvaccination challenge:** After the vaccination procedure, Jenner's subjects were subsequently intentionally exposed to (challenged with) wild-type smallpox and observed for safety and disease outcomes. This would not be permissible today in smallpox research, although human challenge experiments are performed for certain self-limiting or treatable infectious diseases.
- **Presentation of experimental results:** To disseminate the scientific information and advocate for wider vaccination deployment, Jenner presented his work to the Royal Society. He self-published his manuscript after it was rejected by the Society.
- **Branding:** The name "vaccination" was applied to the intervention. *Vacca* is the Latin word for "cow."
- **Antivaccination movement and conflicts of interest:** Jenner experienced significant opposition to his vaccine from groups opposed to the new technique and from individuals who had vested interests in variolation practices and stood to lose their business if the concept of vaccination caught on.

Although Jenner's smallpox challenge experiment would not be approved by today's institutional research boards (IRBs, or ethics committees), certain human challenge studies remain both acceptable and valuable today. The postvaccination challenge experiment can be an important process to efficiently obtain preliminary protective efficacy information for a vaccine and has been shown in early-phase clinical trials to be safe, well tolerated, and immunogenic. Human challenge studies are performed for self-limiting and/or treatable infections to study vaccine protection or to characterize host response to infection in detail (*e.g.*, influenza,[20] primary dengue,[21] norovirus,[22] and malaria[23]). A human challenge experiment can rapidly provide feedback to vaccine developers and public health officials to help prioritize resource-intensive field trial evaluations of promising candidate vaccines. If an encouraging preliminary efficacy signal is observed in a postvaccination human challenge trial, it can provide additional evidence and justification to funding agencies for proceeding to large and expensive field trials. This is particularly advantageous in the context of an epidemic, where rapid vaccine development and deployment are needed for optimal impact on morbidity and mortality. A contemporary example is Zika virus vaccine development—for which postvaccination human challenge experiments are under consideration.

The history of vaccination entered its second phase in the nineteenth century. The miasma theory (which postulated that infectious diseases were caused by a noxious form of bad air) was gradually replaced by the development of the germ theory (according to which infectious diseases were caused by microorganisms too small to be seen without magnification). Animal experiments and laboratory cultivation of microbes were key advances relating to the germ theory in the second half of the nineteenth century. Robert Koch (1843–1910) and the great French chemist Louis Pasteur (1822–1895) made significant contributions through their many key observations and experiments in both agricultural and human infectious diseases and vaccines. Koch's famous four postulates laid out the requirements for establishing causality of infectious diseases by microbes and proved that *Bacillus anthracis* was the cause of anthrax. This provided the first proof of a microbial etiology of a specific disease.[24] Pasteur's work was credited with saving the silkworm industry, reducing the spoilage of wine, stopping epidemics among agricultural herds, and the development of vaccines against human infectious diseases, such as rabies and anthrax. Through attenuation or inactivation of wild-type microbes, Pasteur produced vaccines that induced protection against a number of diseases. He performed a number of classic vaccination and challenge experiments in farm animals; these experiments were designed to show that in experimental challenges, his altered (attenuated or inactivated) cultures of microbes could be delivered as vaccines that offered protection to susceptible animals and herds against organisms that could be devastating, including veterinary pathogens, such as chicken cholera and anthrax, and human pathogens, such as rabies.[24]

Twenty-first century vaccinology is a multitasker's dream, encompassing multiple fields, including microbiology, immunology, medicine, epidemiology, statistics, policy, manufacturing, molecular biology, public health, and ethics. But vaccinology was not always so. Certainly in the 1700s and earlier centuries, the variolators and the great Jenner knew they had created a protected condition in their patients through their intervention only if their patient survived that intervention. But they had no specific knowledge of the killed or attenuated microbe they were administering, of the virulent microbe they were protecting against, or of the immune system changes induced in the bodies of their patients (or even of the existence of the immune system).

In the nineteenth century, Pasteur and Koch certainly knew that microbes were the cause of infectious diseases and that weakened forms of microbes (vaccines) could create a protected state after administration to livestock or humans. But they had no specific knowledge of the changes produced in the vaccine recipient's immune system.

In the early twentieth century, passive immunization was developed as a therapy for infectious diseases. Active immunization involves administering a vaccine to trigger a subsequent cascade of changes in the patient's own immune system leading to a protected state (immunity), whereas passive immunization is direct transfer of the protective immune effector molecules (antibodies) taken from an immune donor patient or animal, without administering a vaccine to the patient. Early work by Emil von Bering (1854–1917) with small-animal serum therapy experiments led to the development of human serum therapies for passive immunization. In 1900, von Bering used horse sera from immune horses to cure and prevent diphtheria caused by *Corynebacterium diphtheria*. The very first Nobel Prize in Physiology or Medicine in 1901 was awarded to von Behring "for his work on serum therapy, especially its application against diphtheria, by which he has opened a new road in the domain of medical science and thereby placed in the hands of the physician a victorious weapon against illness and deaths."[25]

During the second half of the nineteenth century, scientists were discovering the immune system's defense mechanisms. The Nobel Prize in Physiology or Medicine for 1908 was awarded to Ilya Ilyich Mechnikov (1845–1916) and Paul Ehrlich (1854–1915) "in recognition of their work on immunity" and for establishing the concepts of cell-mediated and humoral immunities. Mechnikov described the abilities of certain white blood cells (WBCs) of starfish larvae to perform phagocytosis or engulfment and destruction of harmful bacteria and other microbes. Ehrlich worked on serum therapy against diphtheria with von Bering and speculated that certain WBCs have receptors that bind toxins formed by bacteria, and when these receptors separate from the cells, they become antibodies.

In 1924, Gaston Ramon (1886–1963) developed the method of chemical inactivation of bacterial toxins with formaldehyde and heat to produce toxoids of the pathogenic toxins of diphtheria and tetanus, producing safer vaccine antigens that retained their immunogenic potential.

The Nobel Prize in Physiology or Medicine 1951 was awarded to Max Theiler (1899–1972) "for his discoveries concerning yellow fever and how to combat it." Theiler, through passage of yellow fever virus in mice, developed an attenuated live yellow fever virus variant designated 17D, which became a highly effective vaccine.

Other innovators who exploited the potential of microbial cultivation were Enders et al., who, in the 1940s, discovered methods for cultivation of poliovirus in cell culture. This removed the obstacles in the field of poliovirus vaccine development, which then advanced relatively quickly after decades of very slow progress. Laboratory growth of poliovirus permitted the development of both the inactivated polio vaccine (IPV; Salk, licensed in 1955) and the live attenuated oral polio vaccine (OPV; Sabin, monovalent licensed in 1961, trivalent in 1963).[26] As a result of these vaccines, still in use today, poliovirus type 2 was eradicated in 1999, and no wild-type poliovirus type 3 has been detected since 2012. Only poliovirus type 1 is still endemic in 2016 in just two countries, Pakistan and Afghanistan. The progress toward polio eradication is impressive, as the final steps are being taken, but final eradication will require great persistence in vaccinating the populations in these countries and sustained attention to surveillance.[27]

In their considerations of the power of vaccines and other preventive interventions to protect humans and reduce infectious disease incidences, public health officials have described the stages leading to the ultimate goal of ending human suffering caused by infectious diseases: control, elimination, eradication, and extinction of infectious diseases (Table 90.1).[28] The long road, which will hopefully culminate soon in poliovirus eradication, began with successful poliovirus cultivation in cells in the laboratory. "For their discovery of the ability of poliomyelitis viruses

TABLE 90.1 Stages of Reduction of Infectious Disease Incidence by Vaccination and Other Preventive Interventions

Control. The reduction of disease incidence and prevalence to a locally acceptable level as a result of vaccination and/or other interventions; continued interventions are needed to maintain the reduction. *Example:* diarrheal diseases.

Elimination of disease. Reduction to zero of the incidence of a specified disease in a defined geographical area as a result vaccination and/or other interventions; continued measures are required. *Example:* neonatal tetanus.

Elimination of infection. Reduction to zero of the incidence of infectious disease in a defined geographical area as a result of vaccination and/or other interventions; continued measures to prevent reestablishment of transmission are required. *Example:* poliomyelitis elimination from North America.

Eradication. Permanent reduction to zero of the worldwide incidence of infection caused by a specific agent as a result of vaccination and/or other prevention efforts; interventions are no longer needed. *Example:* smallpox.

Extinction. An infectious agent no longer exists in nature or in the laboratory. *Example:* none.

(Adapted from Dowdle WR. The principles of disease elimination and eradication. Bull W H O 1998;76 Suppl 2:22–5.)

to grow in cultures of various types of tissue," John Enders, Thomas Weller, and Frederick Robbins jointly received the Nobel Prize in Physiology or Medicine in 1954.[29]

Recognition and subsequent exploitation in vaccines of key antigenic substructures rather than whole microbes was another technical advance. For example, the studies of Oswald Avery (1877–1955), Rebecca Lancefield (1895–1981), and others of the polysaccharide capsules of *Streptococcus pneumoniae*[30,31] and the M proteins of the Streptococcus species,[32] respectively, led to the characterization, isolation, and serotyping of these bacterial structures and their recognition as key antigens in immunity to streptococcal diseases. Such observations eventually led to safer vaccination with components (subunits) of pathogens, as opposed to entire microbes (*e.g.,* the bacterial polysaccharide capsules from *S. pneumoniae* or *Neisseria meningitidis,* or the viral surface antigens [hemagglutinins, or HA] in influenza split-virus vaccines). When delivered as vaccines, these isolated microbial components produced protective antibodies and cellular immune responses in vaccinated hosts but did not cause the disease induced by the complete wild-type organisms.

The great influenza A/H1N1 pandemic of 1918 ("Spanish flu") caused an estimated 50 million deaths globally. This led to efforts to develop vaccines against influenza, and in 1933, the virus that causes influenza was identified.[33] That discovery quickly led to the first-generation live attenuated and killed monovalent influenza vaccines. A bivalent vaccine was produced in 1942 after the discovery of a second influenza type, type B. The ability to grow influenza viruses in embryonated hen's eggs facilitated vaccine production and provided the basis for killed and split-virus influenza virus vaccines that came into use in the 1940s.[33] Over the following years that followed, there was increasing awareness of the multiple types of influenza (A, B, and C), subtypes of influenza A (*e.g.,* A/H1N1, A/H3N2), and lineages of influenza B (Victoria and Yamagata). In addition, recognition that influenza virus *antigen drift* (caused by changes in the two surface proteins HA and neuraminidase [NA]) could lead to vaccine mismatch with an altered circulating strain and thus decreased vaccine effectiveness was first described when the US Army's A/H1N1 vaccine, which had been successfully used for several years, became significantly less effective in 1947 as a result of antigenic drift. In 1958, the type A virus, which was circulating at that time, entirely changed its HA and NA proteins to new subtypes, a process termed *antigenic shift.*[33] With increased recognition that it was necessary to regularly update the vaccine composition to match the changing circulating strains, in 1973, the World Health Organization (WHO) began providing annual recommendations for the composition of the seasonal influenza vaccine based on the current circulating subtypes and strains. In 1978, the first trivalent vaccine was licensed (two influenza A strains and one influenza B strain). Beginning in 2013, the WHO began including a second type B strain and recommended four vaccine strains annually (A/H3N2, A/H1N1, B/Victoria, and B/Yamagata), which enabled production and licensure of the new quadrivalent influenza vaccines.

Several other live, attenuated viral vaccines, such as measles, mumps, and rubella vaccine (MMR), were developed in the second half of the twentieth century and became staples of childhood vaccination programs both in the United States and globally. Serial passages of wild-type viruses led to viral adaptation for growth in cell cultures and a diminished ability to cause disease in humans. Importantly, the cell culture–passaged viruses that were useful as vaccines were not only well tolerated and safe in humans, but they retained the ability to produce protective immune responses. Later in the twentieth century, the development of the Oka strain of the varicella-zoster virus led to live attenuated vaccines for both chicken pox in children and herpes zoster in seniors.

KEY CONCEPTS

Historical Highlights

- Edward Jenner promotes use of the smallpox vaccine in 1798.
  - "Morbid matter of various kinds, when absorbed into the system, may produce effects in some degree similar; but what renders the cow-pox virus so extremely singular is that the person who has been thus affected is forever after secure from the infection of the small-pox; neither exposure to the variolous effluvia, nor the insertion of the matter into the skin, producing this distemper."[18]
- World Health Organization certifies eradication of smallpox from the world in 1980.
  - "The world and its peoples have won freedom from smallpox, which was a most devastating disease sweeping in epidemic form through many countries since earliest time, leaving death, blindness and disfigurement in its wake and which only a decade ago was rampant in Africa, Asia and South America."[40]

The decades following World War II have been described as a golden age for vaccinology. Maurice Hilleman (1919–1985) developed a number of vaccines against viruses while working for large pharmaceutical companies during the middle of the twentieth century. His highly productive career included inventions of vaccines against measles, mumps, and rubella; *Haemophilus influenzae* type b; hepatitis A; hepatitis B; and chickenpox.[34]

The polysaccharide vaccines developed for the prevention of bacterial diseases caused by *H. influenzae, N. meningitidis,* and *S. pneumoniae* were welcome advances. Over time it was recognized that these polysaccharide antigens were T-cell independent because they contained no peptide/protein antigens. Furthermore, unlike T cell–dependent protein antigens, these carbohydrate antigens failed to produce B-cell memory responses. The polysaccharide vaccines could not be used in neonates because infants cannot mount an effective immune response against polysaccharide antigens alone. The brilliant technical development that overcame this disadvantage for bacterial polysaccharide vaccines was covalent coupling, or conjugation, of the glycan structure to a protein carrier, such as tetanus or diphtheria toxoids. This maneuver converted the T cell–independent polysaccharide vaccines into T cell–dependent protein–polysaccharide conjugate vaccines, and this resulted in B-cell memory, much improved immunity, usefulness in newborns, and even reduction in nasopharyngeal carriage (not seen with the pure polysaccharide vaccines), thus creating herd immunity.[35] The observation that protein conjugation to polysaccharide improved the immune response was made first by Avery and Goebel in 1929[36] but was not utilized until Robbins et al. conjugated the *H. influenzae* type B polysaccharide.[37]

Increasingly in the late twentieth century and early twenty-first century, the molecular biology revolution and the fundamental dissections of the innate and adaptive immune responses are being exploited to produce new-generation vaccines. Some of these dimensions are discussed further in the sections below.

ACCOMPLISHMENTS OF VACCINATION

It is generally believed that elimination of an infectious disease from circulation in human populations through vaccination can only be achieved in the case of pathogens that have no animal reservoir and vaccination against which induces long-lasting immunity (see Table 90.1). Smallpox eradication was, in fact, achieved after a successful worldwide vaccination campaign and is the signature accomplishment of vaccination. The fields of medicine and public health celebrate this remarkable vaccination success as an example of the power of vaccination to improve human health. Smallpox was an infection that was a scourge of humanity for millennia, disfiguring and blinding survivors and killing 30% of those infected.[38] The last reported case of smallpox in the United States was in 1949. The world's last known patient with naturally occurring smallpox was Ali Maow Maalin in Somalia in 1977.[39] After the disease was eliminated, routine vaccination of the general public against smallpox was discontinued because it was no longer necessary for prevention of this disease. In 1972, the United States halted routine smallpox immunization of the American public. In 1980, the WHO certified that smallpox had been eradicated: "The world and its peoples have won freedom from smallpox, which was a most devastating disease sweeping in epidemic form through many countries since earliest time, leaving death, blindness and disfigurement in its wake and which only a decade ago was rampant in Africa, Asia and South America."[40] Small quantities of smallpox virus are stored in secure research laboratories at the Centers for Disease Control and Prevention (CDC, Atlanta, GA) and in Russia.

It has been over 200 years since Jenner's self-published report on smallpox vaccination.[17] As the dawn of the third millennium approached, the CDC designated vaccination as first on the list of the 10 greatest public health achievements of the twentieth century.[41] In addition to smallpox eradication, the control of many common childhood infections and attendant reductions in morbidity and mortality ranked as very high achievements. Implementation of routine US childhood immunization programs led to major reductions from mid-twentieth century disease peaks to the record low levels of several infectious diseases today (Table 90.2). For example, in the United States, the number of polio, measles, rubella, and mumps cases declined by 100%, 99.9%, 99.9%, and 95.9%, respectively.[42] It is estimated that for each annual birth cohort of approximately four million children in the United States, vaccines in the childhood immunization schedule prevent an estimated 20 million cases of disease and 42 000 deaths.[43] Furthermore, although it is true that a considerable investment of resources is required to complete the annual programs of childhood vaccination, vaccines do result in very significant cost savings, and hence are highly cost-effective interventions. For each yearly US birth cohort, vaccines result in nearly $14 billion in annual net direct cost savings and $69 billion in annual net societal cost savings (which include savings such as reductions in costs of missed work by parents caring for an ill child).[43]

TABLE 90.2 **Historical Comparisons of Morbidity and Mortality for Vaccine-Preventable Diseases in the United States**

Disease	Before Vaccination: Estimated Annual Average Number of Cases	After Vaccination: Annual Cases (Reported or Estimated) in Year 2006	% Reduction
Diphtheria	21053	0	100
Measles	530217	55	99.9
Mumps	162344	6584	95.9
Pertussis	200752	15632	92.2
Paralytic poliomyelitis	16316	0	100
Rubella	47745	11	99.9
Smallpox	29005	0	100
Tetanus	580	41	92.9
Hepatitis A	117333	15296	87
Acute hepatitis B	66232	13169	80.1
Invasive Hib	20000	<50	99.8
Invasive pneumococcal disease	63067	41550	34.1
Varicella	4085120	48445	85

Hib, *Haemplhilus influenzae* type b.

(Adapted from Roush SW, Murphy TV, Vaccine-Preventable Disease Table Working Group. Historical comparisons of morbidity and mortality for vaccine-preventable diseases in the United States. JAMA 2007;298:2155–63.)

Saving lives and reducing morbidity and human suffering is the major accomplishment of vaccination. Although there are many human infectious diseases yet to be controlled by vaccination (some are discussed below), the accomplishments to date are truly spectacular. For example, the CDC estimates that US childhood vaccination programs will prevent 21 million hospitalizations and 732 000 deaths among children born in the last 20 years.[44]

Vaccines are not given solely to protect individuals against diseases. Another purpose of vaccination is to protect communities by reducing transmission of disease-causing microbes from vaccinated persons to unvaccinated persons. The term for this protection is *herd immunity* or *community immunity.*[45] A disease that has been studied closely with regard to community immunity is measles. Measles is very highly contagious and its epidemic form is easily recognizable. Clustering of poor vaccination coverage often occurs in particular communities, and this was highlighted in recent measles outbreaks in the United States.[46]

In a measles outbreak in California in 2014–2015, of the 110 patients, 49 were unvaccinated; in that subgroup, 24% were children too young to be vaccinated, and 67% were intentionally unvaccinated (mostly children) as a result of parental beliefs. Such outbreaks point out the importance of community immunity to protect the vulnerable (unvaccinated) members of our communities. Given that many of the recent measles outbreaks in the United States have been linked to imported cases, another important lesson is that as long as a vaccine-preventable, highly transmissible infectious disease exists anywhere in the world, it remains a potential threat everywhere—and thus vaccination programs will continue to be important to ensure the health of all community members in any part of the world.

Another powerful example of vaccine-induced community immunity comes from pneumococcal vaccines. There are many specific challenges associated with pneumococcal vaccines: the large number of circulating serotypes, the suboptimal immunogenicity of polysaccharide only vaccines, and noninvasive carriage of the organism—all of which lead to significant problems in establishing community immunity. However, in spite of those challenges, introduction of the pneumococcal conjugate vaccines in infants in 2000 not only led to reduction in invasive diseases among the vaccinated population of children but also produced a significant reduction in adults, particularly among seniors, in whom this bacterium frequently causes pneumonia.[1] This result highlights the effectiveness of community immunity produced by vaccines.

Other recently introduced vaccines have made a significant impact in relatively brief periods. The 2006 implementation of routine rotavirus vaccination prevents an estimated 40 000 to 60 000 rotavirus hospitalizations in the United States annually.[47] The human papillomavirus (HPV) vaccine is a recombinant virus-like particle (VLP) vaccine for primary prevention of cancer. The CDC Advisory Committee on Immunization Practices (ACIP) recommended routine HPV vaccination for girls in 2006 and for boys in 2011. Within 6 years of the introduction of vaccine for girls, there was a 64% decrease in the prevalence of the four types of HPV contained in the vaccine among females 14–19 years of age and a 34% decrease among those 20–24 years of age.[48] This example shows the power of an excellent new-generation viral vaccine, even when uptake is incomplete. At the end of the first decade of the current century, cases of hepatitis A, hepatitis B, and varicella have been reported to be at record low levels[49,50] (see Table 90.2).

RECENT CHANGES IN VACCINE DEVELOPMENT STRATEGIES

Vaccines were originally developed with no knowledge of the immunological correlates of protection or, indeed, of the existence of the immune system. Similarly, the earliest vaccines were developed without an in-depth understanding of microbial antigens and epitopes, which are critical to the production of protection against disease. Rather than having a theoretical or purely logical basis, for a century and a half after Jenner, vaccination continued to be based on experience and clinical observation. The observations and experiences led to a concept that was evaluated by a trial-and-error approach. The desired outcomes of the trial-and-error, or empirical, approach to vaccine development were protection against disease and survival after exposure to the pathogen. Vaccines were initially attenuated or killed versions of the whole wild-type human pathogens (*e.g.,* rabies, yellow fever, and influenza viruses). In a few cases, attenuated zoonotic organisms closely related to human pathogens were employed to produce cross-reactive protective responses in humans (*e.g.,* vaccinia, an animal poxvirus utilized as a vaccine against human smallpox, and Bacille Calmette-Guérin [BCG], an agent of bovine tuberculosis [TB] developed as a human TB vaccine).

Later, split-virus vaccines utilized partially purified protein antigens derived from whole killed viruses (*e.g.,* split-virus influenza vaccines). The polysaccharide capsules of bacteria were purified from cultures of serologically distinguishable strains, or serotypes, within single bacterial species leading to polyvalent polysaccharide vaccines (*e.g.,* the 23-valent pneumococcal polysaccharide and the quadrivalent meningococcal polysaccharide vaccine). Bacterial toxins were purified from cultures and made harmless by application of heat or chemical treatment to produce toxoid vaccines (*e.g.,* tetanus and diphtheria vaccines).

During the closing years of the twentieth century and in the early twenty-first century, the advances in genetics, molecular biology, immunology, and microbiology have led to new theory-based (so-called rational) approaches to vaccine development. Perhaps "informed empiricism" is an apt description of much of early twenty-first century vaccine development: experimental approaches that are heavily influenced by our now-imperfect, always evolving knowledge of innate and adaptive immune responses and of microbial antigens but are still dependent on an iterative system of trial and error to discover safe, well-tolerated, and efficacious vaccines. Vaccine developers today also benefit from the cumulative body of knowledge and experience in vaccinology that has accumulated over the past century.

There have been numerous changes in the design of vaccine immunogens. The formation of links between bacterial polysaccharides and protein carriers (*conjugation*) led to dramatic improvements in protection against certain bacterial diseases. Conjugation of *H. influenzae, N. meningitidis,* and, most recently, *S. pneumoniae* polysaccharides to proteins improved these vaccines by converting them from T cell–independent antigens to T cell–dependent antigens.

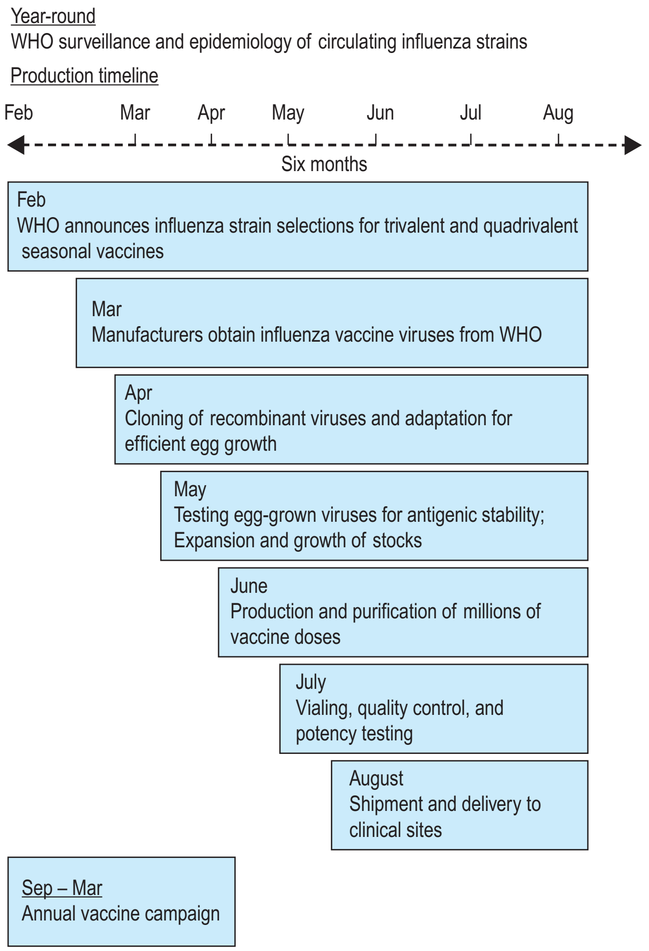
Adjuvants are materials added to vaccine antigens to enhance the recipient's immune response. Recent advances in our understanding of the innate immune system have led to an appreciation that adjuvants act through their effects on innate immunity. Adjuvant-triggered innate signals enhance the quantity and quality of the downstream adaptive immune responses to the vaccine antigen. Today's vaccinologists are developing vaccines with specific molecular adjuvants (*e.g.,* Toll-like receptor-5 [TLR-5], CD-40L, and interleukin-12 [IL-12]), intended to shape the immune response to provide a presumed best fit for protective immunity against the targeted pathogen.

The advent of molecular biology in the mid-twentieth century resulted in many new avenues for vaccine development. Advances in molecular biology allowed for cloning of microbes' genes and their expression in recombinant molecular systems, and vaccines could consequently be designed on the basis of the *in vitro* expression of one or a few genes. For example, the hepatitis B vaccine, originally developed by Hilleman, was purified hepatitis B surface antigen (HBsAg) from the blood of humans with chronic infection. Soon thereafter, a second licensed hepatitis B vaccine was produced in yeast cells through recombinant DNA methods that inserted the HBsAg gene into yeast organisms for expression and purification. The recombinant hepatitis B vaccine is in use in the United States and offers advantages, such as protein purity (because the genes-of-interest are expressed in relative isolation) and vaccine safety (because it is no longer necessary to derive vaccines by partially purifying the HBsAg from paid-donor plasma of humans chronically infected with hepatitis B virus and potentially other viruses).

Another example of a highly effective recombinant vaccine is the HPV vaccine. Recombinant HPV L proteins expressed in recombinant systems form VLPs that are purified and formulated with or without an adjuvant. The most recent polyvalent vaccine expresses VLPs representing nine HPV serotypes. HPV vaccines are truly remarkable for their efficacy and safety and because they offer primary prevention against several types of cancer in both females and males.

It is notable that several countries have recently licensed the first dengue vaccine.[51] It is a recombinant live viral vector that is based on the yellow fever virus vaccine strain 17D but in which the E and preM genes of 17D have been replaced by the E and preM genes of dengue serotypes 1, 2, 3, and 4. Four separate recombinant vectors were developed, one for each of the four dengue serotypes. The fact that the dengue virus and yellow fever virus are both flaviviruses with similar genetic organization facilitated development of this chimeric yellow fever–dengue vaccine. The final formulation is a tetravalent mixture of recombinant viruses representing all four dengue serotypes.

Over the last decade, systems biology, or systems vaccinology, approaches to vaccine development have captured considerable interest.[52-55] The use of new high-throughput assays to assess multiple dimensions of innate and adaptive immune responses has generated very large data sets. In addition to the traditional cellular and humoral immunological assays (antibodies, T cells, B cells), multiple "-omics" assays, including transcriptomics, proteomics, metabolomics, lipidomics, and glycomics, among others, may be performed. These detailed assessments are being applied in a variety of infectious and noninfectious disease states. One advantage for vaccine systems biology studies that enroll generally healthy study participants is the ability to obtain one or more baseline, preperturbation (prevaccination) sets of samples for analysis. Postvaccination changes can therefore be compared with the undisturbed condition. Analyses and integration of the huge amounts of collected data require multidisciplinary collaboration with computational biologists and informaticians. One application of the systems vaccine approach is to dissect the "-omics" responses produced by protective vaccines to identify associated changes at the RNA, protein, metabolite, lipid, and/or glycans levels. The promise, and some might say "hype," of systems vaccinology is both exciting and unfulfilled—and will be realized once actual improvements in human health have been achieved as a result of the systems approach.



WHO surveillance and epidemiology of circulating influenza strains

Feb
WHO announces influenza strain selections for trivalent and quadrivalent seasonal vaccines

Mar
Manufacturers obtain influenza vaccine viruses from WHO

Apr
Cloning of recombinant viruses and adaptation for efficient egg growth

May
Testing egg-grown viruses for antigenic stability; Expansion and growth of stocks

June
Production and purification of millions of vaccine doses

July
Vialing, quality control, and potency testing

August
Shipment and delivery to clinical sites

Sep – Mar
Annual vaccine campaign

FIG 90.2 **An Approximation of the Current Annual Influenza Vaccine Production Process for the Northern Hemisphere.** This is the process for egg-based vaccine production. A few manufacturers' recombinant vaccine techniques have recently eased the tight annual timeline, but most vaccines remain egg based. Development of a universal vaccine would obviate this challenging and time-constrained annual process.

CURRENT RECOMMENDATIONS

Currently in the United States, clear guidelines recommend vaccines for children, adolescents, and adults. Each February, the CDC publishes two immunization schedules[56] based on the recommendations of the CDC's Advisory Committee on Immunization Practices (ACIP). One ACIP schedule of immunizations provides the adult immunization recommendations[57] (Fig. 90.3). The adult schedule has recommendations for each vaccine based on the age of the patient. For example, the ACIP recommends that all adults (persons ≥19 years of age) receive annual influenza vaccination; a tetanus–diphtheria–acellular pertussis (Tdap) vaccine once, followed by tetanus boosters every 10 years; zoster vaccination at age 60 years; and pneumococcal vaccination at age 65 years. It should be noted that although the US Food and Drug Administration (FDA) has licensed a zoster vaccine for use at age 50 years, the ACIP recommendations do not always align completely with FDA indications (hence the ACIP recommendation for the zoster vaccine at age 60 years).

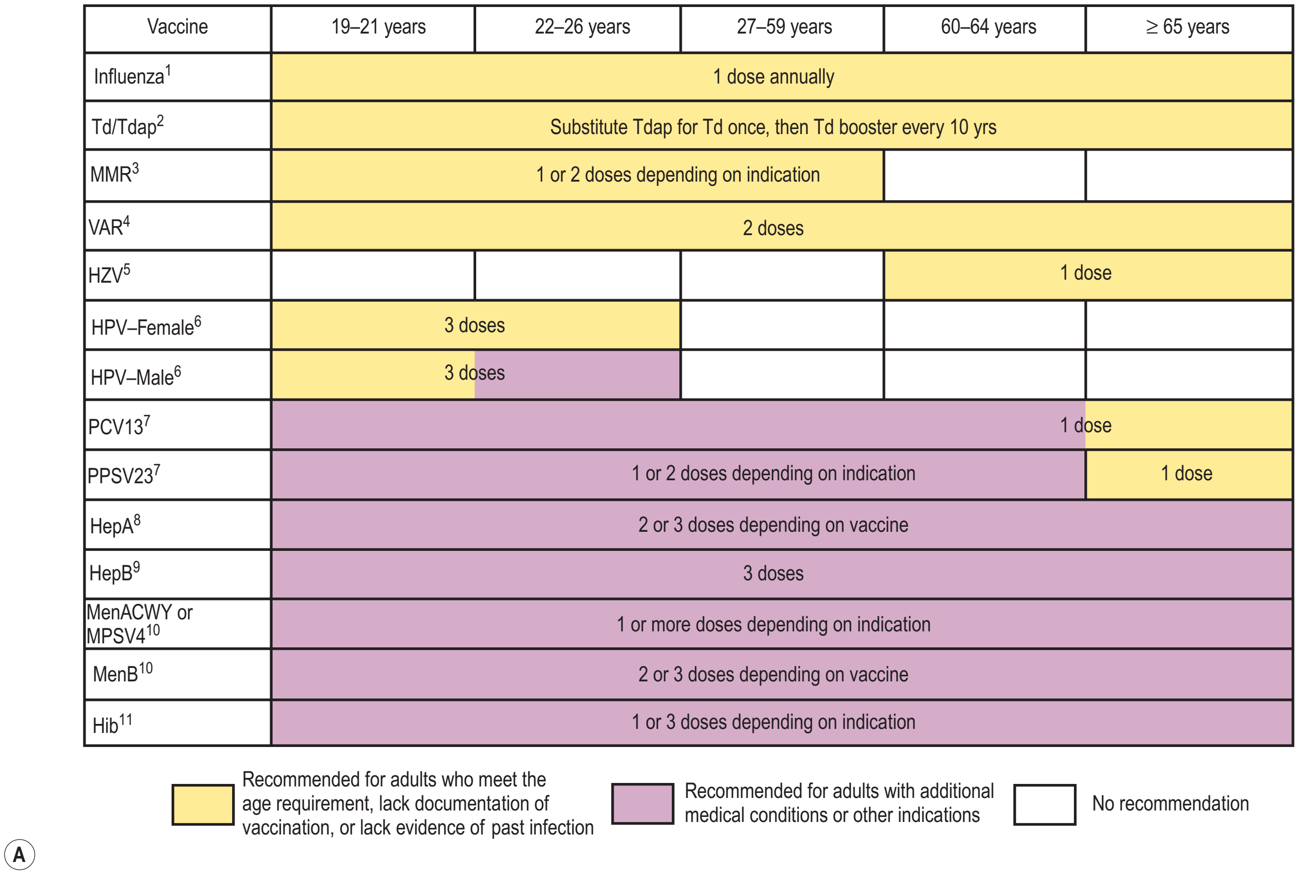
The adult schedule also provides recommendations for vaccines indicated for certain risk factors, including medical conditions (*e.g.*, immunocompromising conditions, kidney failure, diabetes) or life style, occupational, or other conditions (*e.g.*, pregnancy, men who have sex with men, health care personnel). Live vaccines (varicella, zoster, and MMR) are contraindicated in pregnant women, immunocompromised hosts, and those with human immunodeficiency virus (HIV) infection when the CD4 T-cell absolute count is below 200 cells/μL.

The second ACIP immunization schedule of immunizations covers the period from birth to 18 years and includes recommendations for children or adolescents who have not received recommended vaccines[58] (Fig. 90.4). The ministries of health in many European countries publish their own country-specific immunization schedules, and vaccination guidelines that are published by the WHO are utilized by many developing countries. The schedules are generally similar but with some region-specific differences. For example, the 2016 US ACIP immunization schedule for children recommends vaccinations against 10 viral diseases: hepatitis B, rotavirus infection, polio, influenza, measles, mumps, rubella, varicella, hepatitis A, and HPV infection.[58] Preventive viral vaccines in the WHO-recommended routine immunization schedule for children include the same 10 viral vaccines (although four of those, mumps, influenza, varicella, and hepatitis A vaccines, are recommended only for country-specific immunization programs with certain characteristics). The WHO schedule also recommends some additional vaccines (*e.g.*, rabies, yellow fever, Japanese encephalitis, and tick-borne encephalitis vaccines are recommended for certain high-risk populations).[59]

SOME PRESENT AND FUTURE CHALLENGES

The science of vaccinology is exciting and strong, having entered a new era with the high-throughput analytical and data approaches being applied today. Whether these advances in technology, computation, and science will result in improvements in human health remains to be seen. At present there is some (healthy) skepticism about the big data systems biology approach. The field must show that the massive data can be analyzed and integrated and that the approach is more than a fishing expedition but, rather, is an exploratory engine that is hypothesis generating and also leads to models that will be tested and will produce new knowledge, resulting in improved vaccines that benefit human health.

For outstanding scientists to be attracted to and retained in the field of vaccinology, it is essential that science funding agencies provide and maintain robust support for critical discovery research projects (investigator-initiated research projects) that form the engine of innovation that drives all of science. At the same time, targeted big science vaccine program projects and networks have the potential to synergistically pool their approaches in an intense focused effort to tackle major vaccine needs and challenges (*e.g.*, development of HIV and TB vaccines). To translate basic advances and laboratory science into improved health care for patients, having well-trained translational physician-scientists lead clinical–translational human research programs is an additional component of essential infrastructure. There is a need for postdoctoral training programs that focus on vaccinology, but



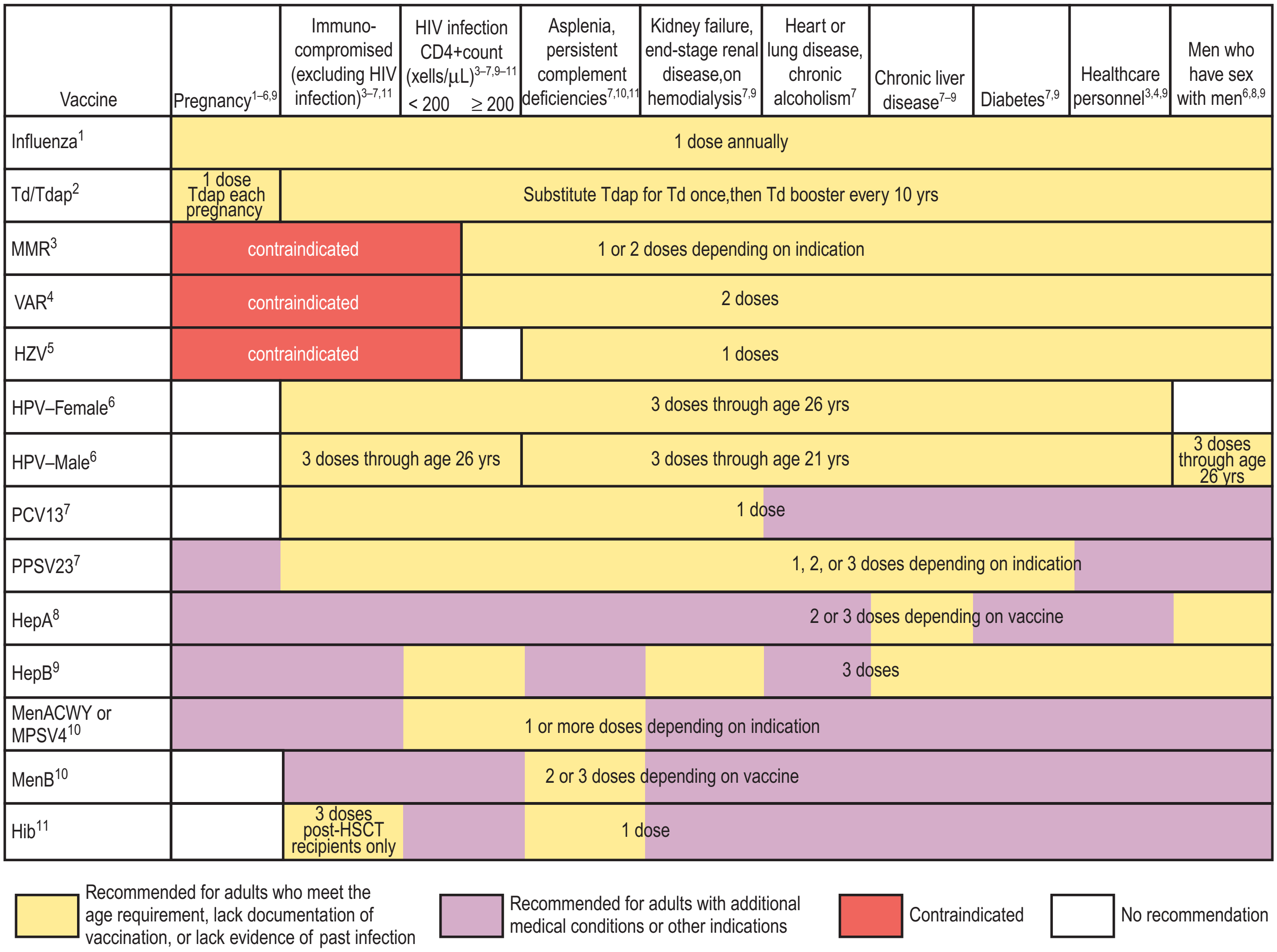
Vaccine	19–21 years	22–26 years	27–59 years	60–64 years	≥ 65 years
Influenza[1]	🟨 1 dose annually	🟨	🟨	🟨	🟨
Td/Tdap[2]	🟨 Substitute Tdap for Td once, then Td booster every 10 yrs	🟨	🟨	🟨	🟨
MMR[3]	🟨 1 or 2 doses depending on indication	🟨	🟨	⬜	⬜
VAR[4]	🟨 2 doses	🟨	🟨	🟨	🟨
HZV[5]	⬜	⬜	⬜	🟨 1 dose	🟨
HPV–Female[6]	🟨 3 doses	🟨	⬜	⬜	⬜
HPV–Male[6]	🟨 3 doses	🟪	⬜	⬜	⬜
PCV13[7]	🟪 1 dose	🟪	🟪	🟪	🟨
PPSV23[7]	🟪 1 or 2 doses depending on indication	🟪	🟪	🟪	🟨 1 dose
HepA[8]	🟪 2 or 3 doses depending on vaccine	🟪	🟪	🟪	🟪
HepB[9]	🟪 3 doses	🟪	🟪	🟪	🟪
MenACWY or MPSV4[10]	🟪 1 or more doses depending on indication	🟪	🟪	🟪	🟪
MenB[10]	🟪 2 or 3 doses depending on vaccine	🟪	🟪	🟪	🟪
Hib[11]	🟪 1 or 3 doses depending on indication	🟪	🟪	🟪	🟪

🟨 Recommended for adults who meet the age requirement, lack documentation of vaccination, or lack evidence of past infection

🟪 Recommended for adults with additional medical conditions or other indications

⬜ No recommendation

FIG 90.3 (A) Recommended Immunization Schedule for Adults Aged 19 Years or Older by Age Group. The US Centers for Disease Control and Prevention (CDC) Recommended Immunization Schedule for Adults Aged 19 Years and Older, United States, 2017, Became Effective in February, 2017 as Recommended by the Advisory Committee on Immunization Practices (ACIP) of the CDC. When Used in Immunization Practice, the Extensive Footnotes Provided With the Schedules (and Referenced in the Figure) Should be Consulted at www.cdc.gov/ vaccines/schedules/hcp/ index.html.



Vaccine	Pregnancy[1–6,9]	Immuno-compromised (excluding HIV infection)[3–7,11]	HIV infection CD4+count (xells/μL)[3–7,9–11] < 200	HIV infection CD4+count (xells/μL)[3–7,9–11] ≥ 200	Asplenia, persistent complement deficiencies[7,10,11]	Kidney failure, end-stage renal disease,on hemodialysis[7,9]	Heart or lung disease, chronic alcoholism[7]	Chronic liver disease[7–9]	Diabetes[7,9]	Healthcare personnel[3,4,9]	Men who have sex with men[6,8,9]
Influenza[1]	🟨 1 dose annually	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨
Td/Tdap[2]	🟨 1 dose Tdap each pregnancy	🟨 Substitute Tdap for Td once,then Td booster every 10 yrs	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨
MMR[3]	🟥 contraindicated	🟥	🟥	🟨 1 or 2 doses depending on indication	🟨	🟨	🟨	🟨	🟨	🟨	🟨
VAR[4]	🟥 contraindicated	🟥	🟥	🟨 2 doses	🟨	🟨	🟨	🟨	🟨	🟨	🟨
HZV[5]	🟥 contraindicated	🟥	🟥	⬜	🟨 1 doses	🟨	🟨	🟨	🟨	🟨	🟨
HPV–Female[6]	⬜	🟨 3 doses through age 26 yrs	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	⬜
HPV–Male[6]	⬜	🟨 3 doses through age 26 yrs	🟨	🟨	🟨 3 doses through age 21 yrs	🟨	🟨	🟨	🟨	🟨	🟨 3 doses through age 26 yrs
PCV13[7]	⬜	🟨 1 dose	🟨	🟨	🟨	🟨	🟪	🟪	🟪	🟪	🟪
PPSV23[7]	🟪	🟨 1, 2, or 3 doses depending on indication	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟪	🟪
HepA[8]	🟪 2 or 3 doses depending on vaccine	🟪	🟪	🟪	🟪	🟪	🟪	🟨	🟪	🟪	🟨
HepB[9]	🟪 3 doses	🟪	🟨	🟨	🟪	🟨	🟪	🟨	🟨	🟨	🟨
MenACWY or MPSV4[10]	🟪 1 or more doses depending on indication	🟪	🟨	🟨	🟨	🟪	🟪	🟪	🟪	🟪	🟪
MenB[10]	⬜	🟪 2 or 3 doses depending on vaccine	🟪	🟪	🟨	🟪	🟪	🟪	🟪	🟪	🟪
Hib[11]	⬜	🟨 3 doses post-HSCT recipients only	🟪	🟪	🟨 1 dose	🟪	🟪	🟪	🟪	🟪	🟪

🟨 Recommended for adults who meet the age requirement, lack documentation of vaccination, or lack evidence of past infection

🟪 Recommended for adults with additional medical conditions or other indications

🟥 Contraindicated

⬜ No recommendation

Ⓑ

FIG 90.3, cont'd (B) Recommended immunization schedule for adults aged 19 or older by medical condition and other indications. The following acronyms are used for vaccines recommended for adults: HepA, hepatitis A vaccine; HepAHepB, hepatitis A and hepatitis B vaccines; HepB, hepatitis B vaccine; Hib, *Haemophilus influenzae* type b conjugate vaccine; HPV vaccine, human papillomavirus vaccine; HZV, herpes zoster vaccine; IIV, inactivated influenza vaccine; LAIV, live attenuated influenza vaccine; MenACWY, serogroups A, C, W, and Y meningococcal conjugated vaccine; MenB, serogroup B meningococcal vaccine; MMR, measles, mumps, and rubella vaccine; MPSV4, serogroups A, C, W, and Y meningococcal polysaccharide vaccine; PCV13, 13-valent pneumococcal conjugate vaccine; PPSV23, 23-valent pneumococcal polysaccharide vaccine; RIV, recombinant influenza vaccine; Td, tetanus and diphtheria toxoids; Tdap, tetanus toxoid, reduced diphtheria toxoid, and acellular pertussis vaccine; VAR, varicella vaccine. (Courtesy CDC; https://www.cdc.gov.)

only a few are in existence. The future of the field depends on attracting and training a next generation of highly qualified vaccinology research leaders.

Patients' and parents' confidence in and uptake of vaccines derives most of all from the advice and education provided by their trusted physicians and other clinicians. But because of the high time pressures of office or hospital encounters, clinicians often find it challenging to prioritize patient education about vaccination. Nonetheless the clinician–patient encounter is the most potent opportunity for education and influence, and this opportunity must be seized for vaccine programs to be successful. Every physician/clinician encounter is an opportunity to review vaccination history, recommend and discuss needed vaccines, and bring our patients' vaccinations up to date.

Training of health care professionals and education of the public are essential dimensions for vaccination to achieve its potential as a driver of improved health. Clearly a vaccine that stays in its vial has no use at all. Health care providers must be inspired and trained to serve as advocates and champions of vaccination; to remember to discuss vaccination at every patient encounter; to stock vaccines in their refrigerators and freezers; and to use electronic reminder systems to ensure compliance with recommended vaccinations based on the patient's age and medical or other indications.

Public awareness and education about vaccination must be enhanced through professional marketing campaigns. The negative consequences of propaganda by antivaccination groups, alternative or delayed schedules of vaccination, and parental hesitancy regarding vaccination can all result in resurgence of infectious diseases, increased morbidity, and increased vaccine-preventable deaths.

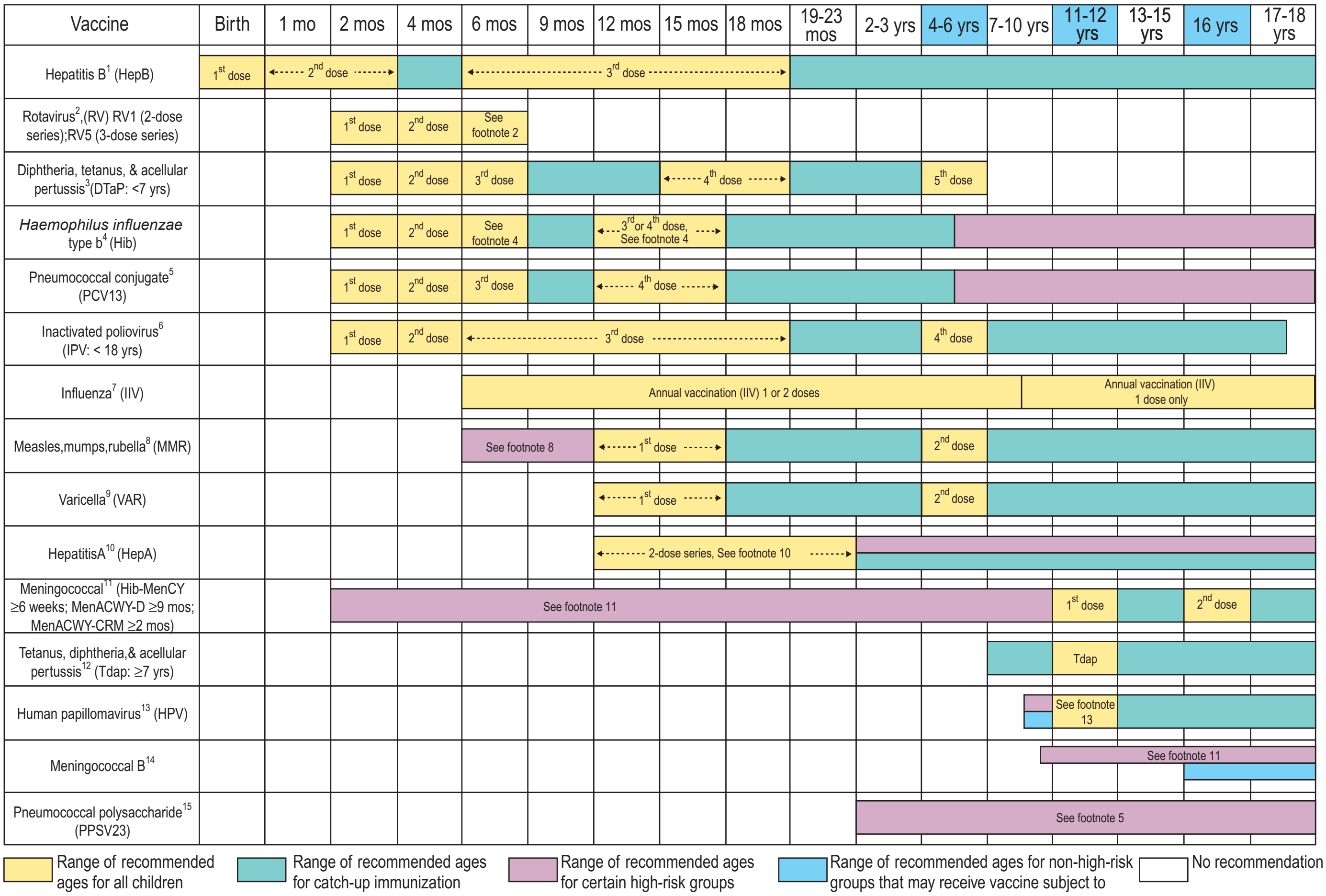
With modern air travel, which makes it possible to go from one continent to another in a matter of a few hours, it is no stretch to say we dwell in a global village, where maintaining good health and avoiding highly contagious infectious diseases depend on the health of communities, whether local or global. Therefore it is in our self-interest to maintain high levels of vaccination in our home communities and to advocate and support vaccination efforts around the globe. Clearly, vaccination is both a personal and a social, or community, endeavor.

A few specific challenges for vaccination are highlighted below to illustrate the general points made above.

A Vaccine Against Human Immunodeficiency Virus

The development of a vaccine against human immunodeficiency virus/acquired immunodeficiency syndrome (HIV/AIDS) has long been recognized as a top HIV research global priority at the US National Institutes of Health (NIH). Strong and simple treatments for those who are living with HIV infection are now in existence in the United States and have been made available even in developing countries. It has been shown that the treatment-as-prevention approach, wherein the viral load is lowered to an undetectable level by antiretroviral treatment of infected persons, results not only in benefit to the infected patient but also in up to 96% reduction in HIV incidence among sexual partners.[60,61]

More recently, antiretroviral drugs (ARDs) have been tested globally and licensed in the United States as a once-daily pill (a combination of tenofovir and emtricitabine) for HIV/AIDS prevention in higher-risk individuals.[62] Known as *preexposure*



Vaccine	Birth	1 mo	2 mos	4 mos	6 mos	9 mos	12 mos	15 mos	18 mos	19-23 mos	2-3 yrs	4-6 yrs	7-10 yrs	11-12 yrs	13-15 yrs	16 yrs	17-18 yrs
Hepatitis B[1] (HepB)	🟨 1st dose	🟨 ←2nd dose	🟨 2nd dose→	🟩	🟨 ←3rd dose	🟨	🟨	🟨	🟨 3rd dose→	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩
Rotavirus[2],(RV) RV1 (2-dose series);RV5 (3-dose series)			🟨 1st dose	🟨 2nd dose	🟨 See footnote 2												
Diphtheria, tetanus, & acellular pertussis[3](DTaP: <7 yrs)			🟨 1st dose	🟨 2nd dose	🟨 3rd dose	🟩	🟩	🟨 ←4th dose	🟨 4th dose→	🟩	🟩	🟨 5th dose					
Haemophilus influenzae type b[4] (Hib)			🟨 1st dose	🟨 2nd dose	🟨 See footnote 4	🟩	🟨 ←3rd or 4th dose, See footnote 4	🟨 3rd or 4th dose, See footnote 4→	🟩	🟩	🟩	🟩/🟪	🟪	🟪	🟪	🟪	🟪
Pneumococcal conjugate[5] (PCV13)			🟨 1st dose	🟨 2nd dose	🟨 3rd dose	🟩	🟨 ←4th dose	🟨 4th dose→	🟩	🟩	🟩	🟩/🟪	🟪	🟪	🟪	🟪	🟪
Inactivated poliovirus[6] (IPV: < 18 yrs)			🟨 1st dose	🟨 2nd dose	🟨 ←3rd dose	🟨	🟨	🟨	🟨 3rd dose→	🟩	🟩	🟨 4th dose	🟩	🟩	🟩	🟩	🟩
Influenza[7] (IIV)					🟨 Annual vaccination (IIV) 1 or 2 doses	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨 / Annual vaccination (IIV) 1 dose only	🟨	🟨	🟨	🟨
Measles,mumps,rubella[8] (MMR)					🟪 See footnote 8	🟪	🟨 ←1st dose	🟨 1st dose→	🟩	🟩	🟩	🟨 2nd dose	🟩	🟩	🟩	🟩	🟩
Varicella[9] (VAR)							🟨 ←1st dose	🟨 1st dose→	🟩	🟩	🟩	🟨 2nd dose	🟩	🟩	🟩	🟩	🟩
HepatitisA[10] (HepA)							🟨 ←2-dose series, See footnote 10	🟨	🟨	🟨 2-dose series, See footnote 10→	🟪/🟩	🟪/🟩	🟪/🟩	🟪/🟩	🟪/🟩	🟪/🟩	🟪/🟩
Meningococcal[11] (Hib-MenCY ≥6 weeks; MenACWY-D ≥9 mos; MenACWY-CRM ≥2 mos)			🟪 See footnote 11	🟪	🟪	🟪	🟪	🟪	🟪	🟪	🟪	🟪	🟪	🟨 1st dose	🟩	🟨 2nd dose	🟩
Tetanus, diphtheria,& acellular pertussis[12] (Tdap: ≥7 yrs)													🟩	🟨 Tdap	🟩	🟩	🟩
Human papillomavirus[13] (HPV)													🟪/🟦	🟨 See footnote 13	🟩	🟩	🟩
Meningococcal B[14]													🟪 See footnote 11	🟪	🟪	🟪/🟦	🟪/🟦
Pneumococcal polysaccharide[15] (PPSV23)											🟪 See footnote 5	🟪	🟪	🟪	🟪	🟪	🟪

🟨 Range of recommended ages for all children

🟩 Range of recommended ages for catch-up immunization

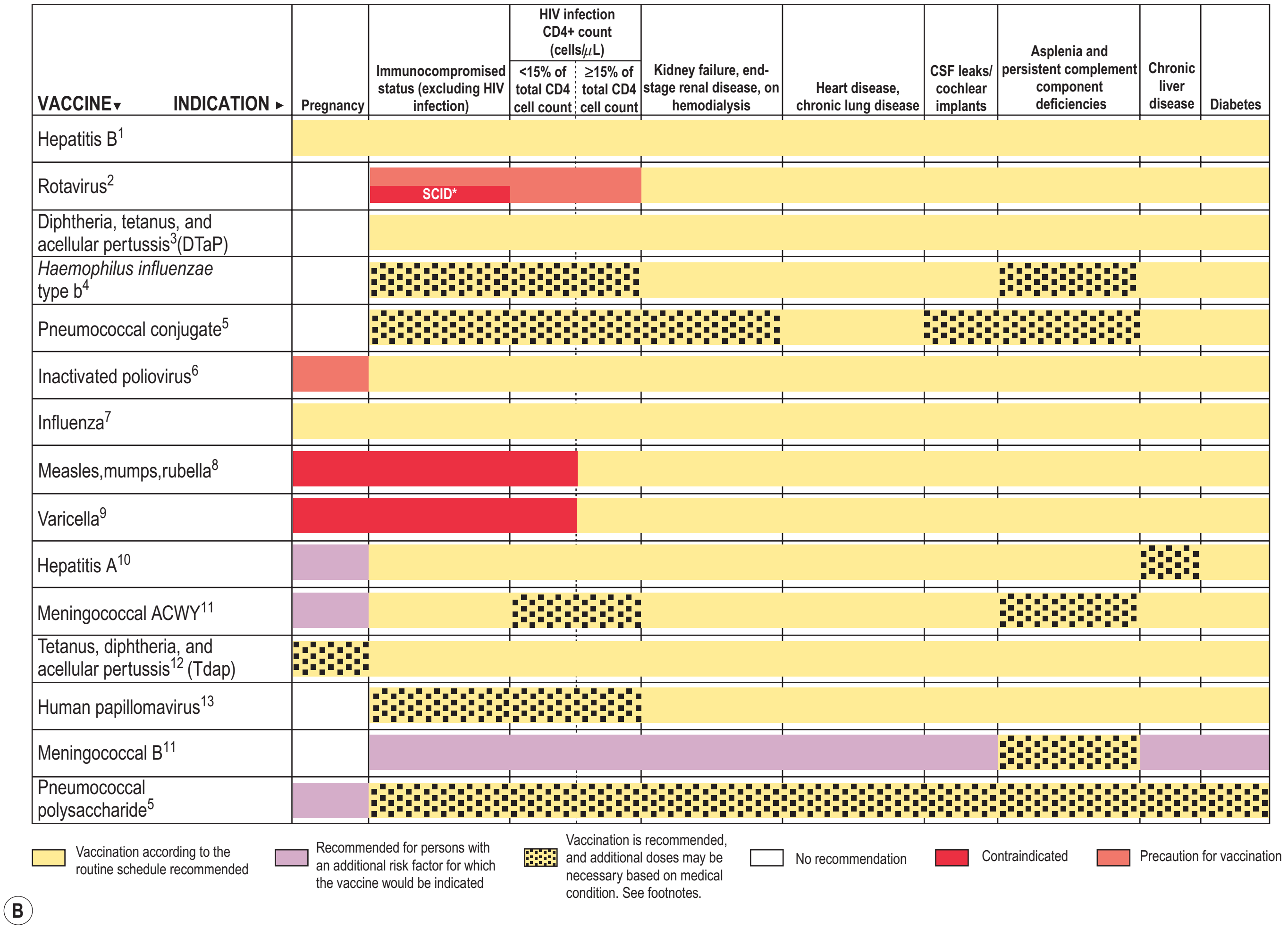
🟪 Range of recommended ages for certain high-risk groups

🟦 Range of recommended ages for non-high-risk groups that may receive vaccine subject to

☐ No recommendation

(A)

FIG 90.4 **CDC Recommended Immunization Schedule for Children and Adolescents Aged 18 Years or Younger, United States, 2017 Became Effective in January, 2017, as Recommended by the CDC's AICP.** When used in immunization practice, the extensive footnotes provided with the schedules (and referenced in the figure) should be consulted at www.cdc.gov/vaccines/schedules/hcp/index.html, which also provides a schedule for persons aged 4 months through 18 years who start late or who are more than 1 month behind. (A) Recommended immunization schedule for children and adolescents aged 18 years or younger by age.



VACCINE ▾ INDICATION ▸	Pregnancy	Immunocompromised status (excluding HIV infection)	HIV infection CD4+ count (cells/μL) <15% of total CD4 cell count	HIV infection CD4+ count (cells/μL) ≥15% of total CD4 cell count	Kidney failure, end-stage renal disease, on hemodialysis	Heart disease, chronic lung disease	CSF leaks/ cochlear implants	Asplenia and persistent complement component deficiencies	Chronic liver disease	Diabetes
Hepatitis B[1]	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨
Rotavirus[2]	☐	🟧 SCID* 🟥	🟧	🟧	🟨	🟨	🟨	🟨	🟨	🟨
Diphtheria, tetanus, and acellular pertussis[3](DTaP)	☐	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨
Haemophilus influenzae type b[4]	☐	▨	▨	▨	🟨	🟨	🟨	▨	🟨	🟨
Pneumococcal conjugate[5]	☐	▨	▨	▨	▨	🟨	▨	▨	🟨	🟨
Inactivated poliovirus[6]	🟧	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨
Influenza[7]	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨
Measles,mumps,rubella[8]	🟥	🟥	🟥	🟨	🟨	🟨	🟨	🟨	🟨	🟨
Varicella[9]	🟥	🟥	🟥	🟨	🟨	🟨	🟨	🟨	🟨	🟨
Hepatitis A[10]	🟪	🟨	🟨	🟨	🟨	🟨	🟨	🟨	▨	🟨
Meningococcal ACWY[11]	🟪	🟨	▨	▨	🟨	🟨	🟨	▨	🟨	🟨
Tetanus, diphtheria, and acellular pertussis[12] (Tdap)	▨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨
Human papillomavirus[13]	☐	▨	▨	▨	🟨	🟨	🟨	🟨	🟨	🟨
Meningococcal B[11]	☐	🟪	🟪	🟪	🟪	🟪	🟪	▨	🟪	🟪
Pneumococcal polysaccharide[5]	🟪	▨	▨	▨	▨	▨	▨	▨	▨	▨

🟨 Vaccination according to the routine schedule recommended

🟪 Recommended for persons with an additional risk factor for which the vaccine would be indicated

▨ Vaccination is recommended, and additional doses may be necessary based on medical condition. See footnotes.

☐ No recommendation

🟥 Contraindicated

🟧 Precaution for vaccination

B

FIG 90.4, cont'd (B) Vaccines that might be indicated for children and adolescents aged 18 years or younger based on medical indications. (Courtesy CDC; https://www.cdc.gov.)

prophylaxis (PrEP), this approach, in an ideal setting with unlimited resources and complete patient compliance, could have a truly dramatic impact on HIV incidence. However, to date, uptake has been low, and patient adherence to treatment continues to be a concern. Many of those most at risk for HIV infection do not have medical insurance and cannot afford to purchase medications.

Although the advances in HIV treatment, treatment-as-prevention, and PrEP have been significant, the numbers of new infections globally remain unacceptably high, with 2.1 million new infections reported in 2015 and a total of 36.7 million people living with HIV infection.[63] In the United States, progress on the prevention of HIV infections through the use of condoms, education, and evidence-based interventions has plateaued at ≈44 000 new infections annually. Another remarkable achievement of ARDs has been the dramatic reduction in AIDS-related deaths in the United States, from a peak of ≈55 000 per year to ≈13 000 per year. As a consequence of the unabated incidence of HIV infection and the reduced mortality, the number of people living with HIV in the United States is now increasing at a rate of ≈30 000 per year.

The global need for an HIV vaccine remains strong. However, there are large scientific challenges, and despite more than 30 years of major efforts since the identification of the viral etiological agent of AIDS in 1984, a licensed HIV vaccine with proven protective efficacy continues to be elusive. The field has completed (only) five efficacy trials, which did not achieve protection in vaccinated higher-risk subjects relative to placebo recipients,[64-68] and in two of the trials (which tested replication-deficient adenovirus serotype 5 recombinant vaccine vectors expressing HIV Gag, Pol, and Nef but not Env), there were reports of higher rates of HIV infection in some subsets of participants in the vaccine groups compared with the placebo groups.[69,70]

Importantly, modest vaccine efficacy was observed in a sixth efficacy trial reported in 2009.[71] This 16 000-person study, known as RV144, which was conducted by the US Army in collaboration with the government of Thailand, evaluated a prime-boost regimen of a nonreplicating canarypox vector prime (expressing HIV Gag, protease and gp120) followed by boosting with the same vector plus a bivalent gp120 protein adjuvanted in alum. To the surprise of some early critics of the trial and many in the field, the RV144 regimen produced modest (31.2%) protection in the general population of Thailand at 3.5 years after vaccination, and 61% protection in the first year after vaccination.[71]

This first evidence of human protection provided by an HIV vaccine proved that development of an HIV vaccine would be possible, and this has significantly reenergized the field. In follow-up laboratory studies, an NIH-funded international collaborative program that assessed the immunological correlates of RV144 infection risk revealed that the level of nonneutralizing immunoglobulin G (IgG) antibody targeting a V1V2 loop antigen from the HIV Env gp120 protein correlated inversely with risk of infection, whereas the level of IgA antibody against gp120 protein directly correlated with infection.[72] The IgG antibody was capable of mediating antibody-dependent cellular cytotoxicity (ADCC), whereas the IgA antibody competed with IgG for binding to gp120 V2.[72] The US National Institute of Allergy and Infectious Diseases (NIAID)—funded HIV Vaccine Trials Network (HVTN), the US Army, and country level and industry collaborators have formed the Pox-Protein Public-Private Partnership (also known as "P5") to plan an intensive series of follow-up human studies to confirm and fully investigate the important leads provided to the field by RV144.[73] Efficacy trials that will test next-generation HIV vaccine regimens of clade C–expressing poxvirus vectors with adjuvanted protein boosts and may include DNA vector priming are being planned for southern Africa. In that region, the HIV clade C (the most prevalent global clade) predominates, and the population has a higher risk of HIV acquisition compared with Thailand, where the clades E and B predominate. The coming decade in HIV vaccine research will be highly interesting. Will the "improved versions" of the RV144 regimen provide protection in southern Africa among the higher-risk, genetically distinct population of Africans exposed to C clade HIV?

Discovery of a vaccine immunogen that induces broadly neutralizing antibodies remains the holy grail of HIV vaccination. Such antibodies do occur naturally in up to 15% of chronically infected persons, but only after years of infection. Although seen in natural infection, no vaccine has been able to induce these broadly neutralizing antibodies in vaccinated humans. However, several broadly neutralizing monoclonal antibodies (mAbs) have been cloned, and a few have been tested for safety and pharmacokinetics in early-phase human trials. The antibody genes are highly mutated from their germline with ≈20–25% mutations, indicating that extensive somatic hypermutation (SHM) has occurred in the lymph-node germinal centers (GCs). They also tend to have short complementarity-determining region 1 (CDR1) arms and as a result of autoreactivity against self lipids, these clones are frequently deleted.[74]

The NIAID, through two of its HIV/AIDS clinical trials networks (the HVTN and the HIV Prevention Trials Network [HPTN]), is conducting a phase IIB efficacy trial of VRC01, one of the broadly neutralizing mAbs. In laboratory virus neutralization assays, this antibody has activity against >90% of 343 strains of HIV-1 across all clades at an IC80 of <1 μg/mL.[75,76] The ongoing VRC01 efficacy trial (or the antibody-mediated protection [AMP] study) is a randomized, double-blind, placebo-controlled clinical trial, where the antibody will be infused intravenously every other month for 18 months at doses of 0 mg/kg (placebo), 15 mg/kg, or 30 mg/kg. If the VRC01 antibody does protect at-risk humans against HIV acquisition, this will provide a clear impetus to the HIV vaccine field to continue and intensify efforts to develop immunogens capable of inducing similar or further improved broadly neutralizing antibodies. Notably, the preparation used in the VRC01 trial is a single antibody, and many experts feel that a combination of mAbs targeting different epitopes on the gp120 envelope protein structure may provide an even stronger result. The AMP study result is eagerly awaited.

Improved Influenza Vaccines

Seasonal influenza A imposes a significant disease burden, with the highest morbidity occurring in children and the highest mortality occurring in persons over 65 years of age. The CDC estimates that in the United States, annually influenza is responsible for 200 000 hospitalizations in infants and children and 3000 to 49 000 deaths (depending on the year), with 90% occurring in seniors. Influenza case numbers peak in February during most years, and 10% of the population can become infected in a season. In the United States, starting from 2010, it is recommended that all persons ≥6 months of age receive an annual influenza vaccine. If fully implemented, this would entail administration of 350 million doses each year.

There are two types of influenza, A and B, responsible for most human disease, and multiple subtypes of influenza A, which are categorized on the basis of the amino acid (AA) sequence homology within the HA and NA proteins. Currently 19 HA and nine NA subtypes are recognized, and these can combine to form many viral subtypes (*e.g.,* H1N1, H3N2, H5N1), which are further defined as specific strains based on AA sequence. Influenza B is further categorized into two lineages, Yamagata and Victoria.

Influenza is a negative-stranded RNA virus that lacks a proofreading function in its viral polymerase; hence it is highly mutable. One of the truisms about influenza is that the only thing predictable about influenza is that it is unpredictable. Because of the RNA genome mutations that continually accumulate during influenza replication, the WHO issues a recommendation every February for a new vaccine strain composition. Vaccine manufacturers in the Northern Hemisphere then race to produce the year's trivalent or quadrivalent seasonal vaccine by late summer to be ready for the winter influenza season. There are multiple challenges and needs with regard to this repeated (endless) annual process of influenza vaccine prediction, production, distribution, implementation, uptake, and protection (Fig. 90.2).

For the segment of the population that has the highest mortality, that is, older adults, currently available vaccines provide only relatively weak protection against influenza diagnosed with polymerase chain reaction (PCR). One approach to this challenge has been vaccination of persons ≥65 years of age with a fourfold higher dose of antigen, 60 μg (rather than 15 μg) for each strain based on HA protein content. This high-dose vaccine was shown to increase both immunogenicity (ref) and efficacy and was licensed for use in seniors in the United States.[11,12]

In European and many other countries, for a number of years, an oil-in-water adjuvant, MF59, has been licensed for use in influenza vaccine formulations specifically for seniors. The tolerability and safety of this adjuvanted vaccine has been demonstrated in tens of millions of recipients, and the increases in immune response and efficacy are well documented.[13] In 2015, on the basis of safety and antibody response data from clinical trials, the US FDA approved this oil-in-water adjuvant, designated as MF59, in a formulation of an influenza vaccine for use in seniors.[14] Therefore for seasonal influenza vaccination of seniors in the United States, clinicians will now, as in many other countries, have an adjuvanted influenza vaccine option that provides enhanced immunogenicity.

Another approach to increased protection of seniors against mortality caused by is better community (herd) immunity. The CDC's ACIP recommendation for universal influenza vaccination of all persons >6 months of age has the potential to reduce mortality in seniors, even when a senior citizen's antibody level remains suboptimal after vaccination. If nearly everyone around a senior citizen is vaccinated and protected, a susceptible senior would be much less likely to come in contact with the influenza virus. Enhancing uptake of the annual vaccine by all children and adults is thus an effective approach to improving protection of the vulnerable population of seniors.

Seasonal influenza vaccine effectiveness varies year to year for each influenza A subtype or influenza B lineage, depending on the degree of match between the vaccine strains and the circulating strains. To allow vaccine manufacturers the 6 months currently required for egg-based vaccine production methods, the vaccine strains for each subtype and lineage must currently be selected in February for the following season's vaccine campaign. The burden of annual revaccination of the entire population against a variable viral target is high, both logistically and financially. Furthermore, uptake of the annual seasonal influenza vaccine in the general population remains suboptimal. Over the past decade, considerable effort has been put into the development of the so-called universal influenza vaccine. The universal vaccine would target conserved regions on the HA protein (*e.g.,* the highly conserved stalk rather than the hypervariable head) or other conserved viral proteins (*e.g.,* the M protein). Some of the universal influenza vaccine candidates are moving into early-phase human safety and immunogenicity clinical trials. If a safe and effective universal influenza vaccine can be developed, the goal would be to replace annual revaccination against the current season's strains with initial vaccination and periodic boosting with a universal vaccine that should protect against any influenza variant that arises each year (*e.g.,* a booster every decade as for tetanus). A further advantage would be pandemic preparedness and prevention because the universal influenza vaccine, by targeting conserved viral domains present on all viruses, should also protect against avian strains, such as A/H5N1 or A/H7N9.[77,78] Finally, recent studies comparing a standard-dose conventional (egg-produced) quadrivalent vaccine to a recombinant quadrivalent HA protein vaccine produced by a cell-line in serum-free medium demonstrated better protection in older adults with the recombinant than the conventional vaccine, thus suggesting the possibility of a newer chapter in vaccine production.[78a]

Zika Virus

A current challenge that also encapsulates the recent history of pathogens reemerging and spreading globally from continent to continent is the ongoing Zika virus epidemic. Zika virus is a reemerging flavivirus closely related to dengue virus. As of 2015, this single-stranded positive-sense RNA virus has been causing an epidemic of mosquito-borne dengue-like illness across South and Central Americas and the Caribbean. Infections in pregnant women, with resulting fetal microcephaly or other birth anomalies, are the most feared and serious consequences of Zika virus infection.[7] A mosquito-borne viral cause of birth defects has never existed before. In general, healthy adults with symptomatic infections experience a mild to moderate self-limiting viral illness that has been described as "mild dengue" and is mostly characterized by fever, rash, conjunctivitis, and arthritis. An increased association of Guillain-Barré syndrome (GBS) with Zika virus infection has been reported in multiple countries, and research is ongoing to determine if a causal relationship exists. Interestingly, it is believed that 80% of Zika virus infections are asymptomatic. In symptomatic infected adults, viremia persists for less than a week in most cases, but longer durations of viral RNA detection are reported in semen and urine.[79,80]

In summer 2016, most Zika virus infections in the continental United States were diagnosed in travelers who returned from Zika virus epidemic areas in Latin America or in their sexual contacts. However, a few cases of local transmission had occurred by midsummer 2016 in southern United States. These southernmost parts of the country will likely experience ongoing limited local transmissions as a result of the broad susceptibility of the population and the presence of the primary mosquito vector *Aedes aegypti.* The projection for only limited local transmission in the United States is based on recent prior experiences with dengue and chikungunya, which are also spread by the same mosquito vector. However, the risk and magnitude of harm that will occur in the United States because of this emerging virus is truly unknown.

Although the virus was first identified nearly 70 years earlier in 1947 in the Zika forest in Uganda, only rare human infections had been documented in Africa and Asia for 60 years. Relatively little scientific work was conducted. Then in 2007, a Zika virus outbreak occurred in the Yap Islands in Micronesia,[81] and in 2013, there was an outbreak in French Polynesia.[82] By 2015, large numbers of persons in Brazil were becoming infected, and in 2016, the virus continued to spread across the Americas and the Caribbean islands. Associations of the Zika virus with increased incidences of microcephaly in newborns and GBS in adults led the WHO to declare a "Global Public Health Emergency of International Concern" on February 1, 2016. Several governments in the affected regions have advised women to delay pregnancy. No vaccine or treatment is currently available.

As a response to the Zika virus epidemic, the development of a vaccine has emerged as a top priority of the US government. Preclinical work has demonstrated efficacy and human vaccine trials are on-going. A WHO website indicates that over 15 pharmaceutical and governmental entities are working on finding a Zika virus vaccine.

The vaccine field has a sense of optimism that a Zika virus vaccine can be developed relatively quickly, given that there are several existing flavivirus vaccines in use around the world. The live attenuated yellow fever virus vaccine 17D, developed by Theiler, has been extensively used since 1927. Chimeric yellow fever virus 17D–dengue virus vaccines have been licensed in several countries, and a live attenuated dengue vaccine was recently shown to be protective in a human challenge study.[21,51] Furthermore, tickborne encephalitis virus vaccines have been in use for decades in Russia and Europe, and a West Nile virus vaccine that expresses viral antigens from a DNA vector was developed by the NIAID and was shown to have acceptable safety and promising immunogenicity in humans.

Other Needed Vaccines

It is beyond the scope of this chapter to describe in detail all the new or improved vaccines that are needed to address significant infectious disease burdens. Brief mention is given to a few more needed vaccines that have been reviewed elsewhere recently.

The Ebola epidemic of 2014–2015 highlighted the importance of having a vaccine ready for the next inevitable outbreak of that recurrent periodic killer disease. Major field trials of candidate Ebola vaccines were launched in 2015, but before the major efficacy endpoints were achieved, the epidemic waned. However, a preliminary result of strong efficacy for the recombinant vesicular stomatitis virus (rVSV)–Ebola glycoprotein vaccine was

ON THE HORIZON

- As valuable as vaccines have been in improving health in the past century, they are likely to be even more valuable in the next century.
- Emerging infections capable of severe morbidity (*e.g.,* Zika virus) or mortality (*e.g.,* Ebola virus and avian influenza virus A/H7N9) continue to emerge and threaten global human health with alarming regularity.
- Having rapid vaccine platforms "on the shelf"—to enable quick production, preclinical testing and human safety, immunogenicity and efficacy testing—is critical to timely and effective responses to outbreaks that may spread globally very rapidly as a result of international air travel.
- Several older technologies for vaccine production (*e.g.,* egg-based production of influenza antigens) or immune response assessment (*e.g.,* hemagglutination inhibition assays) will be replaced by newer, improved methods (*e.g.,* cell culture–derived recombinant antigens and automated high-throughput neutralization assays).
- There will be increased application of new molecular adjuvants (*e.g.,* CD40L, Toll-like receptor-5 [TLR-5], or interleukin-12 [IL-12]) that target specific biochemical pathways to direct and shape the cascade from innate immune response to desired adaptive immunity protective response.
- The worsening problem of antibacterial resistance is inspiring new efforts to prevent morbidity and mortality resulting from multidrug-resistant, hospital-acquired infections through vaccination—that is, it is better to prevent an infection caused by a highly resistant microbe than to try to treat it.

obtained in Guinea, where the epidemic persisted somewhat longer than it did in Sierra Leone or Liberia.[83,84] Even though the progression of candidate Ebola vaccines to field trials was rapid (and impressive), the experience with the timing of the availability of the Ebola vaccine for efficacy testing and field use versus the waning arc of the Ebola epidemic points to a recurring concern. This same concern also arose during the 2009 influenza A pandemic—namely, when major epidemics of highly infectious and lethal infectious diseases occur, can protective vaccines be developed quickly enough to have an impact and save lives? Some who have raised this concern have suggested that one solution is to prepare a library and/or repository of vaccines or platforms for all possible epidemic pathogens in advance so as to ensure readiness.

A major international effort to develop a malaria vaccine has been ongoing for many years, and recent efficacy trials have produced encouraging results.[85,86]

Along with HIV/AIDS and malaria, TB is one of the three top infectious killers each year and remains a major target for development of a vaccine to reduce mortality and morbidity.[87-89]

In the past decade, *Bordetella pertussis* infection, or whooping cough, cases in the United States have increased to levels not observed since the 1950s. The transition from the previously used whole-cell pertussis vaccine to the current acellular pertussis vaccine is thought to be responsible. Although the older vaccine was more reactogenic, it was also more protective. Pertussis vaccine developers and public health officials are working to develop new vaccines and to determine the correlates of protective efficacy and are considering the best options to reduce the incidence of pertussis.[90]

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Kyaw MH, Lynfield R, Schaffner W, et al. Effect of introduction of the pneumococcal conjugate vaccine on drug-resistant Streptococcus pneumoniae. N Engl J Med 2006;354:1455–63.
2. Munoz FM, Bond NH, Maccato M, et al. Safety and immunogenicity of tetanus diphtheria and acellular pertussis (Tdap) immunization during pregnancy in mothers and infants: a randomized clinical trial. JAMA 2014;311:1760–9.
3. Zaman K, Roy E, Arifeen SE, et al. Effectiveness of maternal influenza immunization in mothers and infants. N Engl J Med 2008;359:1555–64.
4. Steinhoff MC, Omer SB, Roy E, et al. Influenza immunization in pregnancy–antibody responses in mothers and infants. N Engl J Med 2010;362:1644–6.
5. Mysorekar IU, Diamond MS. Modeling Zika Virus Infection in Pregnancy. N Engl J Med 2016.
6. Brasil P, Pereira JP Jr, Raja Gabaglia C, et al. Zika Virus Infection in Pregnant Women in Rio de Janeiro - Preliminary Report. N Engl J Med 2016;375:2321–34.
7. Rasmussen SA, Jamieson DJ, Honein MA, et al. Zika Virus and Birth Defects–Reviewing the Evidence for Causality. N Engl J Med 2016;374:1981–7.
8. Driggers RW, Ho CY, Korhonen EM, et al. Zika Virus Infection with Prolonged Maternal Viremia and Fetal Brain Abnormalities. N Engl J Med 2016;374:2142–51.
9. Pilishvili T, Lexau C, Farley MM, et al. Sustained reductions in invasive pneumococcal disease in the era of conjugate vaccine. J Infect Dis 2010;201:32–41.
10. Pilishvili T, Bennett NM. Pneumococcal disease prevention among adults: Strategies for the use of pneumococcal vaccines. Am J Prev Med 2015;49:S383–90.
11. DiazGranados CA, Dunning AJ, Kimmel M, et al. Efficacy of high-dose versus standard-dose influenza vaccine in older adults. N Engl J Med 2014;371:635–45.
12. Dunning AJ, DiazGranados CA, Voloshen T, et al. Correlates of Protection against Influenza in the Elderly: Results from an Influenza Vaccine Efficacy Trial. Clin Vaccine Immunol 2016;23:228–35.
13. Black S. Safety and effectiveness of MF-59 adjuvanted influenza vaccines in children and adults. Vaccine 2015;33(Suppl. 2):B3–5.
14. FDA approves first seasonal influenza vaccine containing an adjuvant. U.S. Food and Drug Administration. Available at: http://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm474295.htm. [Accessed 30 May 2016].
15. Hopkins D. The Greatest Killer, Smallpox in History. Chicago, IL: University of Chicago Press Books; 2002.
16. Smallpox: A great and terrible scourge. NIH. Available at: https://www.nlm.nih.gov/exhibition/smallpox/sp_variolation.html. [Accessed 30 May 2016].
17. Willis NJ. Edward Jenner and the eradication of smallpox. Scott Med J 1997;42:118–21.
18. An Inquiry into the Causes and Effects of the Variolae Vaccinae A Disease Discovered in Some of the Western Counties of England, Particularly Gloucestershire, and Known by the Name of the Cow Pox. Public domain; 1798. Available at: http://www.bartleby.com/38/4/1.html. [Accessed 30 May 2016].
19. Wolfe RM, Sharp LK. Anti-vaccinationists past and present. BMJ 2002;325:430–2.
20. Memoli MJ, Shaw PA, Han A, et al. Evaluation of Antihemagglutinin and Antineuraminidase Antibodies as Correlates of Protection in an Influenza A/H1N1 Virus Healthy Human Challenge Model. MBio 2016;7(2):e00417–16.
21. Kirkpatrick BD, Whitehead SS, Pierce KK, et al. The live attenuated dengue vaccine TV003 elicits complete protection against dengue in a human challenge model. Sci Transl Med 2016;8:330ra36.
22. Bernstein DI, Atmar RL, Lyon GM, et al. Norovirus vaccine against experimental human GII.4 virus illness: a challenge study in healthy adults. J Infect Dis 2015;211:870–8.

23. Roestenberg M, de Vlas SJ, Nieman AE, et al. Efficacy of preerythrocytic and blood-stage malaria vaccines can be assessed in small sporozoite challenge trials in human volunteers. J Infect Dis 2012;206:319–23.
24. Gaynes RP. Germ Theory: Medical Pioneers in Infectious Diseases: ASM Press; 2011.
25. Emil von Behring - Facts. Nobel Media. Available at: http://www.nobelprize.org/nobel_prizes/medicine/laureates/1901/behring-facts.html. [Accessed 1 August 2016].
26. Modlin J, Wenger J. Achieving and maintaining polio eradication–new strategies. N Engl J Med 2014;371:1476–9.
27. Hagan JE, Wassilak SG, Craig AS, et al. Progress toward polio eradication - worldwide, 2014-2015. MMWR Morb Mortal Wkly Rep 2015;64:527–31.
28. Dowdle WR. The principles of disease elimination and eradication. Bull World Health Organ 1998;76(Suppl. 2):22–5.
29. Enders JF, Robbins FC, Weller TH. Classics in infectious diseases. The cultivation of the poliomyelitis viruses in tissue culture by John F. Enders, Frederick C. Robbins, and Thomas H. Weller. Rev Infect Dis 1980;2:493–504.
30. Grabenstein JD, Klugman KP. A century of pneumococcal vaccination research in humans. Clin Microbiol Infect 2012;18:15–24.
31. Watson DA, Musher DM, Jacobson JW, et al. A brief history of the pneumococcus in biomedical research: a panoply of scientific discovery. Clin Infect Dis 1993;17:913–24.
32. Swift HF, Wilson AT, Lancefield RC. Typing Group a Hemolytic Streptococci by M Precipitin Reactions in Capillary Pipettes. J Exp Med 1943;78:127–33.
33. Hannoun C. The evolving history of influenza viruses and influenza vaccines. Expert Rev Vaccines 2013;12:1085–94.
34. Dove A. Maurice Hilleman. Nat Med 2005;11:S2.
35. Plotkin S. History of vaccination. Proc Natl Acad Sci USA 2014;111:12283–7.
36. Avery OT, Goebel WF. Chemo-Immunological Studies on Conjugated Carbohydrate-Proteins: Ii. Immunological Specificity of Synthetic Sugar-Protein Antigens. J Exp Med 1929;50:533–50.
37. Schneerson R, Barrera O, Sutton A, et al. Preparation, characterization, and immunogenicity of Haemophilus influenzae type b polysaccharide-protein conjugates. J Exp Med 1980;152:361–76.
38. Riedel S. Edward Jenner and the history of smallpox and vaccination. Proc (Bayl Univ Med Cent) 2005;18:21–5.
39. Deria A, Jezek Z, Markvart K, et al. The world's last endemic case of smallpox: surveillance and containment measures. Bull World Health Organ 1980;58:279–83.
40. World Health Organization. The Global Eradication of Smallpox: Final Report of the Global Commission for the Certification of Smallpox Eradication, Geneva, December 1979. Geneva, Switzerland: World Health Organization; 1980.
41. CDC. Ten great public health achievements–United States, 1900-1999. MMWR Morb Mortal Wkly Rep 1999;48:241–3.
42. Roush SW, Murphy TV, Vaccine-Preventable Disease Table Working G. Historical comparisons of morbidity and mortality for vaccine-preventable diseases in the United States. JAMA 2007;298:2155–63.
43. CDC. Ten great public health achievements–United States, 2001-2010. MMWR Morb Mortal Wkly Rep 2011;60:19–23.
44. Whitney CG, Zhou F, Singleton J, et al. Benefits from immunization during the vaccines for children program era - United States, 1994-2013. MMWR Morb Mortal Wkly Rep 2014;63:352–5.
45. Potkin A, Orenstein W, Offit PA. Vaccines. 6th ed. Philadelphia, PA: Saunders; 2013.
46. Zipprich J, Winter K, Hacker J, et al. Measles outbreak–California, December 2014-February 2015. MMWR Morb Mortal Wkly Rep 2015;64:153–4.
47. Tate JE, Cortese MM, Payne DC, et al. Uptake, impact, and effectiveness of rotavirus vaccination in the United States: review of the first 3 years of postlicensure data. Pediatr Infect Dis J 2011;30:S56–60.
48. Markowitz LE, Liu G, Hariri S, et al. Prevalence of HPV After Introduction of the Vaccination Program in the United States. Pediatrics 2016;137:1–9.
49. Vogt TM, Wise ME, Bell BP, et al. Declining hepatitis A mortality in the United States during the era of hepatitis A vaccination. J Infect Dis 2008;197:1282–8.
50. Marin M, Zhang JX, Seward JF. Near elimination of varicella deaths in the US after implementation of the vaccination program. Pediatrics 2011;128:214–20.
51. Villar L, Dayan GH, Arredondo-Garcia JL, et al. Efficacy of a tetravalent dengue vaccine in children in Latin America. N Engl J Med 2015;372:113–23.
52. Querec TD, Akondy RS, Lee EK, et al. Systems biology approach predicts immunogenicity of the yellow fever vaccine in humans. Nat Immunol 2009;10:116–25.
53. Li S, Rouphael N, Duraisingham S, et al. Molecular signatures of antibody responses derived from a systems biology study of five human vaccines. Nat Immunol 2014;15:195–204.
54. Mooney M, McWeeney S, Canderan G, et al. A systems framework for vaccine design. Curr Opin Immunol 2013;25:551–5.
55. Kuri-Cervantes L, Fourati S, Canderan G, et al. Systems biology and the quest for correlates of protection to guide the development of an HIV vaccine. Curr Opin Immunol 2016;41:91–7.
56. Immunization Schedules. US Department of Health and Human Services; 2016. Available at: http://www.cdc.gov/vaccines/schedules/hcp/index.html. [Accessed 13 February 2016].
57. Recommended Adult Immunization Schedule—United States - 2016; 2016. Available at: http://www.cdc.gov/vaccines/schedules/downloads/adult/adult-schedule.pdf. [Accessed 13February 2016].
58. Recommended Immunization Schedules for Persons Aged 0 Through 18 Years. United States 2016. US Department of Health and Human Services; 2016. Available at: http://www.cdc.gov/vaccines/schedules/downloads/child/0-18yrs-child-combined-schedule.pdf. [Accessed 13 February 2016].
59. WHO recommendations for routine immunization - summary tables; 2015. Available at: http://www.who.int/immunization/policy/immunization_tables/en/ [Accessed 6 April 2016].
60. Cohen MS, Chen YQ, McCauley M, et al. Antiretroviral Therapy for the Prevention of HIV-1 Transmission. N Engl J Med 2016;375:830–9.
61. Cohen MS, Chen YQ, McCauley M, et al. Prevention of HIV-1 infection with early antiretroviral therapy. N Engl J Med 2011;365:493–505.
62. Koester KA, Grant RM. Editorial Commentary: Keeping Our Eyes on the Prize: No New HIV Infections With Increased Use of HIV Pre-exposure Prophylaxis. Clin Infect Dis 2015;61:1604–5.
63. Global AIDS Update 2016. UN AIDS; 2016. Available at: http://www.who.int/hiv/pub/arv/global-AIDS-update-2016_en.pdf. [Accessed 1 August 2016].
64. Hammer SM, Sobieszczyk ME, Janes H, et al. Efficacy trial of a DNA/rAd5 HIV-1 preventive vaccine. N Engl J Med 2013;369:2083–92.
65. Gray GE, Allen M, Moodie Z, et al. Safety and efficacy of the HVTN 503/Phambili study of a clade-B-based HIV-1 vaccine in South Africa: a double-blind, randomised, placebo-controlled test-of-concept phase 2b study. Lancet Infect Dis 2011;11:507–15.
66. Buchbinder SP, Mehrotra DV, Duerr A, et al. Efficacy assessment of a cell-mediated immunity HIV-1 vaccine (the Step Study): a double-blind, randomised, placebo-controlled, test-of-concept trial. Lancet 2008;372:1881–93.
67. Pitisuttithum P, Gilbert P, Gurwith M, et al. Randomized, double-blind, placebo-controlled efficacy trial of a bivalent recombinant glycoprotein 120 HIV-1 vaccine among injection drug users in Bangkok, Thailand. J Infect Dis 2006;194:1661–71.
68. Flynn NM, Forthal DN, Harro CD, et al. Placebo-controlled phase 3 trial of a recombinant glycoprotein 120 vaccine to prevent HIV-1 infection. J Infect Dis 2005;191:654–65.
69. Moodie Z, Metch B, Bekker LG, et al. Continued Follow-Up of Phambili Phase 2b Randomized HIV-1 Vaccine Trial Participants Supports Increased HIV-1 Acquisition among Vaccinated Men. PLoS ONE 2015;10:e0137666.
70. Duerr A, Huang Y, Buchbinder S, et al. Extended follow-up confirms early vaccine-enhanced risk of HIV acquisition and demonstrates waning effect over time among participants in a

randomized trial of recombinant adenovirus HIV vaccine (Step Study). J Infect Dis 2012;206:258–66.
71. Rerks-Ngarm S, Pitisuttithum P, Nitayaphan S, et al. Vaccination with ALVAC and AIDSVAX to prevent HIV-1 infection in Thailand. N Engl J Med 2009;361:2209–20.
72. Haynes BF, Gilbert PB, McElrath MJ, et al. Immune-correlates analysis of an HIV-1 vaccine efficacy trial. N Engl J Med 2012;366:1275–86.
73. Corey L, Gilbert PB, Tomaras GD, et al. Immune correlates of vaccine protection against HIV-1 acquisition. Sci Transl Med 2015;7:310rv7.
74. Haynes BF, McElrath MJ. Progress in HIV-1 vaccine development. Curr Opin HIV AIDS 2013;8:326–32.
75. Lynch RM, Boritz E, Coates EE, et al. Virologic effects of broadly neutralizing antibody VRC01 administration during chronic HIV-1 infection. Sci Transl Med 2015;7:319ra206.
76. Ledgerwood JE, Coates EE, Yamshchikov G, et al. Safety, pharmacokinetics and neutralization of the broadly neutralizing HIV-1 human monoclonal antibody VRC01 in healthy adults. Clin Exp Immunol 2015;182:289–301.
77. Krammer F. Novel universal influenza virus vaccine approaches. Curr Opin Virol 2016;17:95–103.
78. Cho A, Wrammert J. Implications of broadly neutralizing antibodies in the development of a universal influenza vaccine. Curr Opin Virol 2016;17:110–15.
78a. Dunkle LM, Izikson R, Patriarca P, et al. Efficacy of recombinant influenza vaccine in adults 50 years of age or older. N Engl J Med 2017;376:2427–36.
79. Interim Guidance for Zika Virus Testing of Urine - United States, 2016. MMWR Morb Mortal Wkly Rep 2016;65:474.
80. Matheron S, D'Ortenzio E, Leparc-Goffart I, et al. Long Lasting Persistence of Zika Virus in Semen. Clin Infect Dis 2016;63:1264.
81. Duffy MR, Chen TH, Hancock WT, et al. Zika virus outbreak on Yap Island, Federated States of Micronesia. N Engl J Med 2009;360: 2536–43.
82. Cao-Lormeau VM, Roche C, Teissier A, et al. Zika virus, French polynesia, South pacific, 2013. Emerg Infect Dis 2014;20:1085–6.
83. Henao-Restrepo AM, Longini IM, Egger M, et al. Efficacy and effectiveness of an rVSV-vectored vaccine expressing Ebola surface glycoprotein: interim results from the Guinea ring vaccination cluster-randomised trial. Lancet 2015;386:857–66.
84. An Ebola vaccine: first results and promising opportunities. Lancet 2015;386:830.
85. Birkett AJ. Status of vaccine research and development of vaccines for malaria. Vaccine 2016;34:2915–20.
86. Hoffman SL, Vekemans J, Richie TL, et al. The March Toward Malaria Vaccines. Am J Prev Med 2015;49:S319–33.
87. Evans TG, Schrager L, Thole J. Status of vaccine research and development of vaccines for tuberculosis. Vaccine 2016;34:2911–14.
88. Hoft DF. Tuberculosis vaccine development: goals, immunological design, and evaluation. Lancet 2008;372:164–75.
89. Kaplan G. Rational vaccine development–a new trend in tuberculosis control. N Engl J Med 2005;353:1624–5.
90. Warfel JM, Edwards KM. Pertussis vaccines and the challenge of inducing durable immunity. Curr Opin Immunol 2015;35:48–54.

MULTIPLE-CHOICE QUESTIONS

1. The physician often recognized as the "father of vaccinology" is:
   A. James Phipps
   B. Robert Koch
   C. Edward Jenner
   D. Louis Pasteur

2. The polysaccharide vaccines developed for the prevention of bacterial diseases caused by pathogenic organisms, such as *Haemophilus influenzae, Neisseria meningitidis,* and *Streptococcus pneumoniae,* became much more useful upon conjugation with:
   A. Lipids
   B. Small molecule metabolites
   C. Proteins
   D. Polymers

3. The Advisory Committee on Immunization Practices (ACIP) of the US Centers for Disease Control and Prevention (CDC) in its adult immunization schedule makes annual recommendations for each vaccine based on the age of the patient, and other indications. In the 2017 schedule, these included:
   A. All adults (persons ≥19 years of age) should receive annual influenza vaccination; a tetanus–diphtheria–acellular pertussis (Tdap) vaccine once, followed by tetanus boosters every 10 years; zoster vaccination at 60 years of age; and pneumococcal vaccination at 65 years of age.
   B. Other vaccines indicated for certain risk factors, including medical conditions (*e.g.*, immunocompromising conditions, kidney failure, diabetes) or life style, occupational, or other conditions (*e.g.*, pregnancy, men who have sex with men, health care personnel).
   C. Live vaccines (varicella, zoster, and measles–mumps–rubella [MMR]) are contraindicated for pregnant women, immunocompromised hosts, and in those with human immunodeficiency virus [HIV] infection when the CD4+ T-cell absolute count is below 200 cells/μL.
   D. All of the above are correct.

91

Immunotherapy of Allergic Disease

Anthony J. Frew

Specific-allergen immunotherapy (SIT), or allergen desensitization, is used to treat various forms of allergic disease that involve type I hypersensitivity. SIT involves the administration of allergen extracts to modify or abolish symptoms associated with exposure to relevant allergens. SIT is used for three distinct types of allergic condition: anaphylaxis, rhinoconjunctivitis, and asthma. In anaphylaxis, the patient is completely well between episodes; exposure is infrequent and unexpected but provokes an immediate and potentially catastrophic response. However, since exposure only occurs occasionally, it can be difficult to know whether the treatment has been effective until an accidental exposure event occurs. Allergic rhinoconjunctivitis is clearly driven by allergen exposure, which is low level and may be either continuous or intermittent. Exposure and clinical relevance are most obvious for seasonal rhinitis (caused by pollens or molds), occupational rhinitis, and for rhinitis caused by animal danders. Because there is no seasonal variation in perennial rhinitis, it is more difficult to be sure that allergen exposure drives symptoms—in some cases it clearly does, but in others there may be nonallergic or structural causes for the symptoms. The value of SIT for rhinoconjunctivitis is generally clearer when exposure is predictable, but assessing efficacy is complicated by year-to-year variations in pollen exposure because of weather patterns and individual activity. The third indication for SIT is asthma, where the role of allergen exposure is different from that in rhinoconjunctivitis. As discussed elsewhere, allergen sensitization is a risk factor for developing asthma in childhood, but it is less clear whether allergen exposure has a role in ongoing asthma. At the very least, the fact that allergen avoidance measures are not usually effective in asthma begs the question whether modifying the response to allergen will influence established disease. It is inherently more risky to use SIT to treat patients with asthma compared with treating those who do not have asthma, so the benefits and risks must be weighed carefully when considering the use of SIT to treat patients with asthma.

The general principles of managing any allergic condition are to make an accurate diagnosis, identify relevant trigger factors, institute appropriate interventions to reduce the impact of those triggers, and control symptoms and disease progression. Allergen avoidance measures may help but are rarely sufficient to avoid the need for other therapies. Drug treatments can be very effective, but they only work as long as the drugs are taken. SIT is the only current therapy that modifies the immune response to allergens. It is specific in that treatment is targeted at those allergens recognized by the patient and physician as responsible for symptoms. Although claims have been made for bystander benefits on unrelated allergens, there is little convincing evidence that SIT for one allergy can alleviate symptoms caused by different allergens. Before deciding to use SIT, it is, therefore, essential to assess the patient's condition carefully, especially the role of allergic triggers.

SIT was pioneered over 100 years ago and developed as an empirical science of pollen vaccination, based on the mistaken concept that hay fever was caused by infection.[1] Early studies found that giving large doses of pollen would trigger allergic reactions, so regimes that start with extremely low doses and build up to a maintenance dose that is given every few weeks were developed. A great deal has been learned about the immunological events associated with clinically successful SIT, but uncertainties remain about which of these events is truly important in delivering the clinical benefit. New and modified forms of SIT have been developed to deliver these immunological effects more efficiently, but none of these, as yet, has the efficacy of the standard approach.

Following initial reports from the United Kingdom, SIT was developed in the 1920s–30s in North America, where some differences in practice emerged compared with European practice. In particular, American allergists tend to treat patients for all sensitivities identified on skin testing, using personalized mixtures of extracts prepared from bulk vials, whereas in Europe, patients are normally treated with single allergens, which are supplied directly by the manufacturer. However, pollens from different grasses or trees are often combined. Mixtures of allergens from different sources (*e.g.,* mites and pollens) are used in some parts of Europe as custom mixes from manufacturers. Different approaches are also taken for standardization of extracts. Allergen extracts used in Europe are standardized by their ability to elicit a weal, whereas US standardization has been based on erythema rather than weal.

Whichever form of extract is chosen, patients are started on a very low dose of allergen, and the dose is then increased, usually at weekly intervals, until reaching the maintenance dose, which is then given at 4- to 6-weekly intervals for 3–5 years. Alternative induction protocols may involve several doses on each day (semirush), or the whole series of incremental injections may be given in a single day (rush). The main drawback to rush and semirush protocols is the frequency of adverse reactions, which occur much more often than with conventional protocols. However, full protection can be attained in a few days compared with the 3 months required in the conventional regime. Normally, the doses are given by subcutaneous injection, but since 1990, there has been increasing interest in SIT administered by the sublingual route, which is more convenient for patients and has some advantages in terms of safety.

MECHANISMS OF SIT

KEY CONCEPTS

Possible Mechanisms of Immunotherapy

- Induction of immunoglobulin G (IgG) (blocking) antibodies
- Reduction in specific IgE (long-term)
- Reduced recruitment of effector cells
- Altered T-cell cytokine balance (shift to T-helper cell-1 [Th1] from Th2)
- T-cell anergy
- B-cell suppression

The primary reason for studying the mechanisms of SIT is to identify features that are biologically important and thus to devise new forms of immunotherapy that may improve efficacy, increase safety margins, shorten treatment courses, or achieve more durable results from SIT. At least three distinct phases can be identified in SIT, each with distinct immunological mechanisms.[2] Initially, there is pharmacological desensitization, in which basophils and mast cells are rendered tolerant of the allergen. This is achieved quickly, within a few days if the rush protocol is followed. Clinically, this not only allows the patient to tolerate the maintenance dose of allergen but also protects against acute exposure to allergen (*e.g.,* a wasp sting). This state of pharmacological tolerance is not permanent and rapidly disappears if SIT is stopped. During the first year of maintenance SIT, T-cell tolerance is achieved, with abrogation of late-phase responses to allergen exposure and induction of regulatory T cells (Tregs). Finally, with prolonged treatment, B-cell tolerance develops, and the level of sensitizing antibodies decreases.

Following subcutaneous injection of allergen extracts, a small proportion of the allergenic material is taken up by phagocytic cells and carried to regional lymph nodes. The process is fairly inefficient: <1% of a radiolabeled dose actually reaches the lymph nodes. Almost all studies have shown that SIT induces allergen-specific immunoglobulin G (IgG) antibodies, which increase progressively during treatment. This led to suggestions that SIT might work by inducing antibodies that intercept the allergen and "block" the allergic response. In patients treated for venom anaphylaxis, the development of allergen-specific IgG antibody correlates with clinical efficacy, but for other allergens, the magnitude of the IgG response is not closely related to the degree of efficacy. Moreover, the rise in IgG follows, rather than precedes, onset of clinical benefit. Allergen-specific IgE antibodies increase initially, but the usual seasonal rise in IgE is blunted.[3] Over several years, the amount of allergen-specific IgE declines but does not disappear. In keeping with this, there is little effect on immediate skin test responses to allergen. In contrast, the late-phase skin test response is virtually abolished after successful SIT. Similar patterns are observed for late-phase nasal and airway responses.[4]

SIT also affects allergen-specific T cells. In both skin and the nose, successful SIT is accompanied by a reduction in T-cell and eosinophil recruitment in response to allergen challenge. In parallel, the balance of T-helper cell-1 (Th1) and Th2 cytokine expression (Chapter 16) alters at allergen-challenged sites. Th2 cytokine expression is not affected, but an increased proportion of the recruited T cells express the Th1 cytokines interleukin-2 (IL-2), interferon-γ (IFN-γ), and IL-12. In addition, there is induction of allergen-specific CD4 Tregs that express CD25, FOXP3, and IL-10.[2] IL-10 is induced within a few days of starting SIT and has a number of biological effects that could explain the beneficial effects of immunotherapy. These include modulation of IL-4–induced B-cell IgE production in favor of IgG4, inhibition of IgE-dependent mast cell activation, inhibition of human eosinophil cytokine production and survival, suppression of IL-5, and induction of antigen-specific anergy (Fig. 91.1). Taken together, these findings suggest that SIT modulates allergen-specific T cells, which explains why clinical and late-phase responses are attenuated without much effect on allergen-specific antibody levels. Further work in this area is now concentrating on finding more efficient ways of inducing allergen-specific Tregs. Some caution is needed: Most of the evidence of Treg induction is from peripheral blood rather than from target organs, and although

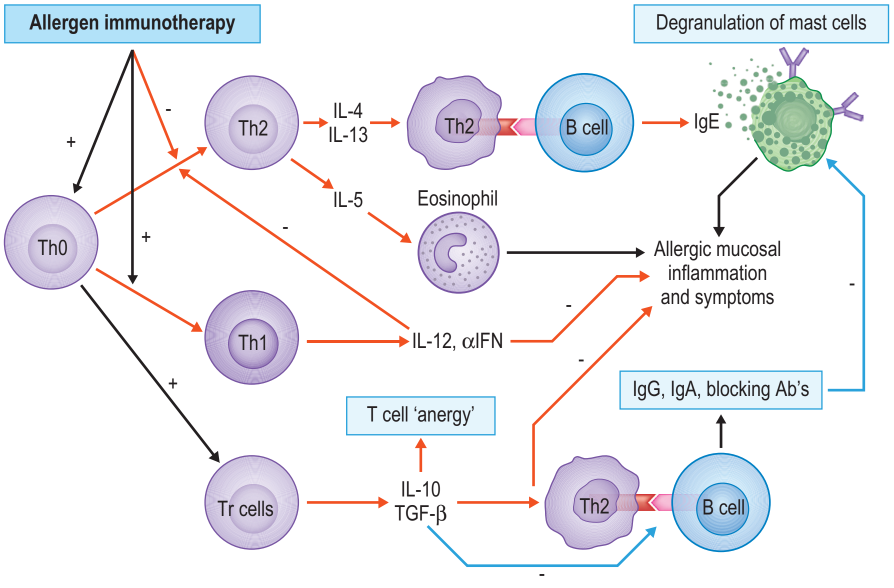


FIG 91.1 Immunological mechanisms of specific immunotherapy.

there are group effects, there is no direct relationship to clinical efficacy in individuals.

SIT for Venom Anaphylaxis

Anaphylaxis to Hymenoptera venom is relatively rare but can be fatal. Venom-specific IgE antibodies can be found in 30–40% of adults for a few months following a sting, but usually these disappear subsequently. Some individuals develop high concentrations of venom-specific antibodies, which may persist for many years without further stings. This group of patients is at risk of anaphylaxis to subsequent stings: a small number die of anaphylaxis each year. Fatality rates are hard to estimate, perhaps 10–20 deaths/year in the United States.

Before embarking on venom immunotherapy (VIT), the patient needs to be carefully assessed, and some account should be taken of the natural history of the venom allergy.[5] Patients who have experienced systemic symptoms after stings are at much greater risk of anaphylaxis on subsequent stings, compared with patients who have only had large local reactions. The frequency of systemic sting reactions in children and adults with a history of large local reactions is about 5–10%, whereas the risk in patients with previous systemic reactions is about 30–70%. In general, there is a lower risk of repeated systemic reactions in children and in those with a history of milder reactions. In adults, the risk of systemic reaction to field stings diminishes over 10–20 years toward 15–30% (Fig. 91.2), but does not return to the background general population prevalence (3%). Children with a history of cutaneous systemic reactions had <5% risk of anaphylaxis during observation for 10–20 years. No test can accurately predict the outcome of the next sting. Live sting challenges have been used in research but are not practical or acceptable in clinical practice. Unfortunately, sting challenges are not absolutely predictive, since patients who do not react to laboratory challenge stings may react to subsequent field stings.

In deciding to recommend VIT, account should be taken of occupational and geographical factors that affect the likelihood of future stings. Bee stings are much more common in beekeepers, their families, and neighbors, whereas wasp stings tend to be sporadic but are an occupational hazard for bakers, gardeners, outdoor caterers, greengrocers, and other similar trades. Other factors to consider are the potential risks of emergency treatment with epinephrine and medical contraindications to VIT (see below). Large local reactions to stings are not regarded as indications for VIT, as they are not life-threatening and do not predict future problems, but there is some evidence they can be attenuated by VIT.[6]

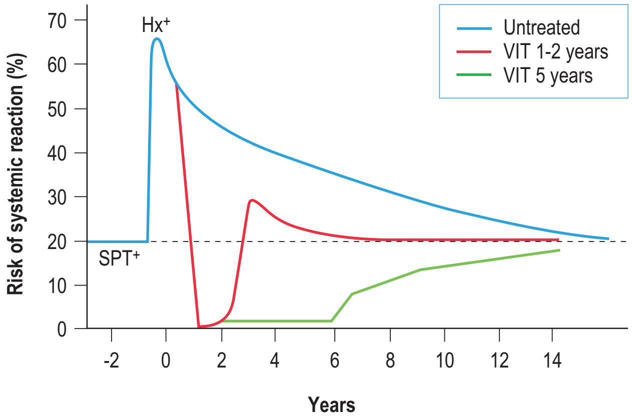


FIG 91. 2 Risk of systemic reaction (SR) from wasp stings over time after an anaphylactic event, and the effect of specific-allergen immunotherapy for 1–2 years or 5 years on the natural history. VIT, venom immunotherapy. (Adapted from Golden DB, et al. Survey of patients after discontinuing venom immunotherapy. J Allergy Clin Immunol 2000; 105: 385–90.)

VIT accelerates the process of risk reduction and confers rapid protection against field and laboratory stings. There is a low risk of systemic reaction (≈10%) for many years after discontinuing VIT. Although most reactions to stings after completing VIT are mild, fatalities have been reported in patients with other risk factors (mastocytosis, severe historical reactions, systemic reaction during VIT). This raises the question whether VIT should be life long, although lifelong treatment would alter the cost–benefit analysis. In children, the chance of systemic sting reactions remains <5% for up to 20 years after discontinuing VIT.

In summary, desensitization with venom preparations provides protection against field and laboratory stings, but some patients will still react despite completing VIT. The residual risk of systemic reactions to stings after VIT is about 10%, but these are usually mild. Patients undergoing VIT for venom anaphylaxis are usually offered injectable epinephrine and other antiallergic medication to use if stung during VIT, but this is unnecessary once they are on maintenance doses.

Assessing Effectiveness in Clinical Trials of SIT for Asthma and Rhinitis

The two main keys to success in designing any clinical trial are application of appropriate criteria for subject recruitment and selection of appropriate clinical endpoints (Table 91.1). In allergic rhinitis, efficacy is generally assessed in terms of symptom scores, recorded daily on diary cards or personal organizers.[7] Several symptom scales are in use, which all rely on a categorical grading system (*e.g.*, 1 = mild, 2 = moderate, 3 = severe, etc.). These measures carry the implicit assumption that the step from 1 to 2 is clinically equivalent to the step from 2 to 3. By assigning numerical values to nonparametric variables, one may either overestimate or underestimate the true effects of therapy. Reduction in rescue medication use is also measured, as, in theory, symptoms can improve because patients use more medication. In practice, most studies that showed improved symptom scores showed parallel reductions in medication use, but some studies showed reductions in medication use without a statistically

TABLE 91.1 Endpoints for Trials of Specific Immunotherapy in Asthma

Spirometry	Air flow obstruction (forced expiratory volume per second [FEV_1], peak expiratory flow rate [PEFR]) Bronchial irritability (histamine/methacholine PC_{20})
Symptoms	Wheeze, nocturnal waking, cough, breathlessness, time off work/school (diary cards or visual analogue scales)
Drug use	Bronchodilator use as surrogate for symptoms
Cost–benefit ratio	Reduction in other drugs, lost work, etc. compared with costs of specific-allergen immunotherapy (SIT) and time lost due to SIT
Safety	Frequency of local and systemic adverse events
Immunology	Allergen-specific immunoglobulin G4 (IgG4) Skin test responses Nasal allergen challenge responses Inhalation allergen challenge responses

significant reduction in symptoms. Combining symptom and medication scores is fraught with difficulty. It is almost impossible to agree on equivalences between different drugs, let alone the relative value of one antihistamine tablet versus a day when symptoms are moderate rather than mild.

When assessing improvement in asthma, it is important to remember that the majority of patients with atopic asthma have relatively mild disease that is usually easy to control with standard doses of conventional antiasthma drugs. Cases of severe asthma are relatively few in number, and the role of allergy in driving their disease is often less obvious than in mild cases. In mild cases, resting spirometry will usually be close to predicted values, so there is little scope for improvement in simple spirometric measures, such as forced expiratory volume per second (FEV_1). Peak expiratory flow rates (PEFRs) may show improvement after SIT, especially the early morning values, which are more likely to be reduced below the patient's best possible value compared with evening peak flow rates. Objective measures of bronchial irritability are more useful markers of improvement and a shift in the concentration of histamine or methacholine needed to induce a 20% fall in FEV_1 (PC_{20}) is commonly used. This measurement can be quite variable within individuals, so the minimum change that would be clinically significant is a change of at least one doubling dilution in the group geometrical mean PC_{20}. Other bronchial challenge tests, such as exercise challenge or adenosine inhalation, may provide additional insights into the ability of an antiasthma treatment to prevent acute attacks. Both these parameters depend indirectly on the PC_{20}, so again, it is important to consider statistical and biological significance separately.

Although these challenge models are useful surrogates for evaluating antiasthma treatment, more weight should be placed on improvements in the symptoms or progression of the disease than on protection against laboratory challenges. Lately, there has been a move to assess efficacy in terms of improvements in quality of life or health gain per unit cost. These analyses are becoming increasingly sophisticated but require careful validation to ensure that the quality-of-life instrument is sensitive to the illness being studied and that the economic model is realistic.

Finally, many trials of SIT for asthma and rhinitis have studied laboratory surrogate markers, such as changes in IgG subclasses, T-cell function, or skin test responses. Although such paraclinical measures do change during SIT and can shed light on the mechanisms of successful SIT, they cannot be used to provide direct evidence of clinical efficacy. Studies that show antibody or T-cell changes without direct clinical evidence of efficacy neither support nor undermine the established evidence that SIT works, albeit in mysterious ways, to improve the symptoms of appropriately selected patients. Recent research on paraclinical measures has identified four areas that may prove useful in future: dendritic cell (DC) markers in sublingual immunotherapy (SLIT)[8]; innate lymphoid cells (ILCs) favoring Th2 development (ILC-2), which have been show to decrease in frequency after successful SIT; decreased expression of diamine oxidase—the enzyme responsible for making histamine; and IgG antibodies that block IgE-mediated functional antigen presentation.[9]

SIT for Allergic Rhinitis

Allergic rhinitis is the main indication for SIT worldwide. As with all uses of SIT, it is important to select patients appropriately. The allergic basis of the rhinitis should be carefully assessed during history taking and skin or blood tests for IgE, and other relevant causes should be excluded. Tests of nasal sensitivity are not used in routine clinical practice but may be useful for assessing effectiveness in clinical trials.

The effectiveness of SIT in seasonal allergic rhinitis has been confirmed in many trials, using grass, ragweed, and birch pollen. Moreover, SIT has been shown to be clinically effective even in patients with severe seasonal rhinitis that is resistant to conventional drug therapy.[10] Data regarding the long-term efficacy of SIT for allergic rhinitis are limited, but formal studies have shown that the effects last for at least 3 years after discontinuing therapy.[11] Longer-term studies are difficult to conduct, but efficacy has been shown for up to 10 years in open studies. The benefits of SIT for perennial rhinitis are less well established than for seasonal rhinitis. In part, this reflects the difficulty in determining the extent to which allergy is responsible for perennial symptoms. Allergy to house dust mite (HDM) is common, but only about 50% of those with positive skin tests have symptoms. Conversely, there are other causes of perennial rhinitis, including vasomotor instability, infection, aspirin sensitivity, and so on. Nevertheless, clinical trials have shown a definite benefit provided the subjects are appropriately selected.[12] Clearer evidence of efficacy has been obtained in rhinitis caused by pet allergy. Several studies have shown a marked improvement in tolerance of cat exposure after SIT, confirmed both on challenge tests and simulated natural exposure.[13]

As with any therapy, the risks and cost-effectiveness of SIT need to be assessed on a case-by-case basis. Current drug therapy for rhinitis can be very effective, but about 60% of patients with seasonal allergic rhinitis report inadequate symptom control, even when taking maximal doses of antihistamines and intranasal glucocorticoids. Others experience nose bleeds caused by intranasal steroids and drowsiness caused by antihistamines. Moreover, we are now more aware of the adverse effects of rhinitis on quality of life. SIT offers a useful option for these patients as well as a logical approach to dealing with the underlying problem.

SIT for Asthma

Immunotherapy has been widely used to treat allergic asthma, but evidence of severe adverse reactions, including a small number of fatalities, has led to SIT being abandoned for asthma treatment in the United Kingdom since 1986, although asthma remains a common indication for SIT in North America and continental Europe.[12,14] There is no doubt that SIT is effective on asthma symptoms and protects against specific and nonspecific triggers. However, the role of allergic sensitization in ongoing asthma is less clear than for rhinitis, and the risks of therapy are undoubtedly greater in asthma. Current drug therapies for asthma aim to suppress the airways inflammation and smooth muscle contraction, which are characteristic features of asthma. None of these treatments is curative, and asthma recurs rapidly upon cessation of treatment. Moreover, current drug therapies are directed against downstream mechanisms rather than against the agents that might cause asthma. Allergen avoidance has been proposed as a potentially useful maneuver in those with allergic asthma; however, although asthma control can be improved by extreme forms of allergen avoidance (*e.g.*, admission to hospital, sending children to holiday homes at altitude), there is little evidence that similar benefits can be achieved by using the type of allergen avoidance that can be achieved in suburban homes. Thus there is scope for improving asthma care and for identifying allergen-specific therapies. SIT offers the possibility of diverting the immune response away from the allergic pattern and toward a more protective or less damaging response.

The efficacy of SIT in adult asthma has been assessed in many trials over the last 60 years. The results of these studies are difficult to interpret, either because poor quality allergen extracts were used or because of poor study design. Many trials were not placebo controlled; they were either open or single-blind studies; and in most cases, only small numbers of patients were treated. A 2010 meta-analysis[14] reviewed all of the relevant papers published between 1954 and 2008. Many of the earlier trials had methodological flaws, but there are sufficient data for general conclusions to be drawn. Symptom scores improved in the treated groups—it was necessary to treat four patients to prevent one from experiencing exacerbation of symptoms, and to treat five to prevent one from needing an increase in medication. SIT reduced the airways response to inhalation of specific allergen and also improved nonspecific bronchial reactivity.

In double-blind placebo-controlled (DBPC) trials in patients with grass pollen asthma, active treatment led to a 60–75% reduction in symptom scores compared with placebo-treated patients. Similar results were found in ragweed allergy with improved peak flow rates during the pollen season and reduced hay fever symptoms and sensitivity to laboratory challenge with ragweed pollen extracts. However, the associated economic analysis indicated that the cost saving in asthma drugs was less than that in SIT. The value of SIT is thus dependent on its ability to improve symptoms and quality of life rather than to reduce the costs of other medications.

In patients with asthma who are sensitive to cats, SIT with standardized aqueous extracts reduces the early asthmatic response to inhaled allergen and attenuates responses to room exposure to cats. Interestingly, these studies showed a clear delay in onset of symptoms and an overall reduction in symptoms and peak flow recordings after simulated exposure to cats, but there was no protection against allergen-induced increases in nonspecific bronchial hyperresponsiveness. Another study of cat SIT for asthma using a depot preparation found a reduction in both specific and nonspecific bronchial reactivity (to cat extract and histamine, respectively). There is, however, no hard evidence that SIT is of benefit to treat asthma caused by dog allergy.

Several DBPC studies have examined the effects of SIT in patients with asthma who are sensitive to HDM. Early studies used tyrosine-adsorbed or depot preparations, which yielded conflicting results, with reductions in drug requirements in children but no improvement in asthma control in adults. Subsequent studies using more modern extracts have found beneficial effects on symptoms, antiasthma medication use, and bronchial hyperresponsiveness.

SIT for Food Allergy

After several false starts, research has now shown that it is possible to tolerize patients to food allergens, such as peanut. Arguments continue over what to call this, as the tolerance achieved seems to be temporary, and patients become fully sensitive again if they stop treatment. However, the possibility to render patients with life-threatening food allergy tolerant of the relevant food is a useful step forward. In line with the clinical time course, it has been shown that basophil sensitivity is suppressed early during peanut tolerization but returns to baseline when treatment is stopped. In contrast, in patients undergoing SLIT for aeroallergen sensitivity their basophil reactivity takes longer to drop away after starting treatment, but once achieved, the effect on basophils is sustained for months or years after ceasing SLIT. This differential effect on surrogate markers suggests that: (i) current forms of oral immunotherapy (OIT) for food allergy are achieving a temporary remission; (ii) strategies that are more efficacious against basophil reactivity may offer a more sustained state of tolerance.[15] In parallel, work is ongoing to develop injection SIT for fish allergies.[16]

Comparison of SIT With Other Therapies

The majority of clinical trials of SIT for asthma have compared SIT either with untreated historical controls or with a matched placebo-treated group. To date, the effectiveness of SIT in asthma has rarely been compared with conventional management (avoidance measures and inhaled or oral antiasthma drugs).

Effects of SIT on the Natural History of Allergic Disease

In recent years, there has been increasing interest in the possibility that SIT may have disease-modifying or preventive effects. These are important not only in their own right as desirable outcomes but also as factors in the economic evaluation of SIT.

Children often start with a limited range of allergic sensitivities and progress over time to develop IgE against a wider range of inhaled allergens. Treatment with SIT may limit this tendency to acquire new sensitizations, although the clinical benefit of this preventive effect is not clear. In the most widely cited study, children with asthma who were sensitized only to HDM were treated with HDM-SIT. Of the 19 treated children, 10 did not develop new sensitizations, whereas all 22 untreated children acquired new sensitivities, in many cases to more than one new allergen.[17] A long-term study of treatment with a grass pollen allergoid has confirmed this observation, with 50% fewer new sensitizations 12 years after completion of the course of SIT.[18] Similar results have been found in a study of children treated with SLIT, in which 3% of SLIT-treated children and 34% of control subjects developed new sensitizations within 3 years.[19] It seems probable that this effect is indirect. In other words, it is unlikely that HDM-SIT directly affects the B cells that recognize cat or grass pollen, but by treating the HDM allergy, SIT may reduce inflammation in the nose and hence modify the likelihood that exposure to other allergens will lead on to sensitization.

Each year, a proportion of patients with allergic rhinitis go on to develop asthma. This rate of progression has been estimated at 5% per annum in college students,[20] but this remains, perhaps surprisingly, an area of considerable uncertainty. Several long-term epidemiological studies are now in progress, under the auspices of the International Study of Asthma and Allergies in Childhood (ISAAC), and these should eventually shed light on the rate of progression at different ages and the extent of regional and international variation. It has been suggested that SIT may modify the natural history of asthma in children who have allergic rhinitis but have not yet developed asthma. Data to support this proposition are limited. In the key study, a group of 205 children aged 6–14 years, without previously diagnosed asthma, were treated with SIT for birch or grass pollen allergy in an open randomized design. Three years after completing treatment, 45% of the untreated group had developed asthma, whereas only 26% of the treated group had asthma. These results have been sustained out to 7 years after completing therapy. Thus in terms of number needed to treat, four children had to be treated to prevent one case of asthma, which makes this an extremely effective therapy.[21] Similar results have been reported with SLIT: After 3 years of SLIT, 8 of 45 SLIT-treated children and 18 of 44 control subjects had developed asthma, representing a relative risk of 3.8 for untreated subjects compared with the actively treated group.[22]

SIT may also modify the progression of established asthma. An early open study using uncharacterized mixed allergen extracts supported this view, with about 70% of treated children losing their asthma after 4 years therapy, compared with about 19% of untreated controls, a result that was sustained up to the age of 16 years. The proportion of children whose asthma was severe at age 16 years was also much lower in the treated group.[23] In this period, when clinical trials were conducted differently, well before the days of ethics committees, Johnstone and Dutton randomized all children attending their clinic between August 1953 and January 1955 to one of four treatments. Neither the subjects nor their parents were aware of the treatment being given. Compared with those who received placebo or very dilute allergens, those treated with the two higher doses of SIT were much more likely to lose their asthma after 4 years, and among those whose asthma persisted, the likelihood of having severe disease was much lower in those who had received active treatment. Although it would be very difficult to conduct such a study today, the data are useful in confirming the dose–response relationship in SIT and its potential for reducing the duration and severity of asthma. A prospective, nonrandomized, open-label study of SLIT has reported that children with asthma who received SLIT had less asthma compared with those on standard medication after 4–5 years, and this difference persisted for 5 years after cessation of treatment.[24]

In contrast, there is no current evidence that SIT influences the evolution of established asthma in adults. In part, this reflects the reluctance of physicians to use SIT to treat patients with severe asthma. Studies that have investigated withdrawal of therapy have reported rapid recurrence of asthma symptoms, although rhinitis symptoms seem to show much more sustained relief after SIT.[25]

In summary, there is persuasive evidence that both subcutaneous and sublingual immunotherapy can modify the course of allergic disease, by reducing the incidence of new sensitizations, and by preventing or slowing the development of clinical asthma. The mechanisms that underlie these observations are not fully understood. It is likely that a combination of immunological and structural alterations is involved. Economic evaluations of the benefits of SIT and SLIT need to take on board these preventive effects to arrive at a full cost–benefit analysis.

SAFETY

The main factor preventing the wider adoption of SIT is the risk of serious adverse reactions. In general, SIT is quite safe in patients who do not have asthma, but a significant number of deaths have been reported in the United Kingdom and the United States when SIT was used to treat patients whose asthma was unstable.[26] The incidence of systemic reactions in patients receiving SIT for asthma varies among study series and has been reported to range from 5% to 35%. In contrast, serious systemic reactions occur after about 1 in 500 injections in patients with rhinitis. Some adverse events are caused by administration of the wrong dose, which underscores the importance of checking doses carefully before each injection. Others occur during changeover from one vial to another. Most allergists, therefore, cut back on the dose when changing vials to minimize this risk. Systemic reactions are much more likely if the patient has an intercurrent viral illness, so doses should be delayed if the patient is unwell or has any signs of active asthma.

On a separate note, there is general agreement that SIT should not be used in patients with autoimmune disorders or malignant disease. Although there is no hard evidence that SIT is actually harmful in these groups, it seems unwise to attempt manipulation of the immune system in such patients, not least because of the risk that spontaneous and unrelated variations in the autoimmune disorder or cancer may be blamed on SIT. Other medical contraindications to SIT include significant coexistent cardiac disease, which may be exacerbated by any adverse reactions to SIT. Patients on beta-blockers should also not receive SIT. Although they are not at increased risk of adverse reactions, their physiological response to the cardiovascular component of anaphylaxis will be impaired, and they will not respond to the epinephrine that is used to treat adverse reactions to SIT.

SUBLINGUAL IMMUNOTHERAPY

High-dose topical immunotherapy regimes were used in the first half of the twentieth century but then lost ground to injection immunotherapy in conventional medical practice. Over the past 20 years, there has been a resurgence of interest in SLIT, which is based on the concept that allergens given via the mucosal surface can induce immunological tolerance, without carrying the same risk of systemic reactions that is found with injection SIT. Moreover, SLIT is more appealing to patients, especially to those who cannot easily find time to attend clinic for injection SIT.[27] In animal models, IgE responses to allergens can be reduced or prevented by oral administration of allergen. Most of these models involve feeding in early life to prevent subsequent acquisition of sensitivity, so this is a different scenario from the usual clinical context of treating established allergic disease. The precise mechanisms by which "oral tolerance" is induced remain unclear, but it seems likely that the route of allergen processing and presentation is a critical determinant of the subsequent T-cell response (Chapter 6). In mice, locally administered allergen is taken up by mucosal DCs and then presented to T cells together with IL-12, biasing the response toward a Th1-like profile and away from the pro-IgE Th2 profile. In contrast to the animal models, the immunological response to SLIT in human studies has been relatively modest. Some changes have been found in skin sensitivity, but most studies have not found any change in systemic parameters, such as specific IgE, specific IgG, or T-cell cytokine balance.[28]

Nevertheless, a body of evidence has accumulated from well-conducted clinical trials, indicating that SLIT can be effective, with up to 30–40% reductions in symptom scores and rescue medication use in seasonal allergic rhinitis.[29] SLIT treatment regimes typically involve a rapid buildup phase followed by treatment about three times per week with rapidly dissolving tablets containing allergen extracts. Some preparations are supplied in liquid form, with a calibrated dropper. Several large clinical trials of SLIT have shown clinically and statistically significant reductions in symptom and medication scores, supporting the conclusions of previous meta-analyses of SLIT, which assessed the earlier, less well-designed studies that were individually small and inconclusive. The meta-analysis estimated the efficacy of SLIT as being about two-thirds of that seen in comparable studies of injection SIT, but the most recent studies have shown levels of efficacy that are equal to the best recent multicenter trials of injection SIT. However, the efficacy may be greater for pollens than for HDM allergies.[30] Local side effects are common with SLIT but are generally well tolerated. Systemic side effects of SLIT are very rare but have been reported, especially in patients who have previously had systemic reactions to injection SIT. SLIT is now being used routinely in some parts of Europe (especially Italy and France), but the doses and regimes being prescribed are often different from those that were used in the clinical trials. Overall SLIT should widen the scope of SIT and allow more patients to be treated. As with all forms of immunotherapy, patient selection remains the key to ensuring that therapy is targeted to those who are likely to benefit from it.

FUTURE DIRECTIONS

There is clearly scope to improve conventional SIT. Possible avenues include the use of recombinant allergens, which should make it easier to standardize allergen vaccines and may allow fine-tuning of vaccines for patients with unusual patterns of reactivity. Most patients with allergies react to the same components of an allergen extract, the so-called major allergens, which are defined as those allergens recognized by over 50% of sera from a pool of patients with clinically significant allergy to the material in question. However, not all patients recognize all major allergens, and some patients only recognize allergens that are not recognized by the majority of sera of patients with allergies. This latter group may not respond to standard extracts but might be better treated by a combination of allergens to which they are sensitive. Until the advent of molecular cloning, this has been impossible to achieve. Now that recombinant allergens for SIT are available, the range of sensitivities can be better characterized, and this may lead to patient-tailored vaccine products. Thus far, clinical trials have confirmed that recombinant allergen cocktails are effective, but these have not yet shown superiority over conventional vaccines.[31] Regulations are the major barrier to this development: Under current regulations, each combination of allergen components would have to be developed as a freestanding pharmaceutical. Until a different regulatory framework is agreed upon, recombinant allergens will only be tested as fixed combination cocktails.

Another alternative approach is to inject allergens directly into lymph nodes, under ultrasound guidance. Proponents of this approach argue that after subcutaneous injection, less than 1% of the injected material reaches the lymphatic system; therefore if the benefit of SIT is driven by the dose, then direct intranodal administration should be more efficient and also safer than conventional SIT. Early studies with doses of about 1% of standard SIT showed immunological efficacy, with some signs of clinical efficacy and reduced risks of side effects.[32] However, subsequent clinical trials by other groups have not demonstrated efficacy.[33,34]

Novel forms of allergenic molecules may be created: for example, a recombinant trimer consisting of three covalently linked copies of the major birch pollen allergen, Bet v 1 has been made. This trimer is much less allergenic, even though it contains the same B-cell and T-cell epitopes as the native molecules and induces Th1 cytokine release and IgG antibodies, analogous to the antibody response to standard SIT.[35] Folding variants and other modifications of the physical structure may also improve the safety of SIT.

Since the epitopes recognized by IgE molecules are usually three-dimensional, while epitopes recognized by T cells are short linear peptide fragments of the antigen (Chapter 6), it should be possible to manipulate allergens and reduce their ability to bind to IgE or to create peptide fragments, which will modulate T cells without the risk of anaphylaxis. Cross-linking allergen proteins with aldehydes reduces their ability to bind antibodies. Such allergoids do not degranulate the basophils of sensitized individuals *in vitro* but are recognized by allergen-specific T cells. Clinically, allergoids have shown efficacy in rhinitis caused by grass pollen or HDM.[36]

Short peptide sequences that should not be recognized by IgE antibodies can be used as vaccines as well.[35] Peptide vaccines can be either natural sequences or altered peptide ligands. If high doses of natural peptides are given, these deceive the T cell into high dose tolerance. Altered peptide ligands induce anergy by providing an incomplete activation signal. Both approaches are affected by the MHC type of the individual undergoing treatment. By sequential alteration of HDM allergen Der p peptides, it is possible to suppress proliferation of T-cell clones recognizing native Der p peptides, as well as suppressing their expression of CD40 ligand and their production of IL-4, IL-5, and IFN-γ. These anergic T cells do not provide help to B cells to switch class to make IgE; importantly, this anergy cannot be reversed by providing exogenous IL-4. In humans, the main focus has been on cat allergen (Fel d1) peptides, which can reduce the level of symptoms on exposure to cat dander.[37] Similar studies with the peptides of phospholipase A2 (PLA2, a major allergen in bee venom) have also been reported. However, to date, peptide vaccines have not shown any greater efficacy compared with conventional vaccines.

In an animal model, intranasal application of genetically produced hypoallergenic fragments of Bet v1 produced mucosal tolerance with significant reduction of IgE and IgG1 antibody responses, as well as reduced cytokine production *in vitro* (IL-5, IFN-γ, IL-10). These reduced immunological responses were accompanied by inhibition of the cutaneous and airway responses that were seen with the complete Bet v 1 allergen. The mechanisms of immunosuppression seemed to be different for the allergen fragments and the whole molecule, in that tolerance induced with the whole Bet v 1 molecule was transferable with spleen cells, whereas that induced by the fragments was not.[38]

From epidemiological and experimental studies, we know that vaccines with mycobacteria have antiallergic properties. In Japan, early vaccination with the Bacille Calmette-Guérin (BCG) vaccine was associated with a substantial reduction in the risk of developing allergy,[39] although similar associations were not evident in studies in Sweden.[40] In an animal model, administration of the BCG vaccine before or during sensitization to ovalbumin reduced the degree of airway eosinophilia that followed subsequent challenge with ovalbumin. This effect was not mediated through any direct effect on IgE production or blood eosinophil numbers but was mediated through IFN-γ and could be reversed by exogenous IL-5.[41]

Two new approaches using DNA vaccines are also undergoing serious consideration. The first of these is a general approach, using CpG oligodeoxynucleotides (ODN), which mimic bacterial DNA, binding to T-cell receptor-9 (TCR-9) and stimulating a Th1-type cytokine response. This technology is essentially a refinement of vaccination adjuvant technology, since CpG ODN have been shown to be the principal activity from complete Freund adjuvant. In a mouse model of asthma, preadministration of CpG ODN prevented both airway eosinophilia and bronchial hyperresponsiveness. Moreover, these effects were sustained for at least 6 weeks after CpG ODN administration.[42] Alternatively, CpG ODN can be coupled to the allergenic protein, which enhances immunogenicity in terms of eliciting a Th1-type response to the allergen but reduces its allergenicity and stimulates Th1 cytokine expression in cultured human peripheral blood mononuclear cells (PBMCs).[43] Initial clinical trials confirmed that the hybrid vaccine elicits a Th1-pattern response,[44] but the results of subsequent trials have been inconclusive. A contrasting approach is immunization with allergen-specific naked DNA sequences. This technology is still in its infancy, but preliminary data suggest that administration of naked DNA leads to production of allergens from within the airway epithelial cells.[45] It seems that the endogenously produced allergen elicits a Th1-type response, due to the different handling pathways for endogenous and exogenous allergens. If this can be reproduced in humans with allergies, this may overcome the existing Th2-pattern response and eliminate the allergy. However, the potential for generating a powerful Th1-type response to ubiquitous agents means that this approach needs careful evaluation in animal models before it can be pursued in humans.

The recent introduction of monoclonal antibodies (mAbs) directed against IgE offers another option. Treatment with anti-IgE reduces immediate and late-phase responses to inhaled allergens and should also reduce the risk of adverse effects from SIT injections. Moreover, when anti-IgE is combined with conventional SIT, the effects on seasonal allergic rhinitis are additive.[46] However, the high cost of anti-IgE and the need for regular injections are likely to limit its use to patients who have severe allergic disease that cannot be managed by other means, and these patients are not generally suitable candidates for SIT.

CONCLUSIONS

ON THE HORIZON

New Technologies for Immunotherapy

- Recombinant allergens
- T-cell peptide vaccines
- T-helper cell-1 (Th1) immunostimulants (*e.g.*, mycobacteria, CpG)
- Allergen-immunostimulant complexes
- Anti–immunoglobulin E (IgE)
- New routes of administration:
  - Sublingual
  - Intralymphatic
  - Epicutaneous
  - Liposome encapsulation

SIT has been established as a treatment for allergic rhinitis and for venom hypersensitivity but is more controversial for treating allergic asthma. When used in appropriately selected patients, SIT is effective and acceptably safe, but care must be exercised in recognizing and treating adverse reactions. Appropriate training of allergists and SIT clinic support staff is essential. Despite a century of use, the precise mechanisms of action of SIT remain uncertain. Current emphasis on the role of Tregs is leading to renewed attempts to simplify SIT regimes and reduce its risks. Future directions in SIT include the development of vaccines that are better standardized and the use of recombinant allergens, both of which should improve the safety profile of SIT. In parallel, and perhaps for the longer term, we should explore the development of general immunomodulatory therapies, which would be particularly advantageous for those patients sensitized to multiple allergens.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Frew AJ. 100 years of immunotherapy. Clin Exp Allergy 2011;41:1221–5.
2. Jutel M, Akdis CA. Immunological mechanisms of allergen-specific immunotherapy. Allergy 2011;66:725–32.
3. Creticos P, Van Metre TE, Mardiney MR, et al. Dose-response of IgE and IgG antibodies during ragweed immunotherapy. J Allergy Clin Immunol 1984;73:94–104.
4. Iliopoulos O, Proud D, Adkinson NF, et al. Effects of immunotherapy on the early, late and rechallenge nasal reaction to provocation with allergen: changes in inflammatory mediators and cells. J Allergy Clin Immunol 1991;87:855–66.
5. Golden DB, Moffitt J, Nicklas RA, et al. Stinging insect hypersensitivity: a practice parameter update 2011. J Allergy Clin Immunol 2011;127:852–4.
6. Golden DB, Kelly D, Hamilton RG, et al. Venom immunotherapy reduces large local reactions to insect stings. J Allergy Clin Immunol 2009;123:1371–5.
7. Bousquet J, Schünemann HJ, Bousquet PJ, et al. How to design and evaluate randomized controlled trials in immunotherapy for allergic rhinitis: an ARIA-GA(2)LEN statement. Allergy 2011;66:765–74.
8. Gueguen C, Bouley J, Moussu H, et al. Changes in markers associated with dendritic cells driving the differentiation of either TH2 cells or regulatory T cells correlate with clinical benefit during allergen immunotherapy. J Allergy Clin Immunol 2016;137:545–58.
9. Shamji MH, Ljorring C, Francis JN, et al. Functional rather than immunoreactive levels of IgG4 correlate closely with clinical response to grass pollen immunotherapy. Allergy 2012;67:217–26.
10. Frew AJ, Powell RM, Corrigan CJ, et al. Efficacy and safety of specific immunotherapy with SQ allergen extract in treatment-resistant seasonal allergic rhinoconjunctivitis. J Allergy Clin Immunol 2006;117:319–25.
11. Durham SR, Walker SM, Varga EM, et al. Long-term clinical efficacy of grass pollen immunotherapy. N Engl J Med 1999;341:468–75.
12. Cox L, Esch RE, Corbett M, et al. Allergen immunotherapy practice in the United States: guidelines, measures, and outcomes. Ann Allergy Asthma Immunol 2011;107:289–99.
13. Varney VA, Edwards J, Tabbah K, et al. Clinical efficacy of specific immunotherapy to cat dander: a double blind placebo controlled trial. Clin Exp Allergy 1997;27:860–7.
14. Abramson MJ, Puy RM, Weiner JM. Injection allergen immunotherapy for asthma. Cochrane Database Syst Rev 2010;(8):CD001186.
15. Gorelik M, Narisety SD, Guerrerio AL, et al. Suppression of the immunologic response to peanut during immunotherapy is often transient. J Allergy Clin Immunol 2015;135:1283–92.
16. Zuidmeer-Jongejan L, Huber H, Swoboda I, et al. Development of a hypoallergenic recombinant parvalbumin for first-in-man subcutaneous immunotherapy of fish allergy. Int Arch Allergy Immunol 2015;166:41–51.
17. Des Roches A, Paradis L, Menardo JL, et al. Immunotherapy with a standardized *Dermatophagoides pteronyssinus* extract. VI. Specific immunotherapy prevents the onset of new sensitizations in children. J Allergy Clin Immunol 1997;99:450–3.
18. Eng PA, Borer-Reinhold M, Heijnen IA, et al. Twelve-year follow-up after discontinuation of pre-seasonal grass pollen immunotherapy in childhood. Allergy 2006;69:198–201.
19. Marogna M, Tomassetti D, Bernasconi A, et al. Preventive effects of sublingual immunotherapy in childhood: an open randomized controlled study. Ann Allergy Asthma Immunol 2008;101:206–11.
20. Horak F. Manifestation of allergic rhinitis in latent sensitised patients. A prospective study. Arch Otorhinolaryngol 1985;242:242–9.
21. Niggemann B, Jacobsen L, Dreborg S, et al. Five-year follow-up on the PAT study: specific immunotherapy and long-term prevention of asthma in children. Allergy 2006;61:855–9.

22. Novembre E, Galli E, Landi F, et al. Coseasonal sublingual immunotherapy reduces the development of asthma in children with allergic rhinoconjunctivitis. J Allergy Clin Immunol 2004;114:851–7.
23. Johnstone DE, Dutton A. The value of hyposensitization therapy for bronchial asthma in children. A 14 year study. Pediatrics 1968;42:793–802.
24. Di Rienzo V, Marcucci F, Puccinelli P, et al. Long-lasting effect of sublingual immunotherapy in children with asthma due to house dust mite: a 10-year prospective study. Clin Exp Allergy 2003;33:206–10.
25. Shaikh WA. Immunotherapy vs inhaled budesonide in bronchial asthma: an open, parallel, comparative trial. Clin Exp Allergy 1997;27:1279–84.
26. Bernstein DI, Wanner M, Borish L, et al. Twelve-year survey of fatal reactions to allergen injections and skin testing: 1990-2001. J Allergy Clin Immunol 2004;113:1129–36.
27. Frew AJ. Sublingual immunotherapy. N Engl J Med 2008;358:2259–64.
28. Scadding G, Durham SR. Mechanisms of sublingual immunotherapy. Immunol Allergy Clin North Am 2011;31:191–209.
29. Calderon MA, Penagos M, Sheikh A, et al. Sublingual immunotherapy for allergic conjunctivitis: Cochrane systematic review and meta-analysis. Clin Exp Allergy 2011;41:1263–72.
30. de Bot CM, Moed H, Berger MY, et al. Sublingual immunotherapy is not effective in house dust mite-allergic children in primary care. Pediatr Allergy Immunol 2012;23:150–8.
31. Pauli G, Larsen TH, Rak S, et al. Efficacy of recombinant birch pollen vaccine for the treatment of birch-allergic rhinoconjunctivitis. J Allergy Clin Immunol 2008;122:951–60.
32. Senti G, Prinz Vavricka BM, Erdmann I, et al. Intralymphatic allergen administration renders specific immunotherapy faster and safer: a randomized controlled trial. Proc Natl Acad Sci USA 2008;105:17908–12.
33. Witten M, Malling HJ, Blom L, et al. Is intralymphatic immunotherapy ready for clinical use in patients with grass pollen allergy? J Allergy Clin Immunol 2013;132:1248–52.
34. Patterson AM, Bonny AE, Shiels WE, et al. Three-injection intralymphatic immunotherapy in adolescents and young adults with grass pollen rhinoconjunctivitis. Ann Allergy Asthma Immunol 2016;116:168–70.
35. Valenta R, Campana R, Focke-Tejkl M, et al. Vaccine development for allergen-specific immunotherapy based on recombinant allergens and synthetic allergen peptides: Lessons from the past and novel mechanisms of action for the future. J Allergy Clin Immunol 2016;137:351–7.
36. Corrigan CJ, Kettner J, Doemer C, et al. Efficacy and safety of preseasonal-specific immunotherapy with an aluminium-adsorbed six-grass pollen allergoid. Allergy 2005;60:801–7.
37. Norman PS, Ohman JL, Long AA, et al. Treatment of cat allergy with T-cell reactive peptides. Am J Respir Crit Care Med 1996;154:1623–8.
38. Wiedermann U, Herz U, Baier K, et al. Intranasal treatment with a recombinant hypoallergenic derivative of the major birch pollen allergen Bet v 1 prevents allergic sensitization and airway inflammation in mice. Int Arch Allergy Immunol 2001;126:68–77.
39. Shirakawa T, Enomoto T, Shimazu SI, et al. The inverse association between tuberculin responses and atopic disorder. Science 1997;275:77–9.
40. Strannegaard IL, Larsson LO, Wennergren G, et al. Prevalence of allergy in children in relation to prior BCG vaccination and infection with atypical mycobacteria. Allergy 1998;53:249–54.
41. Erb KJ, Holloway JW, Sobeck A, et al. Infection of mice with *Mycobacterium bovis*-BCG suppresses allergen-induced airways eosinophilia. J Exp Med 1998;187:561–9.
42. Sur S, Wild JS, Choudury BK, et al. Long-term prevention of allergic lung inflammation in a mouse model of asthma by CpG oligodeoxynucleotides. J Immunol 1999;162:6284–93.
43. Marshall JD, Abtahi S, Eiden JJ, et al. Immunostimulatory sequence DNA linked to the Amb a 1 allergen promotes Th1 cytokine expression while downregulating Th2 cytokine expression in PBMCs from human patients with ragweed allergy. J Allergy Clin Immunol 2001;108:191–7.
44. Creticos PS, Eiden JJ, Broide D, et al. Immunotherapy with immunostimulatory oligonucleotides linked to purified ragweed Amb a 1 allergen: effects on antibody production, nasal allergen provocation and ragweed seasonal rhinitis. J Allergy Clin Immunol 2002;109:743–4.
45. Hartl A, Hochreiter R, Stepanoska T, et al. Characterisation of the protective and therapeutic efficiency of a DNA vaccine encoding the major birch pollen allergen Bet v1a. Allergy 2004;59:65–73.
46. Rolinck-Werninghaus C, Hamelmann E, Keil T, et al. The co-seasonal application of anti-IgE after preseasonal specific immunotherapy decreases ocular and nasal symptom scores and rescue medication use in grass-pollen allergic children. Allergy 2004;59:973–9.

MULTIPLE-CHOICE QUESTIONS

1. Successful immunotherapy may work through induction of:
   A. B-cell tolerance
   B. Allergen-specific immunoglobulin A (IgA) antibodies
   C. FOXP3$^+$ regulatory T cells
   D. Tryptase$^+$ mast cells

2. Venom immunotherapy:
   A. Prevents large local reactions
   B. Accelerates the rate of tolerance
   C. Confers long-lasting efficacy after 1 year of treatment
   D. Permanently prevents systemic reactions

3. Changes in allergen-specific immunoglobulin G4 (IgG4):
   A. Reflect the dose of allergen injected
   B. Predict response at the individual level
   C. Are caused by induction of interleukin-17 (IL-17)
   D. Can be used to predict clinical efficacy

4. Immunotherapy is a proven treatment for allergy to:
   A. Dog dander
   B. Walnuts
   C. Hen's egg
   D. Tree pollen

92

Flow Cytometry

Sergio D. Rosenzweig

Flow cytometry has become a standard laboratory tool in the evaluation of hematopoietic cells, including the identification of leukocyte populations and subpopulations, a method referred to as *immunophenotyping.* The clinical application of this technology has been facilitated by the development of instruments and data analysis systems suitable for routine use in diagnostic laboratories. In addition, the expanded range of monoclonal antibodies (mAbs) specific for lymphocyte (and other hematopoietic cell) surface antigens directly conjugated to a number of different fluorescent indicators (fluorochromes) provide an extensive panel of reagents that facilitate multicolor (polychromatic) studies.

The clinical needs that pushed this technology date back to the emergence of absolute CD4 T-cell counts as a critical measure for disease assessment and follow-up in managing patients infected with human immunodeficiency virus (HIV). Flow cytometry applied in the monitoring of HIV infection was followed by the routine application of cell characterization by flow cytometry in the evaluation of hematologic malignancies and, more recently, in the study of immunodeficiency disorders and other immune-mediated diseases.

Recent advances in instrumentation and fluorochrome chemistry now allow for routine polychromatic flow cytometry studies, with concomitant assessment of cell surface markers and intracellular parameters, including intracellular proteins, phosphoproteins, and cytokines, as well as identification of changes linked to cellular activation and apoptosis. Intracellular flow cytometry also can be applied to evaluate cell cycle status (*i.e.,* G_0-G_1, S, G_2-M) based on DNA staining, which is useful in evaluating tumor cells and assessing the *in vitro* lymphocyte response to various stimuli. Additionally, evaluation of lymphocyte proliferation can be performed with cell tracking dyes that allow quantitation of the rounds of cell division associated with cell activation, and techniques to assess immune cell mediated cytotoxicity have also been developed. Finally, characterization of antigen-specific T cells following immunization or associated with normal and/or abnormal immune responses in association with disease states can be accomplished by using multimer technology as well as intracellular cytokine detection following antigen exposure.

This chapter focuses on the basic concepts of flow cytometry, including instrument characteristics, data management, lymphocyte gating, and directed use of test reagents. In addition, it provides a brief overview of intracellular protein detection, cell activation and cell-mediated cytotoxicity studies, cell cycle analysis, apoptosis detection, and multimer technology, focusing on the appropriate application of these approaches as well as their limitations.

INSTRUMENTATION

The basic components of a flow cytometer, as shown in Fig. 92.1, include an illumination source, an optical bench, a fluidic system, electronics, and a computer.[1] Briefly, stained cells flow into single file by the fluidic system and are interrogated by one or more light sources; these sources generate light signals, which are directed by the optical system to the photodetectors, which, in turn, convert light into electronic signals for storage and subsequent analysis. This process is discussed further in the section below.

The fluidic system lies at the heart of the flow cytometer and consists of isotonic sheath fluid that moves the sample stream containing the cells. This is accomplished by injecting the cell sample into flowing sheath fluid, establishing a hydrodynamically focused single-file flow of cells (particles) that move through the analysis point while maintaining the cell stream in a constant, central location.[2] The centrally focused cell stream ensures that the illumination of all cells is virtually equivalent. Thus the difference in magnitude of the emission signal(s) generated from each cell reflects biological differences between the cells (rather than reflecting the variation in the illumination energy if the cells were not tightly focused). The use of hydrodynamic focusing has the additional advantage of producing little or no change in cell shape, although it may have an effect on cell orientation. The consistency in maintaining cell shape facilitates distinguishing "architectural" differences between specific leukocyte types (see Gating section).[3] However, this method can generate single-file cell rows with precision only up to a flow rate of 60 to 100 μL/min, which can lead to long acquisition times for the detection of very rare events. To overcome this problem, recently introduced flow cytometry instruments utilize acoustic focusing, which align cells through the use of sound waves, allowing sample flow rates of up to 1000 μL/min, without loss of signal quality.[4,5]

Illumination in standard clinical instruments is generated by two or more lasers, each of which provides a specific monochromatic light source (*e.g.,* a sapphire laser generates a 488-nm-wavelength [blue] beam). Modern lasers are small and available in multiple wavelengths, including ultraviolet (350 nm), violet (405 nm), blue (488 nm), green (532 nm), yellow (560 nm), orange (610 nm), and red (633 nm), permitting the simultaneous use of multiple fluorochromes having different excitation

---

This work was supported in part by the Intramural Research Program of the National Institutes of Health Clinical Center.

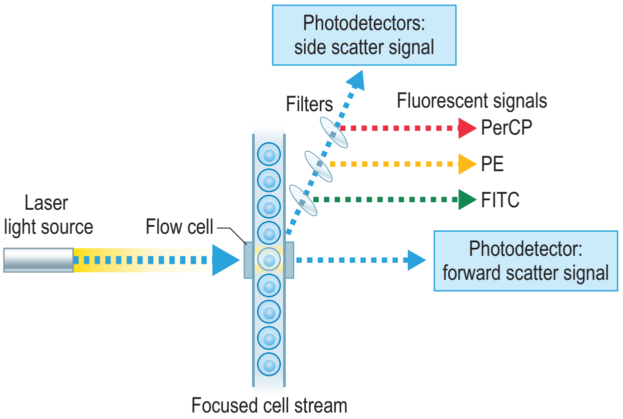


FIG 92.1 Simplified Design of A Flow Cytometer with One Illumination Source (Laser) Set Up to Collect Five Parameters. These include the two nonfluorescent parameters (blue light) forward and side scatter, as well as three fluorescent parameters, green (FITC), orange (PE), and red (PerCP) light.

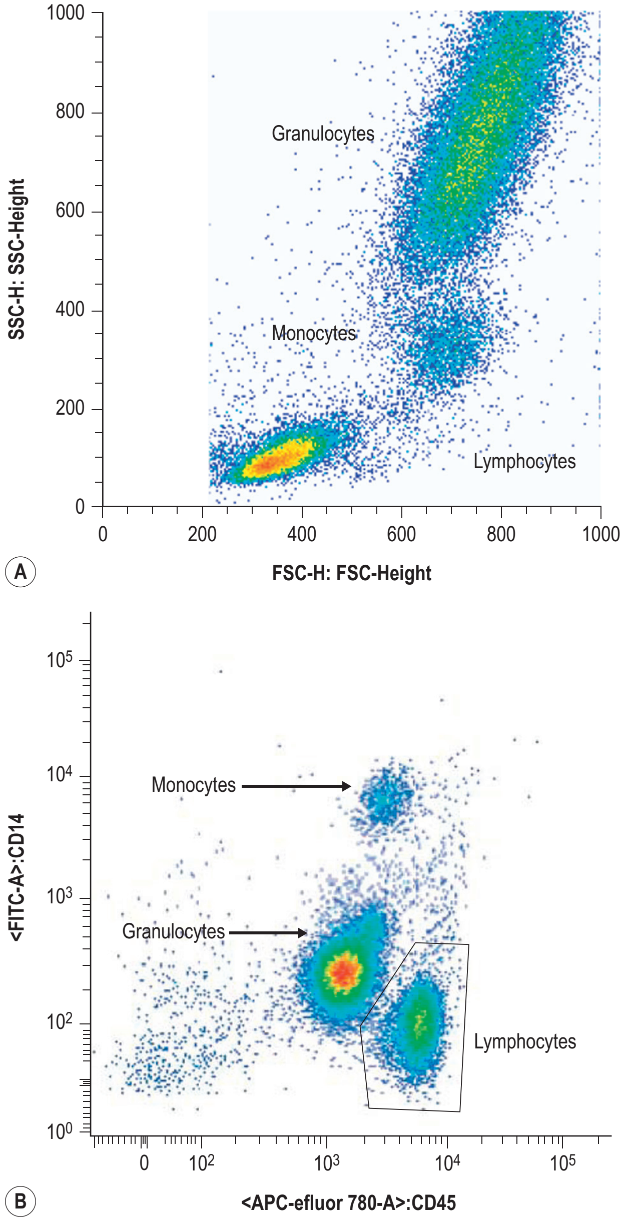


FIG 92.2 (A) Forward- and side-scatter dot plots on a lysed whole blood sample, demonstrating the basic three-part leukocyte differential with lymphocytes, monocytes, and granulocytes. (B) Dot plot with DC45/CD14 gating reagents showing the fluorescence distribution of all the three leukocyte types identified to include lymphocytes, monocytes, and granulocytes, as well as a small number of nonlysed red blood cells and/or debris.

requirements.[6] The point where the light illuminates the cell in analytical instruments occurs within a flow cell, while in cell sorters, the beam intersects cells flowing as a stream in air. The optical bench contains lenses that shape and focus the illumination beam to ensure consistent excitation energy at the analysis point.

The illumination of a cell generates both nonfluorescent and fluorescent signals, which are collected and measured by optically coupling the signal to a detection system consisting of filters, each of which is linked to a photodetector. The filters are chosen to allow the nonfluorescent signals to be measured at the same wavelength as the excitation signal (*e.g.,* 488 nm from a blue light source) for the forward- and side-scatter channels (see Gating section), whereas those for the fluorescence channels utilize specific filters that allow passage of light with wavelengths specific to each fluorochrome (*e.g.,* green, orange, or red; see Fluorochrome section). The number and arrangement of the photodetectors allows for the simultaneous evaluation of multiple colors (parameters) for each cell, with a report describing a modified clinical instrument capable of evaluating 17 or more colors simultaneously from each cell evaluated.[7]

The internal electronics in the flow cytometer provides the system for converting analog light signals (photoelectrons) received at the photodetectors into digital signals for acquisition and storage in a computer. The intensity of these converted signals is measured on a relative scale that is generally set in either 256 or 1024 equal increments (referred to as *channels*) for display and analysis. A number of specialized analysis programs are available, and results are depicted graphically as single-parameter histograms displaying specific light (fluorescence) intensity (*x*-axis) versus cell number (*y*-axis) (Fig. 92.2), or two-color displays where the *x*-axis and the *y*-axis reflect the light intensity of the two colors, and the cell numbers are represented via dot, pseudocolor, contour, or density plots (Fig. 92.3). Most analysis programs enable the operator to evaluate the number and percentage of events, mean and/or median channel fluorescence, and selected statistical measures for each identified cell, and these can be aggregated into specific populations and/or subpopulations of cells. Thus a flow cytometer provides a platform with the capacity to assess multiple pieces of discrete information (parameters) generated from each individual cell contained within a large number of cells present in the test sample, and these are typically accrued at rates of 1000–2000 (or more) cells per second.

FLUORESCENCE REAGENTS

Standard mAb reagents for clinical use are typically directly conjugated to a fluorochrome, a dye that absorbs and emits light of different wavelengths based on the energy lost during the return of excited electrons to their ground state associated with illumination by a specific wavelength of light. Thus the emitted light has a longer wavelength (lower energy) than the wavelength of the excitation beam. The number of commercially available fluorochromes has increased dramatically with the routine use of dye conjugates and instruments with three or more lasers.[8] Commonly used fluorochromes in clinical immunophenotyping include the organic dyes fluorescein isothiocyanate (FITC), phycoerythrin (PE), peridin chlorophyll protein (PerCP), and allophycocyanin (APC). Conjugations of PE and APC to cyanines (Cy5, Cy5.5, and Cy7) and Alexa Fluor dyes produce tandem dyes with additional emission spectra, based on energy transfer from one fluorochrome to the second fluorochrome serving as the source of emitted light. This allows for the simultaneous evaluation of 6–8 colors in most current clinical instruments with only two or three lasers.

One recent advance in the field was the development of a new class of inorganic fluorescent semiconductor nanocrystals, named *quantum dots* (QDs).[9] These particles are perfectly suited for polychromatic flow cytometry, as they have broad excitation spectra (525–800 nm) and sharp, discrete emission spectra that varies depending on their core size.[9] This means that QDs of different sizes (and consequently of different colors) can be excited by the same laser source, allowing simpler multiplexing.[9] In addition, QDs have high quantum yield, high molar extinction coefficients, and extraordinary resistance to photodegradation and chemical degradation. These qualities make them perfectly suitable for use in biological studies, including intracellular *in vivo* imaging, fluorescence resonance energy transfer (FRET) analysis, and dynamic imaging of single proteins for longer periods.[9]

Additional dyes are available for functional studies and include calcium-sensitive dyes (*e.g.,* fluo-3), glutathione-sensitive dyes (*e.g.,* monochlorobimane), and hydrogen peroxide (H_2O_2)–responsive dyes (*e.g.,* dihydrorhodamine 123 [DHR123]).[10,11] Assessment of DNA content can be performed with dyes that intercalate double-stranded DNA and RNA, including propidium iodide (PI) and ethidium bromide.[12] In addition, there are ultraviolet-excited dyes that are highly specific for DNA, including Hoechst 33258 and 4,6-diamidino-2-phenylindole (DAPI); acridine orange is used for simultaneous staining of DNA/RNA.[12]

DATA ANALYSIS

Gating

KEY CONCEPTS

Gating

- Method for defining cell population of interest
- Typically performed using forward scatter and side scatter and lineage-specific antibodies

The proper assessment of specific cell types within a mixture requires initial identification of lineage-specific cells, an approach referred to as *gating.* In practical terms, immunophenotyping focused on lymphocytes requires minimizing the nonlymphocytes included in the evaluation, and this is accomplished by lymphocyte gating. The standard sample for clinical studies is anticoagulated whole blood; directing the study to lymphocytes requires eliminating the great majority of nonlymphocytes from the collected data such that the expression of a percentage for a specific cell subpopulation is an accurate measurement. Without gating, data can also be negatively impacted by coexpression of surface antigens on different cell lineages (*e.g.,* CD4 is found on lymphocytes and monocytes at various densities). In addition, nonspecific binding of monoclonal reagents through Fcγ receptors, and the level of cytophilic human immunoglobulin (Ig), varies between cell types, making appropriate gating crucial to generate valid data. These techniques are also used to focus the evaluation on other hematopoietic cells, including monocytes, granulocytes, eosinophils, erythrocytes, and platelets.

Initial gating to focus on a specific leukocyte population typically involves the use of two nonfluorescent parameters, forward scatter (FSC) and side scatter (SSC) (see Fig. 92.2A).[3] FSC is a reflection of cellular cross-sectional area (direct relationship to cell size), whereas SSC is an indication of the cellular granularity. The combination of these two nonfluorescent parameters provides a three-part differential in red blood cell (RBC)–lysed whole blood that distinguishes between normal lymphocytes, monocytes, and granulocytes. As can be seen in Fig. 92.2A, among leukocytes, lymphocytes have the lowest FSC and SSC, monocytes have higher FSC and SSC, and granulocytes have the greatest SSC. This method is effective in distinguishing a relatively pure population of lymphocytes under most circumstances. However, the presence of nucleated RBCs, large platelets, basophils, or other particulate debris can produce contaminating events (cells) within this “lymphocyte gate.” Furthermore, malignant or activated lymphoid cells may not fit into the previously outlined standard light scatter patterns.

A method for confirming the integrity of the light scatter–based lymphocyte gate uses the directly conjugated monoclonal “gating” reagents anti-CD45 and anti-CD14.[13] These two mAbs more accurately identify the three-part differential. Lymphocytes have the highest level of CD45 binding but are negative for CD14; granulocytes have a lower level of CD45 binding and an intermediate level of CD14 expression; and monocytes have high levels of both CD45 and CD14 expression (see Fig. 92.2B). Importantly, nonleukocytes, including erythrocytes and platelets, are negative for these markers. However, malignant leukocytes that have characteristics of early precursor cells often have altered CD45 and/or CD14 expression, which must be recognized when studying hematological malignancies. Gating reagents provide a reliable means of checking the light scatter–based lymphocyte gate for the frequency of nonlymphocytes within the gate as well as the extent of lymphocyte exclusion from the gate. Guidelines for an acceptable degree of contamination within the lymphocyte gate, as well as the level of lymphocyte exclusion, are contained in the US Clinical and Laboratory Standards Institute guideline for lymphocyte immunophenotyping.[14] With the expanded use of polychromatic flow cytometry, some centers now include anti-CD45 in every tube to refine the gate and prevent cell contamination as described above.

Data Display

KEY CONCEPTS

Data Presentation

- Fluorescence intensity is plotted versus cell number.
- Flow cytometry can present cumulative data on more than one parameter for each cell.
- Multicolor data presentation can increase cell subpopulation resolution.

The simplest method for demonstrating flow cytometry data is the single-parameter histogram (see Fig. 92.2), a graphic presentation of cell number on the *y*-axis versus fluorescence (light) intensity from a single fluorochrome on the *x*-axis. Integration of curve areas provides the number of cells, and often there are two distinct distributions, one referred to as *negative* identifies cells that are not bound specifically by the monoclonal reagent, and the second represents cells bound by the antibody. Negative distribution actually reflects low-level fluorescence resulting from cellular autofluorescence together with any nonspecific binding of the monoclonal reagent(s), and the magnitude of both phenomena varies among different cell types. Interpretation of data is simplified when there are two distinct cell populations (*i.e.*, negative and positive), but the evaluation of two overlapping distributions is more difficult.

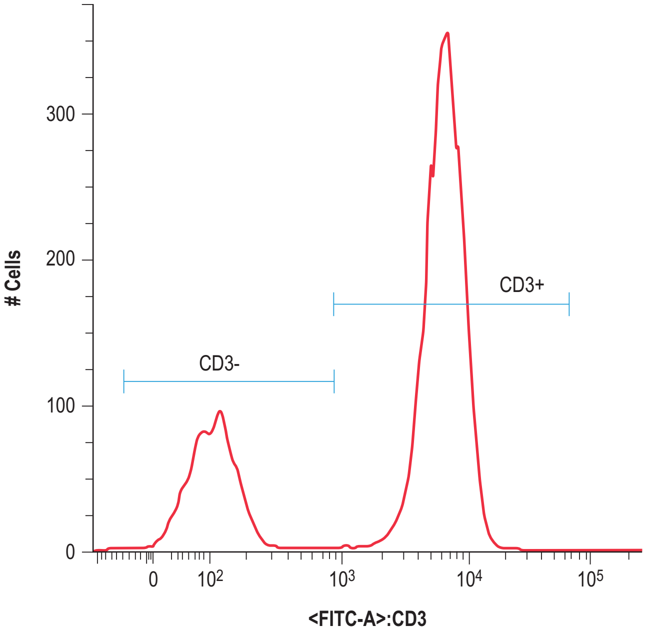


FIG 92.3 Single-parameter histogram for CD3 expression on lymphocytes demonstrating the negative, non–T-cell population (B cells, natural killer [NK] cells), and a positive T-cell population. Integrating the area under each curve would provide the numbers and percentage of cells present in each respective subpopulation.

Multiparameter data can be evaluated by using a series of single-parameter histograms that consider each fluorochrome independently. However, it is more informative to present two parameters simultaneously by using a correlated display (Fig. 92.4), and two-color displays are recommended for clinical flow cytometry.[15] This approach enables the simultaneous visualization of four different populations: A$^+$/B$^-$, A$^-$/B$^+$, A$^+$/B$^+$, and A$^-$/B$^-$. More recently, these displays were developed further to include a mixture of logarithmic (for higher intensity expression) and linear (for lower intensity expression) intensity for each axis to allow for better interpretation of events with very low, zero, or negative fluorescence. This combined-display approach resolves the previous problem of a large number of events being displayed compacted against the axes, even with properly compensated samples, and will be used in the illustrations throughout this chapter.[16]

The simultaneous use of *n* monoclonal reagents can identify a total of 2^n subpopulations. These different subsets can be identified sequentially by first dividing the cells into those that are positive versus those that are negative with one reagent and then evaluating the defined subpopulations for the remaining two reagents using a two-color approach. Alternatively, more modern software can represent multiple populations as polychromatic plots, which can simplify data analysis.[17] The polychromatic approach can provide a means to further resolve subpopulations and has been particularly useful in the evaluation of cellular differentiation, activation, and functional correlates, as well as clarifying overlapping cell subpopulations.

Positive–Negative Discrimination

The evaluation of clinical immunophenotyping data requires establishing criteria for the boundaries between negative or nonstained (autofluorescence plus nonspecifically stained)

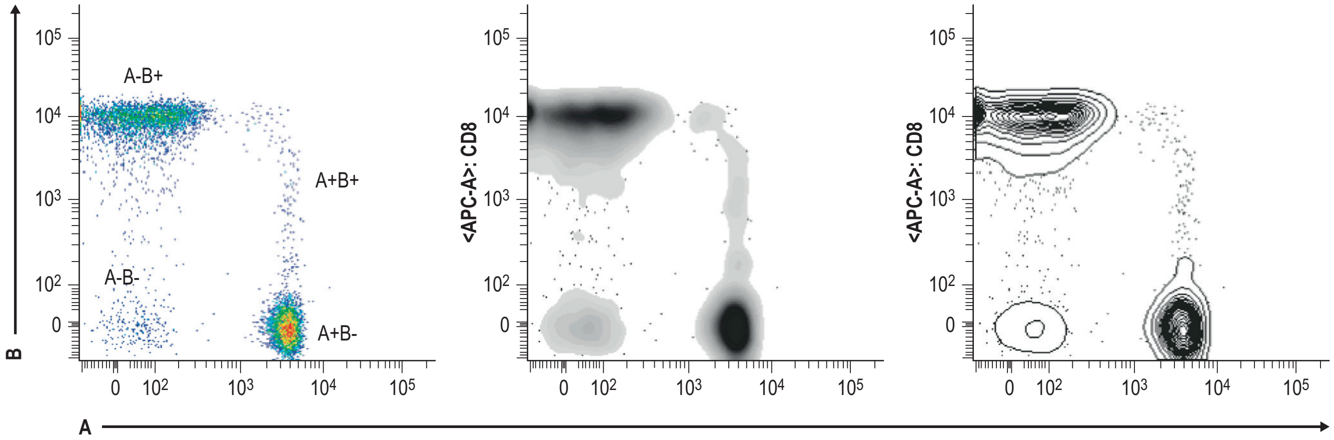


FIG 92.4 Examples of Dot Pseudocolor *(Left),* Density *(Middle),* and Contour *(Right)* Displays Based on the Same Two-Color Parameter Data. All three techniques enable simultaneous evaluation of both parameters, in this case evaluating the expression of markers A and B. These plots identify four populations of cells, those expressing only A or B, those expressing both A and B (very few), and those expressing neither A nor B.

cells and positive (specifically stained) cells. A commonly used approach involves using directly conjugated control mAbs of the appropriate class or subclass (*e.g.*, IgG1, IgG2a, IgG2b, or IgM) that do not specifically react with human lymphocyte surface antigens (commonly referred to as "isotype controls"). The marker (discriminator) is set at the fluorescence histogram channel number such that it includes 98–99% of the negative cells (Fig. 92.5A).

As previously noted, *negative* refers to the aggregate of baseline cellular autofluorescence plus nonspecific reagent binding, and this can vary according to the cell type. For this reason, the use of isotype controls may not correctly identify the positive–negative threshold for specific cell types, particularly when staining dimly expressed proteins. Additionally, to perfectly mimic the specific antibody used, the isotype control would need to have the same antibody to fluorochrome ratio and brightness, something that is not easily accomplished. To overcome these difficulties, an alternative method for positive–negative discrimination, fluorescence-minus-one (FMO), has been developed.[18]

FMO involves the staining of the sample with all the antibodies of interest, except the one targeted for positive–negative threshold. As an example, to define the negative threshold for the protein FAS (CD95) in $CD8^+$ and $CD8^-$ T cells, the FMO control tube would include the cell subset-specific markers (CD3, TCRαβ, and CD8) and omit anti-FAS. After appropriate gating on that population, the threshold can be adequately defined, and it is different for the two exemplified populations (see Fig. 92.5B, *right panel*). One obvious limitation of this method is the higher cost, given that multiple control tubes need to be set up for each sample.

Compensation

The fluorescence signals emitted by different fluorochromes are not completely separated by the filters. This can lead to signal overlap, which is corrected by electronic subtraction of the overlapping signal, a process referred to as *compensation.* The overlap is particularly significant when using multiple fluorochromes, each with different spectral properties.[18] The compensation process involves subtraction of the "spillover" signal detected by the photodetector generated by samples stained with only one fluorochrome. Currently, most flow cytometry analysis software products allow for off-line compensation, where the single reagent stained tubes are used to create a compensation matrix, which is then applied to all the tubes in the experiment. This allows for much simplified compensation procedures, without the need for any hardware compensation during data collection.

Quality Control

Quality control is a critical component of clinical flow cytometry to ensure optimal results.[19] This includes monitoring instrument setup and performance, optimizing sample preparation and reagents, and standardizing controls and data interpretation. Quantitative flow cytometry based on a fluorescence standard curve provides quantitative data in units referred to as *molecular equivalents of soluble fluorochrome* (MESF). When properly constructed standard curves are used, quantitative data for different reagents can be generated and compared. Finally, participation in interlaboratory proficiency testing surveys, such as the triannual samples provided by the College of American Pathologists (CAP), is an important additional measure to monitor laboratory performance, and this is mandated in US clinical laboratories by the Clinical Laboratory Improvement Amendment of 1988 (CLIA 88).

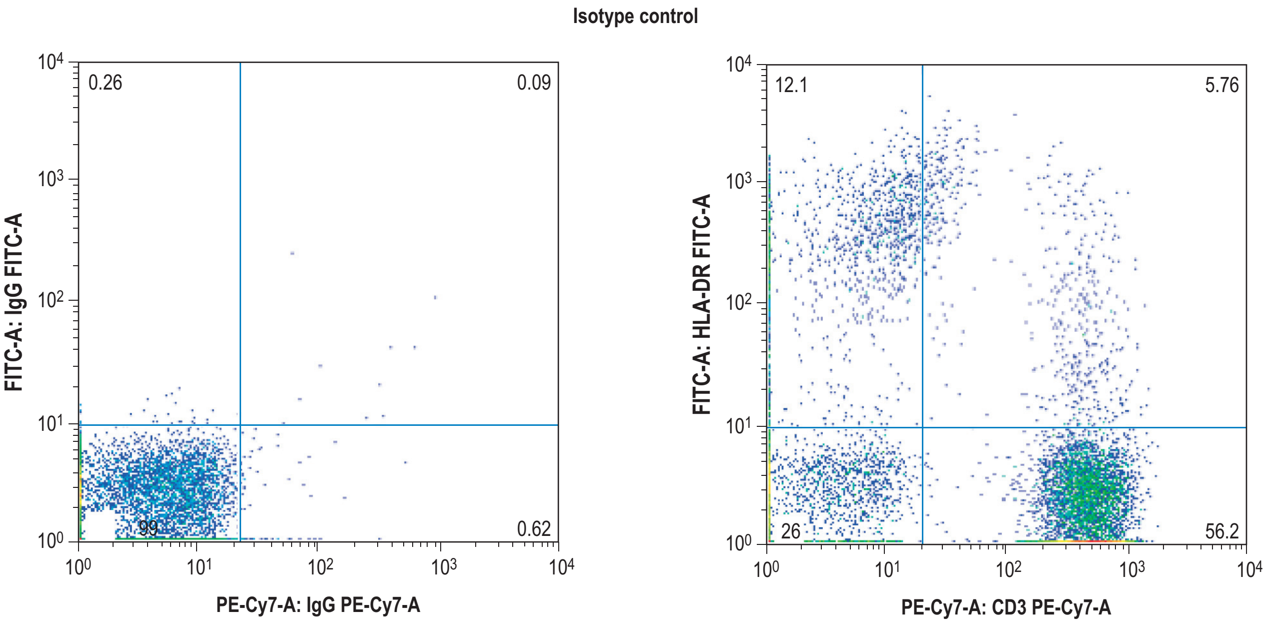
Methods

Whole-blood lysis represents the most common technique used for sample preparation and consists of mixing a fixed volume of anticoagulated whole blood (or bone marrow) with one or more directly conjugated mAbs, followed by incubation at a designated temperature and time.[20] Next, RBCs are lysed, and the sample is washed and then run into the flow cytometer, usually following fixation in paraformaldehyde to reduce the risk of infection. Nonlysed cells that remain include all peripheral blood leukocytes as well as any nonlysed RBCs, platelets, and debris. The heterogeneity of the sample necessitates careful lymphocyte gating (see Gating section above) to generate accurate immunophenotyping data. The advantages of the whole-blood-lysis method include fewer preparation steps, less sample handling, and a lower likelihood of differential lymphocyte loss that can occur when density gradient techniques are used to prepare mononuclear cells for analysis. Alternative sources of cells (*e.g.*, bone marrow, bronchial alveolar lavage fluid, fine-needle aspirates) can also be evaluated with flow cytometry.[21] Patient studies must be determined with the same methods and reagents as used in the determination of the control ranges to ensure comparability. The number of events (cells) analyzed typically ranges from 10 000–20 000 in routine clinical studies but must be increased when evaluating very small subpopulations of cells to produce statistically relevant data necessary in rare-event analysis.

The application of control ranges must take into account the fact that significant changes take place in lymphocyte distribution and development during childhood, as well as changes in lymphocytes that occur among the very old.[22,23] In addition, since immunophenotypic differences can be induced by drugs, including tobacco products, information on current medications should be obtained, whenever possible. Other factors can also have an impact on lymphocyte distribution, including race, gender, diurnal variation, and recent or intercurrent infections.[24]

The choice of immunophenotyping reagents depends on the cells being targeted for study and the question being asked. However, regardless of the specific setup, the inclusion of a tube with gating reagents (anti-CD45 and anti-CD14) to confirm the integrity of the standard lymphocyte gate is recommended.[13] In addition, control reagents should be included to establish the fluorescence intensity of negative cells. Important internal controls include a pan-T-, B-, and natural killer (NK)-cell marker for every sample (Table 92.1), based on the principle that the whole is the sum of its parts. Thus the total percentage of lymphocytes in the gate determined by the gating reagents should approximately equal the sum of the percentages for T, B, and NK cells. A technical or biological explanation must be identified when this relationship does not hold. Biological explanations for a significant difference would include the presence of immature or malignant cells that were not identified by standard pan-T-, B-, and NK-cell reagents. Also, if the gating reagents (CD45/CD14) had not been included, contaminating cells (*e.g.*, myeloid precursors, nucleated RBCs, large platelets) with FSC and SSC characteristics similar to those of lymphocytes could not be ruled out. Potential technical problems include reagent or fluorochrome degradation, failure to add a reagent, and a host of others. Evidence for any major technical error should result in repeating the study.

Additional data that can be used for controls depend on the setup. For example, the availability of multiple antibodies that


Isotype control

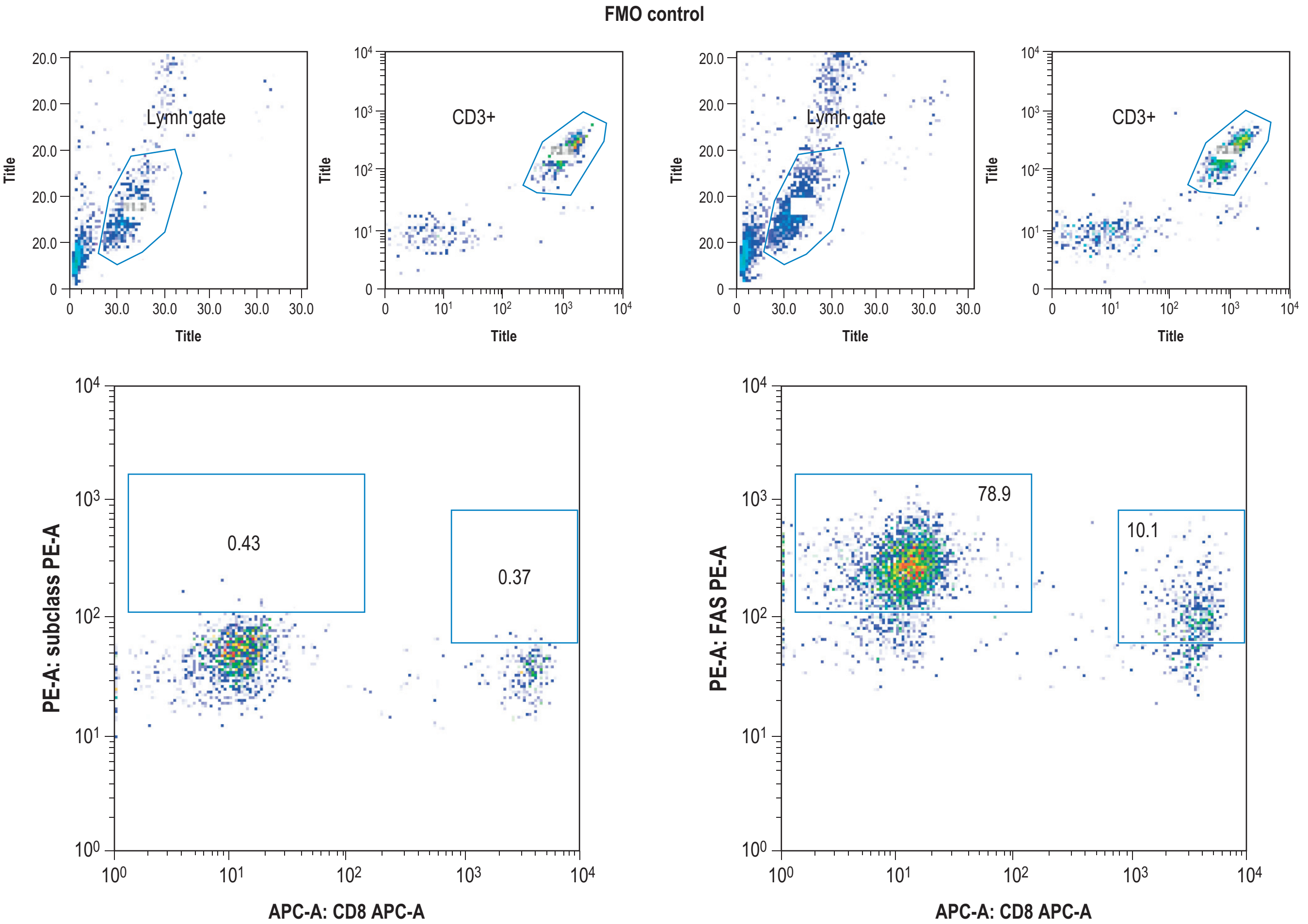

FMO control

FIG 92.5 **Positive–Negative Discrimination Strategies.** (A) Nonspecific antibodies (isotypes) were used to stain the sample shown on left and the positivity threshold determined and applied to the sample on the right. (B) Fluorescence-minus-one (FMO). The positivity threshold was determined by staining the sample on the left with the population-specific markers (CD3, TCRαβ, and CD8) and omitting FAS. The panel on the right contains FAS. Note that the positivity threshold is slightly different for $CD8^+$ and $CD8^-$ cells.

TABLE 92.1 Selected Lymphocyte Surface Antigens for Immunophenotyping

T Cells

Pan-T cell: CD3, CD2, CD7, CD5
Major T-cell subset: CD4, CD8
Surface antigens associated with function: CD28, CD31, CD38, CD45RA, CD45RO, CD62L, CXCR7
Activation antigens: CD25, CD40L, CD69, CD71, HLA-DR

B Cells

Pan-B cell: CD19, CD20, surface immunoglobulin
Major B-cell subset: CD5, CD21
Surface antigens associated with function: CD27, CD40, CD80, CD86
Activation antigens: CD23, CD25

NK Cells

Pan-natural killer (NK) cell: CD16, CD56
NK subset: CD2, CD8, CD57

identify a similar cell subpopulation can serve as a useful check (*e.g.*, total T cells by comparing CD3 and CD5 or CD2; total B cells by comparing CD19 and CD20). In addition, the use of specific reagents in >1 tube enables comparison between the repeat values as a measure of consistency. The application of internal checks should be performed by the flow operator as a simple means of confirming the validity of the data. Insights regarding unusual biological findings may also be uncovered through this type of evaluation (*e.g.*, the presence of an increased population of $CD4^-$/$CD8^-$ double-negative T cells).[25]

The challenge in performing immunophenotyping is to accurately identify cells with specific surface characteristics (antigens). As previously noted, the capacity to discriminate cell subpopulations is often enhanced through the directed use of antibody combinations. The typical data generated consist of the percentage of negative versus positive cells when using one reagent and multiple subpopulations when using more than one reagent. Regardless of the experimental design, it is important to consider not only the percentage of cells within each subpopulation but also the absolute number of cells. This is most commonly obtained by multiplying the relevant percentage from the flow cytometer by the absolute lymphocyte count obtained by using a white blood cell (WBC) count and differential. For example, when assaying for CD4 T-cell counts, the percentage of $CD4^+$ cells is multiplied by the peripheral absolute lymphocyte count to yield the absolute CD4 count. A potential problem with this approach is the requirement for two separate procedures (*i.e.*, dual platforms) to generate the final result. This introduces the possibility of additive error, based on the inherent errors of the two different methods. It also has fueled a search for approaches that facilitate performing both tasks by flow cytometry (*i.e.*, a single platform). One alternative involves the inclusion of a fixed number of fluorescent beads (in a defined volume) in each tube as a reference standard to generate absolute numbers without requiring the use of the complete blood count (CBC) and differential to generate a lymphocyte count. An alternative approach involves the use of impedance-based cell counting in the flow cytometer to generate an absolute lymphocyte count (dependent on a fixed volume of sample being run) and then generating both percentage and absolute number of cells in each specific population or subpopulation. Regardless of the approach, the reporting of both percentages and absolute numbers is necessary when immunophenotyping peripheral lymphocytes.

The objective of evaluating malignant cells is often to characterize the lineage and differentiation level of the abnormal cells, rather than quantifying subpopulations. The pattern of reactivity combined with fluorescence intensity is often useful in identifying leukemic patterns, whereas the absolute number of cells usually is not required. However, flow-cytometric detection and quantitation of rare abnormal cells can be useful in evaluating for minimal residual disease after therapy in lymphoproliferative disorders.

PRACTICAL APPLICATIONS OF FLOW CYTOMETRY

Immunophenotyping Studies

CLINICAL RELEVANCE

Immunophenotyping Studies

- Immunotyping studies can be used to identify cell subsets, lineage, stage of cell differentiation, state of cell activation, and clonality.
- Lymphocyte results should be checked with T cells + B cells + natural killer (NK) cells = 100%.
- Immunophenotyping studies are not the equivalent of lymphocyte function studies.

Most immunophenotyping studies aim to enumerate specific cell subpopulations, evaluate for the presence or absence of particular surface antigens, identify the differentiation level of specific cells, determine cell lineage, evaluate for functional correlates based on specific antigen expression, examine evidence of cell activation, and/or establish evidence of monoclonality.

Quantification of a particular cell subpopulation can be readily accomplished by flow cytometry together with a WBC count and cell differential. The absolute lymphocyte count is used to generate population or subpopulation cell numbers based on the percentages of the respective cell types generated with the flow cytometer. The evaluation of absolute CD4 T-cell counts has formed the basis for monitoring patients with HIV infection.[26] The quantitation of $CD34^+$ hematopoietic stem cells (HSCs) in donor peripheral blood or bone marrow is used in many cellular reconstitution protocols.[27] Subpopulation characterization can also be useful in the evaluation of patients with clinical history and laboratory findings suggestive of immune deficiency.[28] This has taken on a higher level of significance in the setting of newborn screening for severe T-cell immune deficiency via the T-cell receptor excision circle (TREC) assay, which, when abnormal, is typically followed by immunophenotyping to evaluate for naïve T cells based on CD45RA, CD127, or CD62L expression.

When evaluating for the presence or absence of surface proteins associated with specific functional attributes, it is important to realize that this approach does not assess the actual functional status of cells. This point is clearly illustrated by the finding of normal B-cell numbers in most patients with common variable immune deficiency despite the fact that these patients fail to produce Igs normally.[29] However, changes in the characteristics of the B cells, particularly relative to memory B cells, provides potential insight into different phenotypes of this disorder and provides additional support for heterogeneity of patients with the disease.[29-31] Because of the limitations of immunophenotyping, when evaluating the status of the immune system, it is common practice to perform cell function testing in parallel.

Flow cytometry can be used to test for the presence or absence of a specific cell surface antigen. An example of this type of application is in the evaluation of a patient with a history of recurrent skin infections, delayed wound healing, and persistent granulocytosis, which suggests a diagnosis of leukocyte adhesion deficiency type 1 (LAD-1).[28,32] This disorder results from a defect in the gene encoding CD18, preventing the expression of three different heterodimeric adhesion molecules (β_2 integrins) each containing CD18 (Chapter 22). This disorder can usually be diagnosed by studying granulocytes (and lymphocytes) for the expression of CD18 (as well as the three isoforms of CD11). Patients often have decreased rather than absent CD18 expression and confirmation of the diagnosis can be accomplished by demonstrating failure of CD18 (and CD11a, 11b, 11c) upregulation following granulocyte activation.[28]

Immunophenotyping can also help address questions regarding the level of cell differentiation. Antibodies specific to proteins expressed by early (precursor) cells represent one approach and would include evaluating for the thymocyte marker CD1 or the pre-B-cell marker CD10 (CALLA), to cite a few. However, defining the developmental level of a particular cell population or subpopulation is best accomplished by using a panel of reagents that span the natural history of the cell lineage. This approach represents the standard for testing leukemias and lymphomas. In addition to defining the presence or absence of specific antigens, evaluating their level of expression is also valuable, which may be altered in abnormal cells. Malignant cells may also express antigens associated with different lineages and have altered FSC and SSC characteristics, as well as diminished or absent CD45 expression, requiring modified gating approaches.

Issues of monoclonality can be dealt with by using flow cytometry when analyzing B cells and, in some circumstances, when studying T cells. Normally, B cells are a heterogeneous mixture of mutually exclusive κ or λ light-chain-positive cells. Measuring the distribution of κ or λ light-chain-expressing B cells or plasmocytes can be informative with respect to the presence or absence of monoclonality.[33] The capacity to evaluate T-cell monoclonality by flow cytometry is less definitive and consists of using T-cell antigen receptor β-variable (Vβ) chain-specific reagents looking for evidence of significant overrepresentation of one Vβ chain family. This approach currently consists of setting up a series of tubes, each with three different Vβ family–specific mAbs, one conjugated with FITC, one with PE, and the third with FITC plus PE. This combination enables distinguishing the frequency of each of the three different Vβ families per tube (green$^+$, orange$^+$, green$^+$/orange$^+$) and represents a flow-cytometric method to complement polymerase chain reaction (PCR)–based spectratyping.[34]

The state of lymphocyte activation can be addressed by evaluating for the presence of surface antigens that are found only on activated cells or are upregulated following activation. These include receptors for specific growth factors (*e.g.,* interleukin-2 [IL-2] receptor α chain, CD25), receptors for critical elements required for cell growth (*e.g.,* transferrin receptor, CD71), ligands for cell–cell communication following activation (CD40 ligand [CD152] on activated CD4 T cells), and surface antigens that are upregulated as a result of activation (*e.g.,* adhesion molecules, HLA-DR, CD69). In addition, the memory status of both T cells and B cells can be assessed on the basis of differential surface molecule expression associated with prior antigen encounter. This enables a distinction to be made between naïve T cells that express CD45RA, CD31 (recent thymic emigrants), CD62L, and CXCR7 and memory T cells that express the alternative CD45 isoform, CD45RO (and varied CD62L, or CXCR7, depending on whether the cells are central or effector memory cells).[35] In addition, memory B cells can be detected by the expression of CD27 and can be further divided into isotype-switched and isotype-nonswitched memory cells on the basis of their pattern of surface immunoglobulin expression.[29,36]

Defects associated with familial lymphohistiocytosis (FLH) are generally associated with abnormal NK cell function. Many of the FHL-causing defects can be determined by flow cytometry. For example, signaling lymphocyte activation molecule (SLAM) associated protein (SAP) and X-linked inhibitor of apoptosis (XIAP) intracellular staining can be used to evaluate for X-linked lymphoproliferative (XLP) disorders types 1 and 2, respectively. Likewise, lack of intracellular perforin expression in NK cells would be indicative of hemophagocytic lymphohistiocytosis (HLH) type 2. Additionally, the evaluation of CD107a surface expression, which is normally expressed on cytoplasmic granules and upon incubation with specific target cells (*e.g.,* K562 cells) gets expressed on the surface of NK cells, is useful in determining the underlying genetic defect causing FHL.[37,38] Specifically, lack of CD107a upregulation is suggestive of syntaxin-11 or MUNC-13.4 defects.

INTRACELLULAR EVALUATION

Cellular Activation

CLINICAL RELEVANCE

Intracellular Flow Cytometry

- Activation-directed studies:
    - Calcium flux
    - Intracellular protein phosphorylation
    - Oxidative burst: neutrophils
- Intracellular cytokine studies:
    - Clarify the T-helper-1 (Th1)/Th2/Th17 status of an immune response
    - Can be assessed in an *in vitro* antigen-specific response
    - Can be combined with evaluation of cell surface studies

Ligand binding and transmembrane signal transduction resulting in cellular activation can be evaluated by using flow cytometry. Changes in intracellular ionic calcium concentration (Ca^{2+}) can be used to monitor cell activation after ligand binding. These changes are associated with the activation of phospholipase C and protein kinase C. In general, three reagents have been used to measure Ca^{2+}: quin 2, indo-1, and fluo-3. Quin 2 has a low excitation coefficient and is not useful for flow cytometry; indo-1 requires ultraviolet excitation; fluo-3 can be excited by 488 nm but does not permit ratiometric analysis. Nevertheless, because of its ease of use, fluo-3 is currently the most widely used probe for intracytoplasmic Ca^{2+} evaluation by flow cytometry. Strict attention must be paid to loading conditions, the presence or absence of free Ca^{2+} in the medium, experimental temperature, baseline measurements, and calibration. This approach can be combined with cell surface marker or cell cycle evaluation.[10]

Intracellular pH changes related to cellular activation also can be evaluated. The most useful probe for pH is SNARF-1.[10] This probe can be excited at 488 nm and allows for ratiometric analysis with detection wavelengths set for 575 and 640 nm. Glutathione (glutamylcysteinylglycine [GSH]), an important antioxidant generated during cell activation, that can be measured by using flow cytometry.[10] The fluorescent probe monochlorobimane is commonly used for this measurement, but it is complicated by the need to determine GSH by an independent method, such as high-performance liquid chromatography (HPLC).

Additional approaches for evaluation of cellular activation include assessment of intranuclear markers (Ki-67, PCNA) as well as surface proteins that are upregulated following cellular activation (*e.g.,* CD69, CD25, CD71).[39] Actual cell division can be evaluated by using lipophilic membrane dyes (*e.g.,* PKH26, CFSE, Cell Trace Violet), also referred to as *cell tracking dyes,* which lose 50% of their fluorescence with each round of cell division.[40] This approach has become more common in the clinical assessment of lymphocyte function because of the capacity to evaluate specific lymphocyte subpopulations responding to mitogenic and antigenic stimuli. Lipophilic membrane dyes also can be used to label target cells in cell-based cytotoxicity assays.[41] Recently, an approach to evaluate lymphocyte proliferation following cell stimulation has been described by using the modified nucleoside EdU. Detection of DNA synthesis induced by the different activating agents is measured using a copper-catalyzed click chemistry, which results in EdU being covalently bonded to a fluorescent azide.[42] This approach allows for assessment of cell proliferation at the cell population or subpopulation (*e.g.,* CD3, CD4, CD8) level and can be used in association with mitogen and recall-antigen stimulation.

Functional evaluation of cell activation can be accomplished with flow cytometry–directed detection of the generation of phosphorylated intracellular proteins associated with specific activation signals. An example of this is the detection of phosphorylated signal transducer and activator of transcription 1 (STAT1) following interferon-γ (IFN-γ) stimulation of monocytes, which has been found to be equal to or more sensitive than immunoblotting.[43] This type of assay requires fixation and permeabilization to allow entry of the specific reagent and now has been extended to a number of additional intracellular proteins that are phosphorylated following exposure of selected cells to specific stimuli. Currently, a number of intracellular signaling proteins that undergo phosphorylation following a specific activation signal can be assessed with flow cytometry by using commercially available reagents, some of which are offered in kit form.

The assessment of oxidative burst following cell stimulation plays a central role in neutrophil function testing where the hydrogen peroxide-sensitive dye DHR123 is used. This procedure involves loading granulocytes with the dye, stimulating with phorbol myristate acetate (PMA), and evaluating for fluorescence with flow cytometry.[11,44] This test has proven to be extremely accurate in diagnosing patients with chronic granulomatous disease (CGD) and carriers of X-linked CGD.[44] A major advantage is its sensitivity, which allows for detection of one normal cell in a population of 1000 abnormal cells. This makes assessment of oxidative burst a useful tool in monitoring allogeneic granulocyte survival after transfusion into patients with CGD, as well as a means of following donor chimerism in the setting of allogeneic stem cell transplantation. It also provides a method to identify corrected cells following gene therapy in CGD and has utility in predicting outcome.[45,46]

Intracellular Cytokine Detection

Flow cytometry affords a platform to evaluate cytokine production at the single-cell level by using cytokine-specific, directly conjugated mAbs after fixation and permeabilization of cells.[47] This approach allows for simultaneous detection of two or more intracellular cytokines in combination with cell surface markers or other intracellular markers. Important aspects of intracellular cytokine detection include use of a protein transport inhibitor during activation, use of proper controls, and choice of antibodies. As there is little or no spontaneous cytokine production in circulating human lymphocytes, intracellular cytokine detection requires *in vitro* activation. Initial experience was based on supraphysiological stimulation with use of PMA and ionomycin, but antigen-specific activation systems have also proven to be feasible. It should be emphasized that regardless of the activation method, the duration of activation is an important variable, as individual cells reach maximum cytokine production at different times. In addition, different cytokines have different optimal periods of activation. It is recommended that a proper kinetic profile be established for the biological system or clinical condition being studied.[47]

To increase the amount of intracellular cytokines, inhibitors of intracellular protein transport (*e.g.,* monensin or brefeldin) are commonly used, which leads to accumulation of proteins within the cell. Nonspecific binding of antibody reagents is an issue, as permeabilization allows access not only to the cytokine of interest but also to other proteins present in much greater quantities within the cell than on the cell surface. In addition, fixation further increases nonspecific binding, and use of a negative-control sample, which contains an excess of unlabeled or "cold" anticytokine antibody, as well as a subclass-matched or FMO-control sample provides optimal control. When the conjugated anticytokine is added to the negative-control sample, it can only bind to other proteins in a nonspecific manner, thereby providing a measure to discriminate between specific and nonspecific binding. The use of directly conjugated anticytokine antibodies not only simplifies the staining procedure but also provides the best distinction between specific and nonspecific binding. Because the fixation agent may change the native state of certain epitopes, it is also important to use antibodies that recognize antigens after fixation when combining cell surface characterization with intracellular cytokine evaluation.

One of the main applications of intracellular cytokine detection by flow cytometry has been the study and refinement of the T-helper (Th)1/Th2/Th17 paradigms. It has recently become clear that the regulated secretion of cytokines can be used to study the response of individual T cells to both polyclonal stimuli and specific antigens. Measuring antigen-specific T-cell cytokine expression in response to specific antigen offers a useful alternative to the multimer-based approach (discussed below) to quantify the frequency of antigen-specific T cells (Fig. 92.6).[48]

Cell Cycle Analysis

CLINICAL RELEVANCE

Cell Cycle Analysis

- Useful for screening percentage of S phase and aneuploidy
- Can be combined with cell surface studies
- Can be combined with markers of apoptosis

In addition to surface immunophenotyping and cytoplasmic characterization, flow cytometry is also used in cell cycle analysis. PI is the most commonly used fluorochrome because of its optimal linear DNA-binding capacity in a variety of different

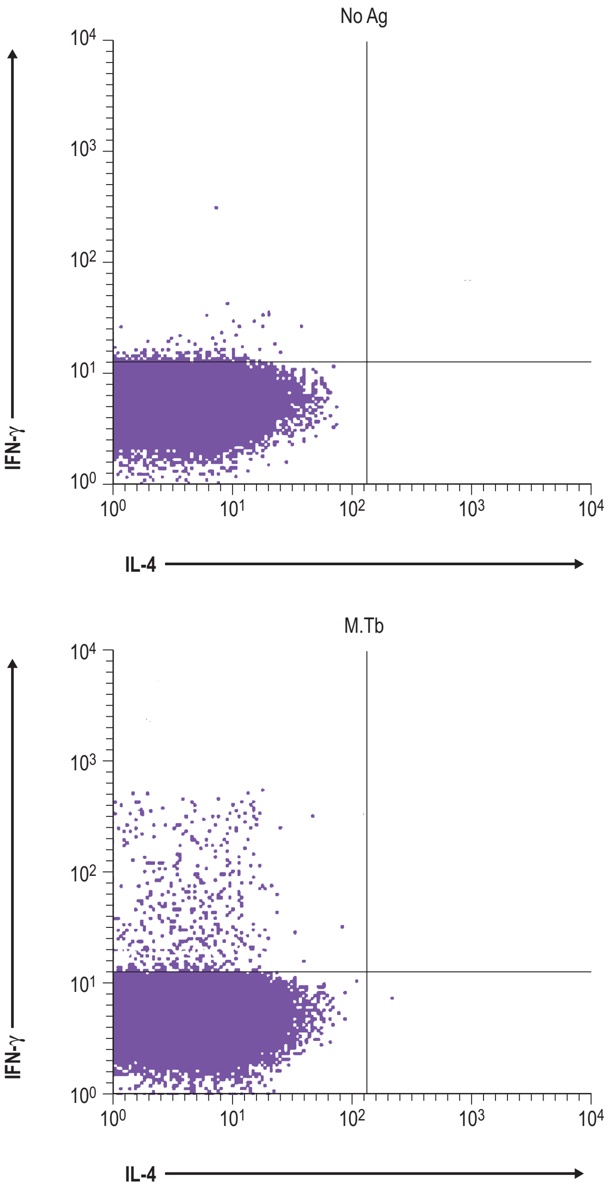


FIG 92.6 Two-color dot plots of CD3$^+$ T cells evaluated for intracytoplasmic interferon (IFN)-γ and interleukin (IL)-4 expression. The donor had a positive skin test to purified protein derivative (PPD) and demonstrated a T-helper-1 (Th1) pattern of cytokine expression (IFN-γ) in response to *Mycobacterium tuberculosis* antigen, with an absence of a Th2 cytokine pattern (IL-4). (Courtesy of Calman Prussin, MD.)

cell types. Thus a single-parameter histogram of DNA content using PI readily permits the determination of cell cycle compartments, expressed as the percentage of cells in G_0–G_1, S, and G_2–M (Fig. 92.7A). In addition to these conventional parameters, the presence or absence of aneuploidy can be determined by inspection of the G_0–G_1 peak and/or use of a DNA index (ratio of abnormal DNA content to a diploid DNA standard). Also, elevation in the S and/or G_2–M phase can be detected. The optimal display of these data uses a combination of SSC versus DNA content. Cells observed on the histogram in the area below the level of G_0–G_1 may be undergoing apoptosis.[49] When dealing with DNA staining, a consistent cellular source of DNA (*e.g.*, chicken erythrocytes) should be used as an internal reference for evaluating DNA content and evaluating the cell cycle distribution.

It should be noted that several different computer algorithms have been developed to determine the relative proportion of each cell cycle compartment, and the selection of a software program is not a trivial process. The major instrument manufacturers supply cell cycle analysis programs, and third-party programs are also available. Generally, the optimal program should be capable of modeling two or more aneuploid populations, subtracting debris (particularly if formalin-fixed, paraffin-embedded [FFPE] archival material is used) and accurately estimating S-phase cells.[49,50] The combination of a surface marker and the cell cycle has been very useful in differentiating normal cell populations from tumor cell populations. One example is the use of anti-κ-, anti-λ-, or B-cell reagents to separate the aneuploid B-cell clone from the remaining normal, reactive B cells in a lymphoid cell mixture. Another uses cytokeratin as a marker to distinguish between tumor cells and inflammatory cells.

The other major event that has occurred in cell cycle analysis has been the development of technology using the incorporation of bromodeoxyuridine.[51] This thymidine analogue is used to directly determine the percentage of S-phase cells. Additionally, when used in kinetic studies, it permits determination of individual times for the components of the cell cycle and determination of the growth fraction. Finally, recent developments have resulted in the availability of two anticyclin reagents to evaluate cell cycle transition points in malignant cells.[52]

APOPTOSIS DETECTION

Flow cytometry has become the method of choice for the detection and quantification of cellular apoptosis,[53] in part because of its capacity for rapid assessment of a large number of cells and samples. Many distinct features of an apoptotic cell can be evaluated by using flow cytometry based on light scatter, plasma membrane changes, mitochondrial transmembrane potential, DNA content, and DNA integrity.

The light scattering properties of a cell undergoing programmed cell death are the simplest attributes that can be assessed by flow cytometry. Dying cells typically shrink, producing a loss in FSC and, despite an initial transient increase in SSC, ultimately demonstrate a decrease in SSC (see Fig. 92.7B). The use of light scatter can be combined with cell surface staining to help characterize the dying cells. However, scatter changes alone are not specific to apoptosis and should be accompanied by an additional characteristic associated with cell death. Live cells have phospholipids asymmetrically distributed in inner and outer plasma membranes, with phosphatidylcholine and sphingomyelin on the outer surface and phosphatidylserine (PS) on the inner side. Early during apoptosis cells lose asymmetry, exposing PS on the outside. Annexin V is a protein that binds preferentially to negatively charged phospholipids, such as PS, and directly conjugated annexin V is a useful reagent for the specific detection of apoptotic cells.[53]

Another characteristic of plasma membranes associated with live cells is that they exclude charged cationic dyes, such as PI and 7-amino-actinomycin-D (7-AAD). Consequently, only cells in a late stage of apoptosis, with ruptured cell membranes, will take up these dyes. Thus the combined use of cationic dyes (*e.g.*, PI) with annexin V allows for discrimination between live cells

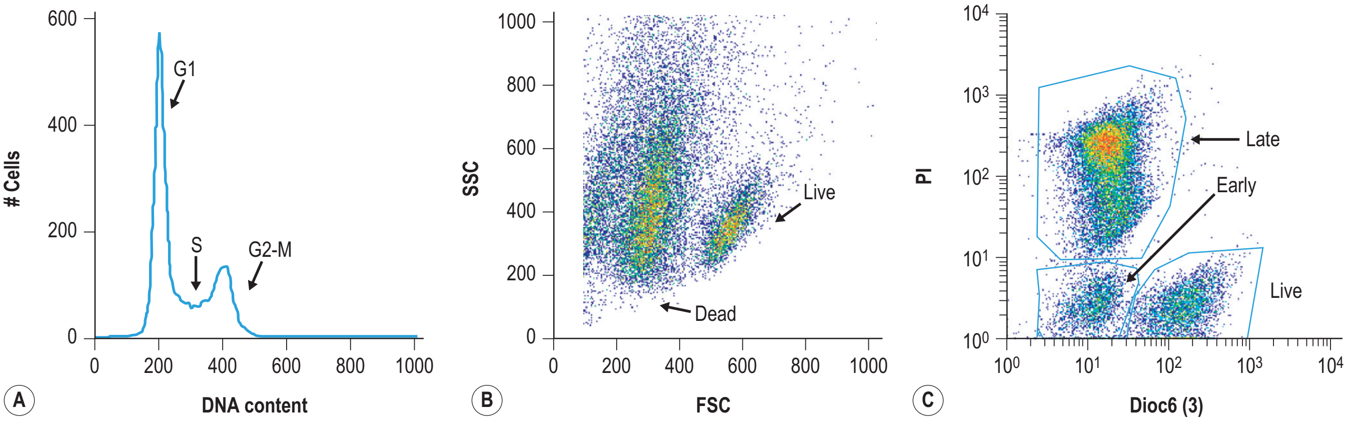


FIG 92.7 (A) Assessment of DNA content as a reflection of cell cycle demonstrating cells in G_1, S, and G_2–M phases. (B) Assessment of live versus dead cells based on forward-(FSC) and side-scatter characteristics (SSC). (C) Assessment of cell apoptosis using propidium iodide (PI) and Dioc6 (3) identifying cells that recently initiated apoptosis (early), cells that are dead (late) and cells that are alive (live).

(annexin V negative/ PI negative), early apoptotic cells (annexin V positive/ PI negative), and late apoptotic cells (annexin V positive/ PI positive).

Assessment of mitochondrial transmembrane potential ($\Delta\psi m$) is yet another technique used to identify apoptotic cells. Cells decrease $\Delta\psi m$ very early in the apoptotic process, before rupture of the plasma membrane, losing the ability to accumulate potential-dependent dyes, such as rhodamine 123, JC-1, or 3,3''-dihexyloxacarbocyanine iodide (Dioc6[3]). These dyes can also be used with PI to detect cells in the different stages of apoptosis (see Fig. 92.7C).

Measurement of DNA content can also be employed to distinguish live cells from dead cells, as described above (see Cell cycle analysis). This kind of analysis has to be done by using a linear scale, not a logarithmic scale, to discriminate dying cells from debris. DNA cleavage also exposes –OH termini associated with the DNA breaks, and these can be detected via the attachment of fluorochrome-conjugated deoxynucleotides, in a reaction catalyzed by exogenous TdT, a technique called *TUNEL*.

PEPTIDE–MHC MULTIMERS

KEY CONCEPTS

Peptide–MHC Tetramers

- Useful for assessing the number of antigen-specific T cells
- Can be directed at both $CD4^+$ and $CD8^+$ T cells
- Requires information about the antigenic peptide and human leukocyte antigen (HLA) (major histocompatibility complex [MHC]) restriction

In contrast to B cells, direct visualization of antigen-specific T cells *ex vivo* has, until recently, been unsuccessful. In 1996, Altman et al. introduced a novel flow cytometry–based methodology that enables the direct visualization and quantification of antigen-specific T cells.[54] Soluble peptide–MHC multimers are generated such that multiple T-cell receptors (TCRs) are engaged at the same time, which means that the avidity of these multimeric ligands for the peptide-specific TCR is greatly increased. The methodology involves engineering a biotinylation recognition sequence on the –COOH terminus of the extracellular domain of one chain of the MHC molecule, which, after combining with a specific antigenic peptide, is bound by avidin or streptavidin. As both avidin and streptavidin have four biotin-binding sites, the result is a tetrameric peptide–MHC complex that serves as a ligand for T cells specific for both the peptide and the MHC. Flow-cytometric detection is achieved by labeling streptavidin with a fluorochrome. The major pitfall of this approach is the need to know the antigen-derived peptide and its HLA restrictions, as well as the HLA type of each subject studied. Since the initial report, an increasing number of tetramer-based studies have been performed. Most have focused on the MHC class I–mediated immune response, in both mice and humans, to a variety of infectious agents, including cytomegalovirus (CMV), HIV, Epstein-Barr virus (EBV), and others. Since the initial description with class I–restricted recognition, detection of antigen-specific CD4 T cells with tetramers of soluble MHC class II molecules and covalently linked peptide has also been reported.[55]

In addition to demonstrating the feasibility of this approach, published studies have provided several new insights into the MHC class I–mediated immune response (Chapter 5). For example, it has become clear that the extent of the MHC class I–mediated cellular response is much greater than previously estimated. Furthermore, the extensive proliferation of CD8 T cells during an acute infection is not the result of bystander activation but represents an expansion of antigen-specific CD8 T cells. Peptide–MHC tetramer assays have shown promise in the study of the kinetics of primary and secondary immune responses and have resulted in better understanding of such concepts as immunodominance and clonal exhaustion.

An obviously attractive aspect of this technology is that tetramer staining can be combined with a variety of cell surface and intracellular phenotypic and functional markers. Indications that the phenotype of antigen-specific T cells varies among individuals and among different phases of the immune response are already present. In addition, tetramer-positive T cells can be sorted for further analysis, such as cytotoxicity assays or *in vitro* expansion. Tetramer-based technology has not only proven useful for the study of the immune response to infectious agents, it has also been applied to the study of oral tolerance, autoimmune conditions, and tumor immunology. It is likely that this highly sensitive and specific technology and other approaches that define the antigen-specific response will find many more applications and will lead to new discoveries and a reassessment of certain existing concepts.[56]

CONCLUSIONS

ON THE HORIZON

- The combination of flow cytometry and mass spectrometry expands the number of cell markers analyzed simultaneously (≥35 currently performed), including extracellular and intracellular targets to provide an extensive evaluation of functional markers.
- Flow cytometry utilizing spectral rather than conventional light analysis has the potential to allow more "colors" to be accurately evaluated without the need for complex compensation schemes to handle the overlapping light emissions from conventional fluorochromes. This could facilitate more extensive polychromatic studies to be performed easily in the clinical laboratory.

Flow cytometry has become readily available in clinical laboratories, and the application of this technology has moved forward along with significant improvements in instrumentation and the availability of an array of monoclonal reagents. Properly performed, flow cytometry can provide rapid and accurate lymphocyte subpopulation identification. The primary clinical indications of immunophenotyping remain quantifying CD4 T-cell counts in HIV infection, lineage assignment in leukemias and lymphomas, evaluation of other hematological cell types, and assessment of CD34 expression to identify stem cells for transplantation. Additional uses include characterization of immune deficiency disorders, evaluation of immune-mediated inflammatory diseases, and assessment of patients after organ transplantation. Generally, immunophenotyping does not represent a diagnostic procedure but, rather, plays a part in the evaluation and understanding of complex disorders and the longitudinal evaluation of immunomodulatory therapy.

It is critical to recognize that immunophenotyping is a means of identifying cells but is not directed at cell function. The expansion of flow-cytometric techniques to evaluate intracellular characteristics, assess intracellular changes associated with activation, characterize apoptosis, and identify antigen-specific T cells has moved this technology into the cell function arena. These newer approaches expand the utility of flow cytometry as a valuable method for the characterization of various aspects of immune function.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Shapiro HM. How flow cytometers work. 4th ed. Hoboken, NJ: Wiley; 2003.
2. Kachel V. Hydrodynamic properties of flow cytometry instruments. New York: Wiley Liss; 1990.
3. Thompson JM, Gralow JR, Levy R, et al. The optimal application of forward and ninety-degree light scatter in flow cytometry for the gating of mononuclear cells. Cytometry 1985;6:401–6.
4. Goddard G, Martin JC, Graves SW, et al. Ultrasonic particle-concentration for sheathless focusing of particles for analysis in a flow cytometer. Cytometry A 2006;69:66–74.
5. Ward M, Turner P, DeJohn M, et al. Fundamentals of acoustic cytometry. Curr Protoc Cytom 2009;49:1.22.21–21.22.12.
6. Chattopadhyay PK, Hogerkorp CM, Roederer M. A chromatic explosion: the development and future of multiparameter flow cytometry. Immunology 2008;125:441–9.
7. Perfetto SP, Chattopadhyay PK, Roederer M. Seventeen-colour flow cytometry: unravelling the immune system. Nat Rev Immunol 2004;4:648–55.
8. Giepmans BN, Adams SR, Ellisman MH, et al. The fluorescent toolbox for assessing protein location and function. Science 2006;312:217–24.
9. Chattopadhyay PK, Yu J, Roederer M. Application of quantum dots to multicolor flow cytometry. Methods Mol Biol 2007;374:175–84.
10. Rabinovitch PS, June CH, Kavanagh TJ. Introduction to functional cell assays. Ann N Y Acad Sci 1993;677:252–64.
11. Vowells SJ, Sekhsaria S, Malech HL, et al. Flow cytometric analysis of the granulocyte respiratory burst: a comparison study of fluorescent probes. J Immunol Methods 1995;178:89–97.
12. Darzynkiewicz Z, Halicka HD, Zhao H. Analysis of cellular DNA content by flow and laser scanning cytometry. Adv Exp Med Biol 2010;676:137–47.
13. Loken MR, Brosnan JM, Bach BA, et al. Establishing optimal lymphocyte gates for immunophenotyping by flow cytometry. Cytometry 1990;11:453–9.
14. US Clinical and Laboratory Standards Institute. Clinical applications of flow cytometry: quality assurance and immunophenotyping of peripheral blood lymphocytes. Villanova, PA: NCCLS Publication H24-T; 1992.
15. Stewart CC, Stewart SJ. Multiparameter data acquisition and analysis of leukocytes. Methods Mol Biol 2004;263:45–66.
16. Parks DR, Roederer M, Moore WA. A new "Logicle" display method avoids deceptive effects of logarithmic scaling for low signals and compensated data. Cytometry A 2006;69:541–51.
17. Roederer M, Moody MA. Polychromatic plots: graphical display of multidimensional data. Cytometry A 2008;73:868–74.
18. Roederer M. Spectral compensation for flow cytometry: visualization artifacts, limitations, and caveats. Cytometry 2001;45:194–205.
19. Owens MA, Vall HG, Hurley AA, et al. Validation and quality control of immunophenotyping in clinical flow cytometry. J Immunol Methods 2000;243:33–50.
20. Schwartz A, Marti GE, Poon R, et al. Standardizing flow cytometry: a classification system of fluorescence standards used for flow cytometry. Cytometry 1998;33:106–14.
21. Yamada M, Tamura N, Shirai T, et al. Flow cytometric analysis of lymphocyte subsets in the bronchoalveolar lavage fluid and peripheral blood of healthy volunteers. Scand J Immunol 1986;24:559–65.
22. Reichert T, DeBruyere M, Deneys V, et al. Lymphocyte subset reference ranges in adult Caucasians. Clin Immunol Immunopathol 1991;60:190–208.
23. Shearer WT, Rosenblatt HM, Gelman RS, et al. Lymphocyte subsets in healthy children from birth through 18 years of age: the Pediatric AIDS Clinical Trials Group P1009 study. J Allergy Clin Immunol 2003;112:973–80.
24. Levi FA, Canon C, Touitou Y, et al. Seasonal modulation of the circadian time structure of circulating T and natural killer lymphocyte subsets from healthy subjects. J Clin Invest 1988;81:407–13.
25. Fleisher TA, Oliveira JB. Autoimmune lymphoproliferative syndrome. Isr Med Assoc J 2005;7:758–61.
26. Yarchoan R, Venzon DJ, Pluda JM, et al. CD4 count and the risk for death in patients infected with HIV receiving antiretroviral therapy. Ann Intern Med 1991;115:184–9.
27. Sekhsaria S, Fleisher TA, Vowells S, et al. Granulocyte colony-stimulating factor recruitment of CD34+ progenitors to peripheral blood: impaired mobilization in chronic granulomatous disease and adenosine deaminase—deficient severe combined immunodeficiency disease patients. Blood 1996;88:1104–12.
28. Oliveira JB, Notarangelo LD, Fleisher TA. Applications of flow cytometry for the study of primary immune deficiencies. Curr Opin Allergy Clin Immunol 2008;8:499–509.

29. Wehr C, Kivioja T, Schmitt C, et al. The EUROclass trial: defining subgroups in common variable immunodeficiency. Blood 2008;111:77–85.
30. Rosel AL, Scheibenbogen C, Schliesser U, et al. Classification of common variable immunodeficiencies using flow cytometry and a memory B-cell functional assay. J Allergy Clin Immunol 2014;135:198–208.
31. Bonilla FA, Barlan I, Chapel H, et al. International consensus document (ICON): common variable immunodeficiency disorders. J Allergy Clin Immunol Pract 2016;4:38–59.
32. Etzioni A. Defects in the leukocyte adhesion cascade. Clin Rev Allergy Immunol 2010;38:54–60.
33. Berliner N, Ault KA, Martin P, et al. Detection of clonal excess in lymphoproliferative disease by kappa/lambda analysis: correlation with immunoglobulin gene DNA rearrangement. Blood 1986;67:80–5.
34. Pilch H, Hohn H, Freitag K, et al. Improved assessment of T-cell receptor (TCR) VB repertoire in clinical specimens: combination of TCR- CDR3 spectratyping with flow cytometry-based TCR VB frequency analysis. Clin Diagn Lab Immunol 2002;9:257–66.
35. Sallusto F, Geginat J, Lanzavecchia A. Central memory and effector memory T cell subsets: function, generation, and maintenance. Annu Rev Immunol 2004;22:745–63.
36. Avery DT, Ellyard JI, Mackay F, et al. Increased expression of CD27 on activated human memory B cells correlates with their commitment to the plasma cell lineage. J Immunol 2005;174:4034–42.
37. Alter G, Malenfant JM, Altfeld M. CD107a as a functional marker for the identification of natural killer activity. J Immunol Methods 2004;294:15–22.
38. Marcenaro S, Gall F, Martini S, et al. Analysis of natural killer-cell function in familial hemophagocytic lymphohistiocytosis (FHL): defective CD107a surface expression heralds Munc13-4 defect and discriminates between genetic subtypes of the disease. Blood 2006;108:2316–23.
39. Mardiney M 3rd, Brown MR, Fleisher TA. Measurement of T-cell CD69 expression: a rapid and efficient means to assess mitogen- or antigen-induced proliferative capacity in normals. Cytometry 1996;26:305–10.
40. Allsopp CE, Langhorne J. Assessing antigen-specific proliferation and cytokine responses using flow cytometry. Methods Mol Med 2002;72:409–21.
41. Slezak SE, Horan PK. Cell-mediated cytotoxicity. A highly sensitive and informative flow cytometric assay. J Immunol Methods 1989;117:205–14.
42. Yu Y, Arora A, Min W, et al. EdU incorporation is an alternative non-radioactive assay to [(3)H]thymidine uptake for in vitro measurement of mice T-cell proliferations. J Immunol Methods 2009;350:29–35.
43. Fleisher TA, Dorman SE, Anderson JA, et al. Detection of intracellular phosphorylated STAT-1 by flow cytometry. Clin Immunol 1999;90:425–30.
44. Vowells SJ, Fleisher TA, Sekhsaria S, et al. Genotype-dependent variability in flow cytometric evaluation of reduced nicotinamide adenine dinucleotide phosphate oxidase function in patients with chronic granulomatous disease. J Pediatr 1996;128:104–7.
45. Heim KF, Fleisher TA, Stroncek DF, et al. The relationship between alloimmunization and posttransfusion granulocyte survival: experience in a chronic granulomatous disease cohort. Transfusion 2001;51: 1154–62.
46. Kuhns DB, Alvord WG, Heller T, et al. Residual NADPH oxidase and survival in chronic granulomatous disease. N Engl J Med 2010;363:2600–10.
47. Foster B, Prussin C, Liu F, et al. Detection of intracellular cytokines by flow cytometry. Curr Protoc Immunol 2007;Chapter 6:Unit 6.24.
48. Suni MA, Picker LJ, Maino VC. Detection of antigen-specific T cell cytokine expression in whole blood by flow cytometry. J Immunol Methods 1998;212:89–98.
49. Shankey TV, Rabinovitch PS, Bagwell B, et al. Guidelines for implementation of clinical DNA cytometry. International Society for Analytical Cytology. Cytometry 1993;14:472–7.
50. Braylan RC, Benson NA, Nourse V, et al. Correlated analysis of cellular DNA, membrane antigens and light scatter of human lymphoid cells. Cytometry 1982;2:337–43.
51. Rothaeusler K, Baumgarth N. Assessment of cell proliferation by 5-bromodeoxyuridine (BrdU) labeling for multicolor flow cytometry. Curr Protoc Cytom 2007;Chapter 7:Unit 7.31.
52. Gong J, Bhatia U, Traganos F, et al. Expression of cyclins A, D2 and D3 in individual normal mitogen stimulated lymphocytes and in MOLT-4 leukemic cells analyzed by multiparameter flow cytometry. Leukemia 1995;9:893–9.
53. Telford WG, Komoriya A, Packard BZ, et al. Multiparametric analysis of apoptosis by flow cytometry. Methods Mol Biol 2011;699:203–27.
54. Altman JD, Moss PA, Goulder PJ, et al. Phenotypic analysis of antigen-specific T lymphocytes. Science 1996;274:94–6.
55. Mallone R, Nepom GT. MHC Class II tetramers and the pursuit of antigen-specific T cells: define, deviate, delete. Clin Immunol 2004;110:232–42.
56. Kern F, LiPira G, Gratama JW, et al. Measuring Ag-specific immune responses: understanding immunopathogenesis and improving diagnostics in infectious disease, autoimmunity and cancer. Trends Immunol 2005;26:477–84.

MULTIPLE-CHOICE QUESTIONS

1. In flow cytometry compensation refers to a process that:
   A. Distinguishes between leukocyte populations
   B. Eliminates overlapping fluorescent light signals
   C. Provides the level of the cellular autofluorescence
   D. Converts analog light signals into digital signals

2. A flow cytometry setup with a combination of five different monoclonal reagents, each with a different fluorochrome, generates how many potential cell subsets?
   A. 8
   B. 16
   C. 32
   D. 64

3. The finding of a CD3 cell count of 285/mm^3 and no CD4/CD45RA cells in an infant is most consistent with which diagnosis?
   A. Dedicator of cytokinesis 8 (DOCK8) deficiency
   B. CD40 ligand deficiency
   C. Severe combined immunodeficiency
   D. Wiskott-Aldrich syndrome

4. Which test would be indicated in a child with persistent systemic nontuberculous infection not responsive to multiagent antimicrobial drugs?
   A. Dihydrorhodamine (DHR) assay
   B. Expression of CD18 on granulocytes
   C. Level of T-helper-1 (Th1) cell–based on intracellular interferon-γ (IFN-γ) expression
   D. Expression of the IFN-γ receptor type 1 on monocytes

93

Assessment of Functional Immune Responses in Lymphocytes

Roshini Sarah Abraham

The immune response is mediated by a complex network of cells, soluble and membrane-bound biological mediators and receptors interacting within the context of specific tissues and organs to protect against pathogens. Although the immune system has been broadly divided into innate and adaptive components, there is considerable overlap and interaction between the two, despite the highly specialized functions and kinetics of each in the immune response. T and B cells form the pillar of the adaptive immune response, whereas natural killer (NK) cells are the effector lymphocytes of the innate immune response. There are approximately 2×10^{12} lymphocytes in the body, and typically the response to a foreign antigen is in the context of initial activation of innate immune components. T cells recognize antigen primarily in the context of antigen-specific T-cell receptors (TCRs) and specific peptides presented by molecules of the major histocompatibility complex (MHC), either class I or class II (Chapters 5 and 6). T cells are also capable of responding nonspecifically to polyclonal stimulators, such as mitogens (*in vitro*) or superantigens (*in vivo*) that do not initiate an antigen-specific proliferative response. T cells upon activation also produce cytokines that are crucial to their effector functions including activating B cells and inducing antibody production, promoting differentiation of cytotoxic T cells, activating macrophages, and promoting activation and migration of inflammatory cells. Cytokines play a role in the induction of the T-cell response when produced by antigen-presenting cells (APCs) at the time of antigen recognition. B cells recognize antigen directly via the B-cell receptor (involving surface immunoglobulin) and produce antibodies with help from T cells (T-dependent antigens) or from the innate immune system (T-independent antigens). More recently, the granularity of the antibody response has been further delineated, and T-dependent antibody responses are divided into type 1, with help provided by T-follicular helper cells (Tfhs), and type 2, with help provided by NKTfh cells. Similarly, T-independent responses are classified into three groups: type 1, induced by recognition of microbial antigens by the Toll-like receptors (TLRs) in the absence of Bruton's tyrosine kinase (BTK); type 2, which requires the presence of BTK; and type 3, which has neutrophil B-helper cells.[1,2]

Regulatory (Treg) cells have been shown to play a critical role in the control of both physiological and pathological immune responses in a variety of contexts (Chapter 18). Treg cells exert a direct inhibitory effect on the development of autoimmune disease because their absence leads to the development of a severe phenotype, as manifested by the FOXP3 deficiency, IPEX (immune dysfunction/polyendocrinopathy/enteropathy/X-linked)[3] (Chapter 35). Treg cells can suppress the proliferation of antigen-stimulated naïve T cells, as demonstrated by *in vitro* studies, and induction of FOXP3 in conventional naïve T cells imparts suppressive function *in vitro* and *in vivo,* resulting in the generation of induced-regulatory (iTreg) T cells.[4] Besides producing antibodies to neutralize pathogens as well as presenting antigen to T cells, B cells also exert immunomodulatory control of the immune response, primarily via interleukin (IL)-10 production. IL-10–producing B cells, classified as regulatory B (Breg) cells, play an important role in immune homeostasis and in protection against autoimmune responses and inflammatory damage. Breg cells, in addition to producing IL-10, secrete other cytokines that act on other effector T-cell subsets, Tregs, APCs such as dendritic cells (DCs), and macrophages.

NK cells are considered to be innate immune effector cells, but they straddle the threshold of innate and adaptive immunity. NK cells are directly involved in cytotoxicity and cytokine secretion upon activation, but they also function indirectly by regulating APC and the effector T-cells' response. NK cell activation is controlled by synergistic signals from activating and inhibitory receptors.[5] The characteristics, responses, and assays to measure the activity and function of each of these subsets are described in detail in the sections below.

T-CELL RESPONSE

T cells form the cellular arm of the adaptive immune response, and each T cell has unique specificity derived from the presence of a functional, antigen-recognizing receptor on the cell surface. Naïve T cells derived from the thymus can migrate through the peripheral circulation and secondary lymphoid tissues (spleen and lymph nodes), but they cannot effectively participate in the immune response to pathogens. To do so, naïve T cells must undergo a defined process of cellular activation, which involves recognition of antigen by the TCR in the context of MHC molecules. The majority of circulating T cells express the α/β TCR, whereas a smaller proportion of cells express the γ/δ TCR. On the surface of the T cell, the TCR associates with the CD3 complex, comprised of four distinct subunits (γ, δ, ε, and ζ); the cytosolic components of the CD3 complex are involved in the intracellular propagation of signals after ligation of the TCR. Besides the CD3 complex, the TCR also clusters with a CD4 or CD8 coreceptor, depending on the type of T cell; CD4 T cells recognize antigen (Ag) in the context of MHC class II, whereas CD8 T cells recognize antigen presented on MHC class I. The first signal or "cognate" signal, which is recognition of peptide-MHC complex by the TCR, results in actin-mediated reorganization of the cytoskeleton in both the T cell and APC to form the immunological synapse (Fig. 93.1). The synapse consists of the

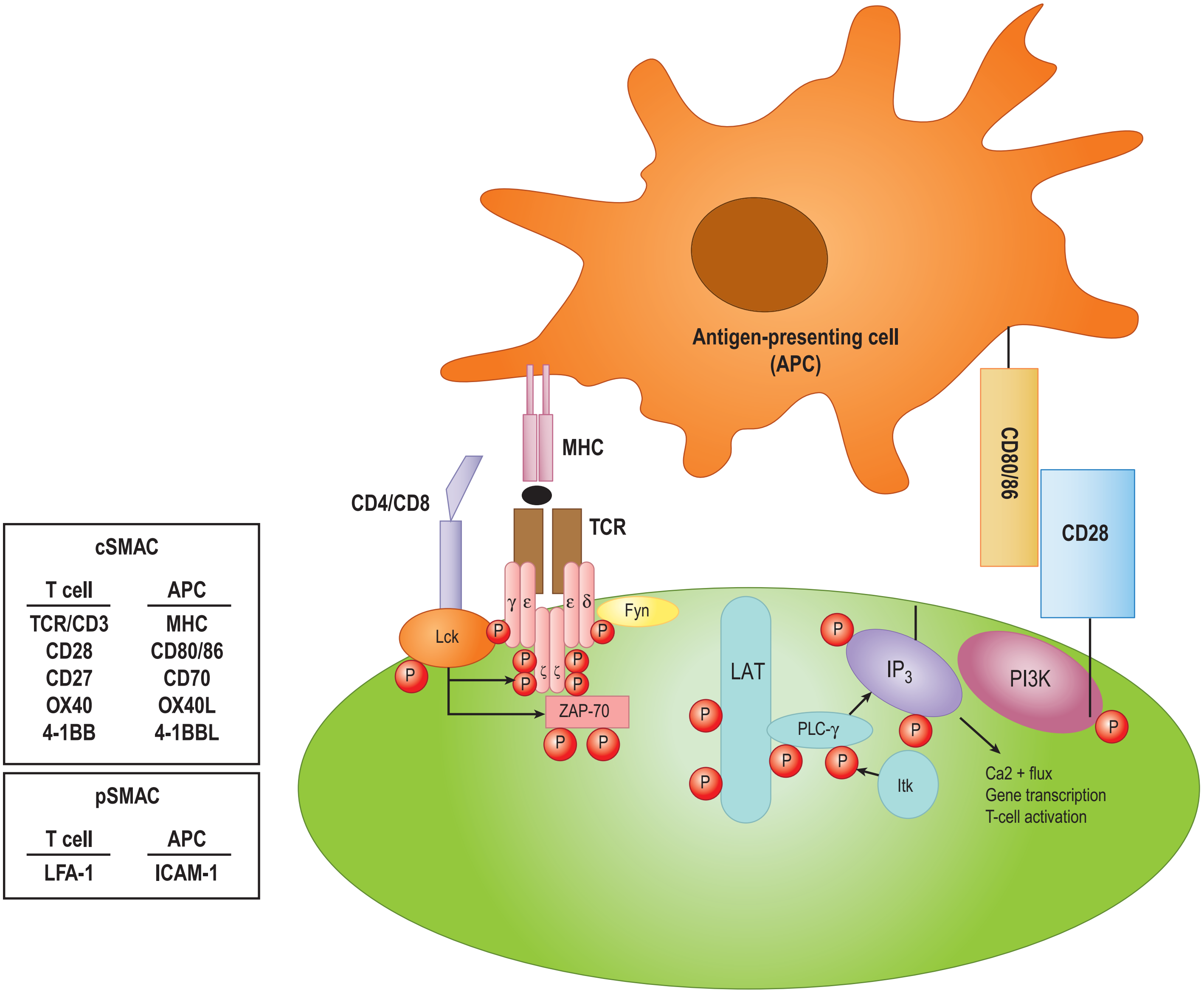

Antigen-presenting cell
(APC)
MHC
CD4/CD8
TCR
CD80/86
CD28
γ ε
ε δ
Fyn
Lck
P
ζ ζ
ZAP-70
LAT
PLC-γ
IP_3
PI3K
Itk
Ca2 + flux
Gene transcription
T-cell activation
cSMAC
T cell
APC
TCR/CD3
MHC
CD28
CD80/86
CD27
CD70
OX40
OX40L
4-1BB
4-1BBL
pSMAC
T cell
APC
LFA-1
ICAM-1

FIG 93.1 Schematic of T-Cell Activation. T cells recognize antigen in the context of major histocompatibility complex (MHC) molecules on antigen-presenting cells (APCs). Sustained interaction between a T cell and APC results in the formation of the immunological synapse with membrane reorganization and ordering of key molecules on both cells into a supramolecular activation complex (SMAC) comprised of c (central) SMAC and p (peripheral) SMAC. Cognate interaction with appropriate costimulatory signals results in intracellular signaling events that ultimately result in T-cell activation. LAT, linker for activation of T cells; PLC-γ, phospholipase C-gamma; TCR, T-cell receptor; IP3, inositol 1, 4, 5 triphosphate; P, phosphorylation; PI3K, phosphatidylinositol 3′-hydroxyl kinase; Lck and Fyn are Src-family nonreceptor protein tyrosine kinases; Itk, IL-2–inducible T-cell kinase; ZAP-70, ζ-chain–associated protein kinase 70. (Adapted from Figure 1, Pennock ND, et al. *Adv Physiol Educ* 2013;37:273–83.)

TCR along with other costimulatory molecules and adhesion receptors, which results in the formation of a large macromolecular structure called the supramolecular activation complex (SMAC). The unique organization of the SMAC results in prolonged and stronger intracellular interactions and appropriate downstream signaling activity, including phosphorylation of the CD3 receptor components. The complete culmination of the TCR-induced signals results in IL-2 production and secretion, which drives T-cell proliferation in an autocrine and paracrine manner.

It is important to recognize that various T-cell processes and stages of differentiation are specifically regulated by metabolic pathways.[6] For example, naïve T cells rely on oxidative phosphorylation for their metabolism, with additional glycolytic metabolism during times of proliferation. On the other hand, activated effector T cells demonstrate aerobic glycolysis and increased oxidative phosphorylation. T-helper cell-1 (Th1), Th2, and Th17 CD4 effector T cells use glycolytic metabolism, whereas regulatory T-cell and memory T-cell development is enhanced by fatty acid oxidation and catabolic metabolism. In particular, memory T cells are quiescent and mainly use oxidative phosphorylation; however, on antigenic rechallenge, the use of oxidative phosphorylation and glycolysis is rapidly facilitated.[6] While naïve T-cell activation and differentiation into effector cells is essential for a normal immune response during exposure to acute antigenic challenge, chronic antigenic stimulation via persistent antigen exposure and/or inflammation causes an alteration of memory T-cell differentiation, resulting in a state of T-cell exhaustion. During T-cell exhaustion there is a progressive diminution of effector functions, along with upregulation of several inhibitory receptors, metabolic changes, and a failure to transition to a quiescent state and to obtain antigen-independent memory T-cell homeostatic response. Exhausted T cells have only a limited ability to clear pathogens or tumors; however, there is some evidence that this is a reversible phenomenon, at least at a global level.[7]

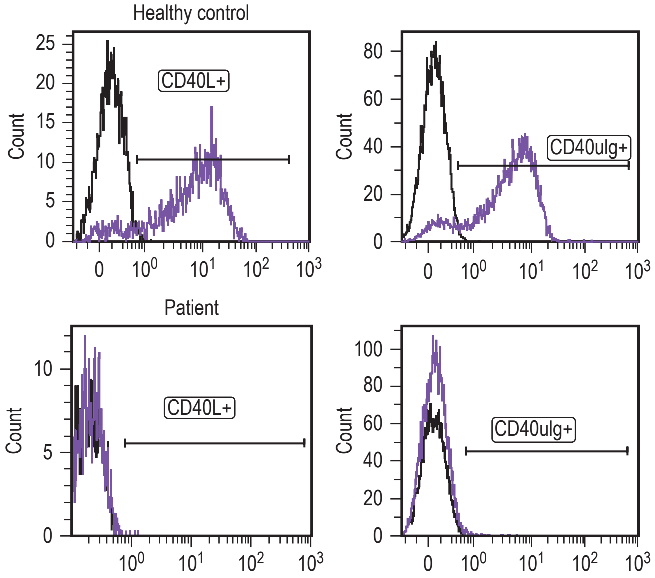


FIG 93.2 Expression of CD40L on Activated CD4 T Cells. CD40L is an early activation marker expressed on CD4 T cells, and it is important for T-cell costimulation and isotype class-switching. In healthy individuals, on T-cell activation CD40L (CD154) is expressed on the cell surface of CD4 T cells and is capable of binding a soluble form of the receptor CD40μIg, which enables assessment of its function *in vitro* (top panel). In patients with X-linked hyper-IgM syndrome due to CD40L deficiency, either there is no protein expression of CD40L on the cell surface and therefore no binding to the soluble receptor (lower panel) or, as is the case with 20% of patients, the CD40L expressed on the cell surface is nonfunctional and cannot bind to the receptor (CD40μIg).

CLINICAL RELEVANCE

Lymphocyte Function Evaluation

- Assessment of lymphocyte function typically refers to measurement of T-cell function via cellular proliferation to nonspecific and specific stimuli; however, it is much broader and more comprehensive and refers to any method that assesses any of the various effector or regulatory aspects of any lymphocyte subset.
- The robustness of clinical interpretation of lymphocyte function assays is dependent on the analytical procedure, sample type, timing of collection, and clinical context, among other factors. Assays performed in clinical diagnostic immunology laboratories are standardized to minimize inconsistencies in these variables.

MEASUREMENT OF T-CELL FUNCTION VIA ACTIVATION MARKERS

Activation of T cells results in the expression of several induced markers, which can be used to ascertain the competence of T cells participating in the immune response. These include CD69, CD154 (CD40L), MHC class II, and CD25, which are expressed in a sequential manner and can be assessed by flow cytometry.[8] In contrast to *in vivo* activation of T cells and subsequent expression of these markers, *in vitro* assessment of T-cell function involves activation of T cells with nonspecific and polyclonal stimulants such as phorbol myristate acetate (PMA) or phytohemagglutinin (PHA), which results in expression of these markers in a kinetically regulated fashion.[8] The expression of CD40L on activated T cells, besides being useful as a global marker for T-cell activation, is used more specifically for the diagnosis of a primary immunodeficiency, X-linked hyper-IgM syndrome or CD40L deficiency (Chapter 35). CD40L on activated CD4 T cells interacts with CD40, expressed constitutively on B cells, and participates in isotype (class)-switching as well as providing costimulatory help to T cells. The expression of CD40L on activated CD4 T cells can easily be ascertained in the laboratory using an *in vitro* T-cell stimulation protocol (Fig. 93.2). In this assay, CD69 is used as a control for early T-cell activation. The majority of patients (80%) with mutations in *CD40LG* have absent protein expression on activated CD4 T cells; however, 20% of mutations may remain permissive for protein expression, but with aberrant function. These patients can be identified by incorporating an additional component in the assay using a soluble form of the receptor CD40-muIg to assess binding of the ligand (Fig. 93.2). The flow cytometric assay for CD40L expression and function offers a rapid diagnostic test for XL-Hyper IgM syndrome[9].

ASSESSMENT OF CELLULAR VIABILITY IN LYMPHOCYTES

The flow cytometry–based assays also allow determination of viable cells in the sample after peripheral blood mononuclear cell (PBMC) isolation from whole heparinized blood, which is particularly important in the interpretation of results (Fig. 93.3). The use of annexin V in a flow cytometry assay enables visualization of apoptotic cells. An early event in cell apoptosis is the translocation of membrane phosphatidylserine (PS) to the cell surface. Annexin V is a calcium (Ca^{2+})-dependent phospholipid-binding protein, which has a strong affinity for PS, and a fluorochrome-conjugated reagent directed at annexin V is used in a flow assay. In addition to visualizing apoptotic cells, dead cells can be assessed by simultaneously using 7-amino-actinomycin-D (7-AAD), which is a membrane-impermeable dye that is generally excluded from viable cells. When internalized into dying or dead cells, it binds to double-stranded DNA by intercalating between base pairs in guanine (G)-cytosine (C)-rich

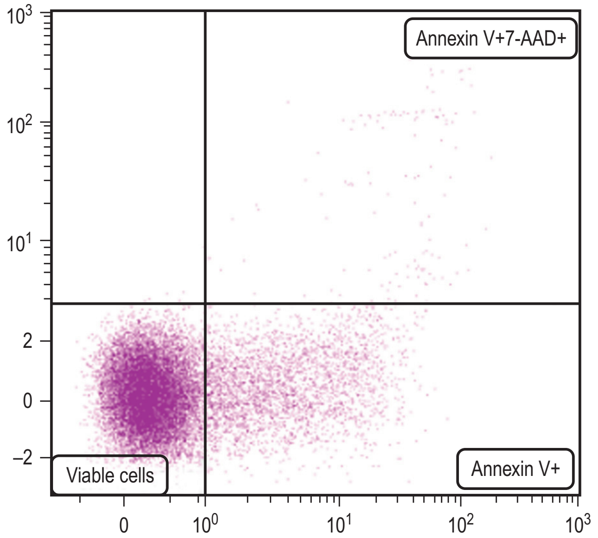


FIG 93.3 Assessment of Cellular Viability for Lymphocyte Function Studies. For all functional assessments of immune response, it is essential and useful to measure the proportion of viable cells at the initiation of the analysis and even at the end of the analysis, if required. Measurement of cell apoptosis and death can be achieved by flow cytometry analysis of annexin V positivity (apoptotic) and annexin V 7-amino-actinomycin-D (7-AAD) dual positive (dead) cells. Cells that are negative for both these markers are viable cells, and their frequency can be assessed in the flow assay.

regions. These two dyes can be combined to provide information on the proportion of viable cells in the starting cell mixture for proliferation assays (Fig. 93.3).

MEASUREMENT OF T-CELL COMPETENCE VIA PROLIFERATION

Mitogens are very potent stimulators of T-cell activation and induce polyclonal T-cell proliferation. It has been suggested that mitogens can induce T-cell proliferative responses even if the lymphocytes are incapable of responding adequately to antigenic (physiological) stimuli. Therefore abnormal T-cell responses to mitogens are considered a diagnostically less sensitive but more specific test of aberrant T-cell function. Lectin mitogens have been shown to bind to the TCR, thereby activating quiescent T cells. Mitogenic stimulation induces increased intracellular calcium (Ca^{2+}) in T cells, which is essential for T-cell proliferation. Whereas PHA is a strong T-cell mitogen (Fig. 93.4), PWM is a weak T-cell mitogen that also induces B-cell activation and proliferation with a different time line for maximal stimulation. Mitogens such as PHA activate T cells by binding to cell membrane glycoproteins, including the TCR-CD3 complex. In addition, there are a number of mitogenic or comitogenic antibodies, including those directed against the CD3 coreceptor that can stimulate T-cell proliferation. Typically, anti-CD3 antibodies provide an initial activation signal and provide a variable proliferative response (Fig. 93.5). Addition of a costimulatory antibody (anti-CD28) to anti-CD3 results in enhanced proliferation (Fig. 93.5). An exogenous T-cell growth factor, such as IL-2, can also be used as an alternate to anti-CD28 costimulation, and in patients with suspected IL-2 receptor (IL-2R)–associated signaling defects, it may be more helpful from a diagnostic perspective than the use of anti-CD28 (Fig. 93.5). IL-2, an autocrine cytokine, has been demonstrated to be critical in T-cell proliferation and in regulation of T-cell growth through binding to a heterotrimeric receptor complex consisting of 3 chains—α, β, and γ (IL-2Rα, IL-2Rβ, and IL-2Rγ)—on the surface of T cells. Triggering of the TCR leads to synthesis of IL-2 in certain T-cell subsets with induction of high-affinity IL-2Rs on antigen- or mitogen-activated T cells; the binding of IL-2 to the IL-2R ultimately leads to T-cell proliferation. The use of exogenous IL-2 in association with anti-CD3 allows discrimination of T cells that cannot proliferate to other mitogenic signals but can respond to a potent growth factor such as IL-2.

KEY CONCEPTS

T-Cell Activation and Function

- Cognate recognition of peptide major histocompatibility complex (MHC) complex on the APC by the T-cell receptor results in formation of the immunological synapse.
- Activated T cells express early activation markers, such as CD69 and CD40L. Other T-cell activation markers include CD25 and MHC class II (human leukocyte antigen–D related [HLA-DR]).
- T cells can be stimulated to proliferate using nonspecific stimulants such as plant lectins (mitogens) and cross-linking of the CD3 coreceptor, along with other costimulatory molecules (*e.g.*, anti-CD28) or in the presence of exogenous interleukin-2 (IL-2).
- The magnitude of antigen-specific T-cell proliferation is dependent on the starting frequency of antigen-specific T cells.
- Flow cytometry–based assays of T-cell proliferation offer the advantages of high resolution, accounting for cellular dilution due to T-cell lymphopenia and single-cell analysis.

Antigens including candida antigen (CA) and tetanus toxoid (TT) have been widely used to measure antigen-specific recall (anamnestic) T-cell responses when assessing cellular immunity. In fact, this may be more revealing about cellular immune compromise than assessing the response of lymphocytes to mitogens because the latter can induce T-cell proliferative responses even if those T cells are incapable of responding adequately to antigenic (physiological) stimuli. Therefore abnormal T-cell responses to antigens are considered a diagnostically more sensitive, but less specific, test of aberrant T-cell function. Antigens used in recall assays measure the ability of T cells bearing specific TCRs to respond to antigenic peptides presented by APCs. The antigens used for assessment of the cellular immune response are selected to represent antigens, seen by a majority of the population either through natural exposure (CA) or as a result of vaccination (TT) (Fig. 93.6).

In addition to measuring cellular proliferation as a readout for T-cell function, the production of cytokines by activated T cells is another important component for evaluating T-cell functional activity. T cells typically are not monofunctional, producing a single cytokine on activation; rather, they are multi- or polyfunctional, and the range of cytokines are produced in a sequential manner rather than simultaneously, although a population of stimulated T cells will have individual T cells producing different cytokines in a temporally regulated manner. These cytokines can be measured by intracellular flow cytometry after T-cell activation with mitogens, in both CD4 and CD8

T-cell subsets or alternatively in the culture supernatant of activated cells. Typically, the cytokines measured in *in vitro* stimulation assays include IL-2, interferon (IFN)-γ, IL-4, IL-5, IL-6, IL-17, and tumor necrosis factor (TNF)-α.

The process of T-cell activation as described above results in IL-2 production that drives T-cell proliferation. Measuring T-cell proliferation after stimulation *in vitro* with various specific and nonspecific stimuli has been a staple of the diagnostic immunology repertoire for several decades. T-cell proliferation assays can be divided into three categories: use of nonspecific mitogenic stimuli such as PHA, pokeweed mitogen (PWM), and concanavalin A (Con A); physiological global stimuli based on CD3 coreceptor cross-linking (using anti-CD3, either soluble or bead-bound) with or without CD28 cross-linking (using anti-CD28, either soluble or bead-bound); and with or without exogenous IL-2. Antigen-specific stimulation depends on the use of common recall antigens.

The commonly used method for assessing T-cell proliferation for decades involved measurement of ^{3}H-thymidine (^{3}H-T) incorporated into the DNA of proliferating cells. The results are expressed as either counts per minute (cpm) or disintegrations per minute (dpm) of both activated and nonactivated cells (background) cultured for a fixed period of time, usually 72 hours for mitogens. A reference range based on proliferation results from a group of control subjects should be provided along with the patients' results (both background and poststimulation results). Although this method remains widely employed, several disadvantages exist, including but not limited to the use of radioactive material (^{3}H-T); its inability to discriminate responder cell subsets or to account for cellular dilution, which is particularly relevant in the context of T-cell lymphopenia; and the lack of information on the contribution of cell death after stimulation and its impact on the final result. To overcome the intrinsic shortcomings of the ^{3}H-T method, newer flow cytometry assays are currently being used in the clinical diagnostic setting and are gaining popularity because of the additional information they provide.

The flow cytometric methods for measuring cell proliferation include the use of fluorescent dyes to identify proliferating cells. One of the more commonly used dyes, carboxyfluorescein diacetate succinimidyl ester (CFSE), must be carefully used when measuring lymphocyte proliferation *in vitro* since it can be toxic to cells and nonoptimal labeling conditions can impact measurement of cell proliferation and interpretation of results. CFSE is a fluorescent cell-membrane permeable dye similar in physical properties to the commonly used fluorochrome, fluorescein isothiocyanate (FITC). During cell proliferation, the intensity of staining in daughter cells is half that of parent cells, allowing visualization of the number of rounds of cell divisions associated with the successive decrease in fluorescence. The disadvantages to using CFSE in the clinical laboratory are its photo-instability, limitation in the number of cell divisions being identified (≤7 cell divisions), and interpretation requires measuring loss of signal rather than gain of signal. In addition, as noted above, CFSE at concentrations of 37 nM to 10 μM can be toxic to cells resulting in increased cell death. Furthermore, it can modulate expression of activation markers, resulting in a decrease in CD69, HLA-DR, and CD25 expression. Finally, it has been reported that there is an increase in the number of false-positive results with CFSE making it suboptimal for measuring lymphocyte proliferation in patients with severe cellular immunodeficiencies. Alternatives to CFSE include related compounds, such as Cell Trace Violet (CTV) and Cell Proliferation Dye eFluor 670 (CPD), which have different excitation and emission spectra compared with CFSE. These dyes also can be used for tracking lymphocyte proliferation status *in vivo,* in animal models, permitting measurement of up to 11 cell divisions.*

Another alternative to tritiated thymidine is the use of a thymidine analog, 5-ethynyl-2′-deoxyuridine (Edu®), which can combine with a fluorescent azide in a copper-catalyzed cycloaddition reaction (referred to as "Click" chemistry) and permits flow cytometric evaluation of lymphocyte proliferative responses by assessing its incorporation into cellular DNA. Edu®- labeling has been shown to be a fast and sensitive method for measuring cell proliferation and also facilitates identification of dividing cells. It is relatively more photo-stable than CFSE and is added to cells after completion of the stimulation period with the measurement being comparable to the thymidine method in regards to a gain of signal is the endpoint. The Edu® assay has not shown the limitations of the CFSE method in the clinical laboratory and has been used to evaluate lymphocyte proliferation in a spectrum of patients, including those with severe combined immunodeficiencies (SCIDs). In fact, this assay has been particularly useful in discriminating between functional T cells and nonfunctional T cells in the context of severe T cell lymphopenia, which cannot be achieved with the standard thymidine assay (Figs. 93.4-93.6). All the proliferation data shown in this chapter utilize this Edu®-based measurement of T cell proliferation, and results are typically provided for both CD45 bright positive (total lymphocytes) and CD3 T cells, with the former being more representative of the data generated with the standard thymidine assay.*

MEASUREMENT OF CELL-MEDIATED CYTOTOXICITY

CD8 T cells are considered the representative cytotoxic T cell in the immune system, and cellular cytotoxicity is a mechanism to eliminate cells infected with intracellular pathogens, allogeneic cells, or tumor cells. CD8 T cells, like CD4 T cells, recognize antigen via the TCR and kill target cells via cytotoxic protein granule exocytosis and/or cytokine production. Over the past few decades, the cytotoxic potential of CD4 T cells has been described, particularly in viral infections. Although cytotoxic CD4 T cells are rare in the circulation of healthy individuals (<2%), they can account for substantial proportions of total CD4 T cells in certain viral infections, including but not limited to HIV.[10] The CD4 cytotoxic T cells appear to be a lineage of memory CD4 T cells that contain cytotoxic granules with perforin and granzymes, and they are thought to arise in the context of chronic or potent activation with viral infections such as cytomegalovirus (CMV), Epstein-Barr virus (EBV), human immunodeficiency virus (HIV), or certain autoimmune diseases.[11]

Cellular cytotoxic activity has been conventionally measured by release of chromium (^{51}Cr) from labeled target (T) cells cultured with various ratios of effector (E) cells (varied E:T ratio). The percentage lysis is developed using the level of radioactivity in the supernatant of target cells in comparison to

---

*This section has been reproduced in part with permission from ASM Press (Abraham RS, Lymphocyte Activation) (copyright permission obtained)

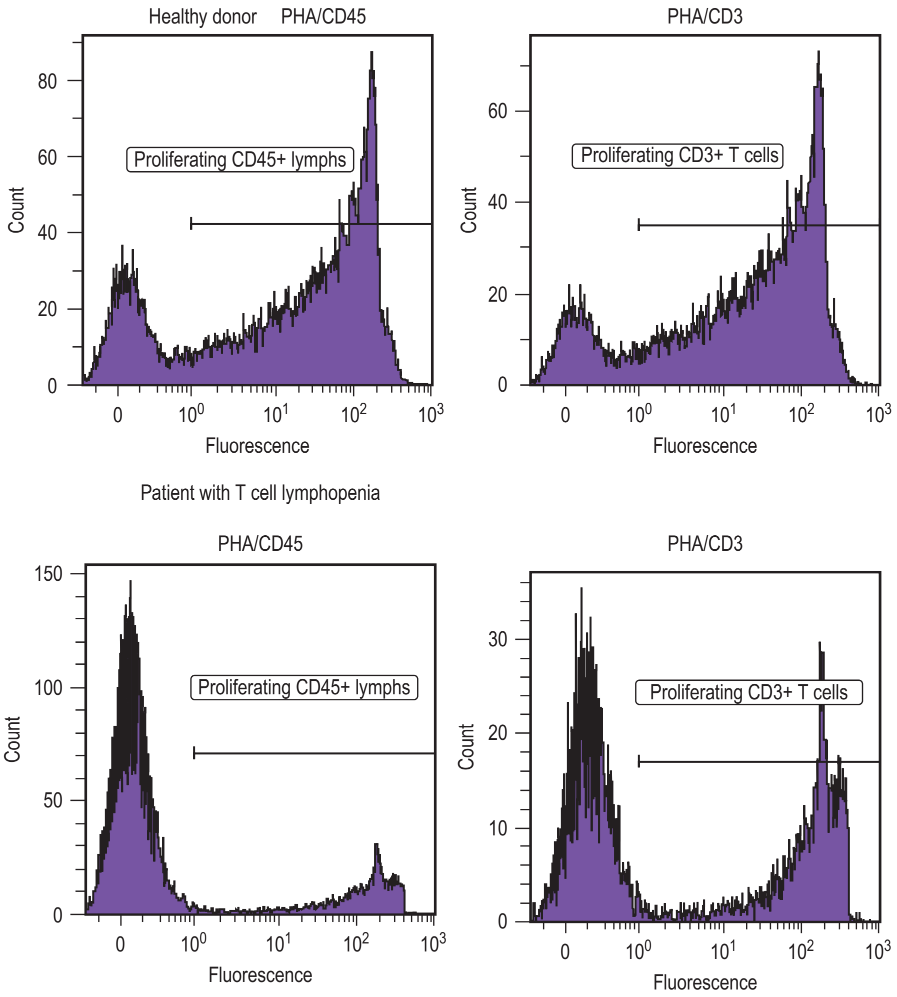


FIG 93.4 T-Cell Proliferation to Mitogenic Stimulants. T cells can respond polyclonally to *in vitro* stimulation with various mitogens. T-cell proliferation assessment by flow cytometry allows analysis in total lymphocytes (CD45$^+$) and T cells (CD3$^+$). The top panel demonstrates T-cell proliferation to phytohemagglutinin (PHA) in both cell subsets. In patients with T-cell lymphopenia, where cellular dilution may be a concern, single-cell analysis by flow cytometry allows discrimination of functional versus nonfunctional T cells. In the lower panel, a patient with T-cell lymphopenia has abnormal proliferation when total lymphocytes are assessed; however, the response is normal when the T-cell compartment is specifically evaluated. Therefore this patient, who would have been classified as abnormal based only on the CD45$^+$ lymphocyte response, can be reclassified as having normal T-cell proliferation based on the CD3$^+$ T-cell response, but with significant T-cell lymphopenia.

KEY CONCEPTS

Cellular Cytotoxicity

- CD8$^+$ T cells and natural killer (NK) cells are involved in killing of cellular targets and contain intracellular granules with cytotoxic proteins, such as perforin and granzymes.
- Cytotoxic CD4$^+$ T cells are a subset of memory T cells with cytolytic potential and are usually observed in circulation in the context of chronic viral infections, *e.g.,* cytomegalovirus (CMV), human immunodeficiency virus (HIV).
- Tetramer-based assays have been used to quantify and delineate function of antigen-specific CD8$^+$ T cells (also CD4$^+$ T cells).
- Cellular degranulation results in expression of CD107a on the cell surface of CD8$^+$ T cells and NK cells and is often used as a surrogate of cytotoxic activity.
- NK cells recognize target cells that lack major histocompatibility complex (MHC) class I, *e.g.,* viral cells or tumor cells that have downregulated MHC class I.
- The majority of circulating NK cells are cytotoxic (CD3$^-$CD16^{++}CD56$^{+/-}$), while a minority are cytokine producing (CD3$^-$CD56^{++}).
- Interleukin (IL)-2 augments NK cell–mediated cytotoxicity (lymphokine-activated killing) and also induces interferon (IFN)-γ secretion by NK cells.
- Regulatory T cells control NK cell activation and cytotoxic function by limiting availability of IL-2.

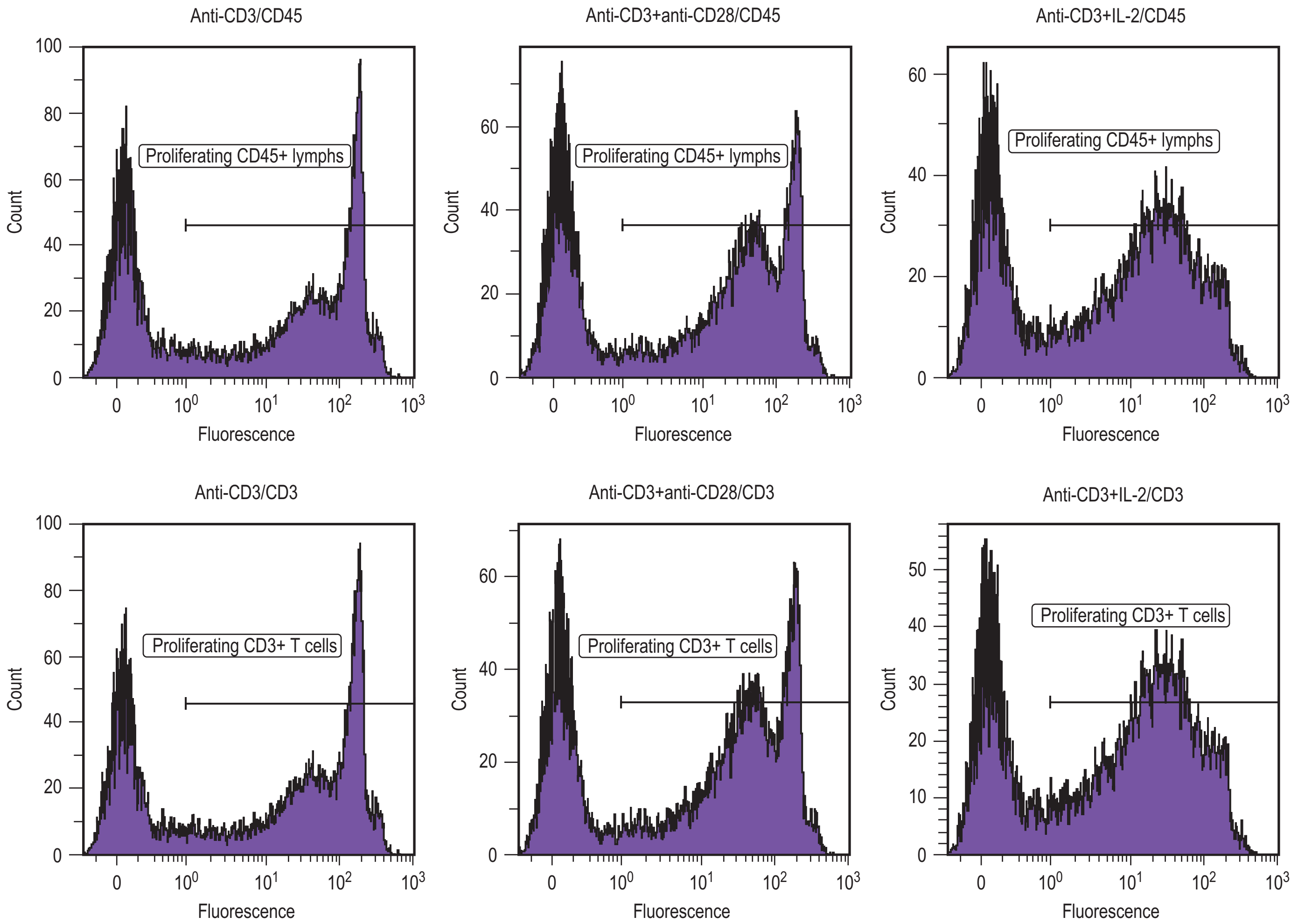


FIG 93.5 **T-Cell Proliferation to Anti-CD3 Stimulants.** Assessment of T-cell proliferation to CD3 coreceptor cross-linking provides a more physiological, yet global, evaluation of T-cell function. T cells are stimulated with either soluble anti-CD3 alone or soluble anti-CD3 plus anti-CD28 or with soluble anti-CD3 plus interleukin (IL)-2. In each case, proliferation is measured in total $CD45^+$ lymphocytes and $CD3^+$ T cells. An example of anti-CD3–stimulated T-cell response is shown in a healthy donor.

controls consisting of target cells cultured alone (no lysis) and target cells exposed to hypotonic conditions (*e.g.*, distilled water, saponin) representing 100% lysis. Similar to the radioactive assay for measuring T-cell proliferation described previously, there are several limitations to the radioactive cytotoxicity assay, including the hazard of radioactive material, the difficulty in labeling target cells with ^{51}Cr, and bulk quantitation of cytotoxic activity, which does not allow single cell–specific analysis. Alternative assays have been developed, including flow cytometry and ELISPOT-based methods.[12] More recently, a method of assessing human CD8 T-cell cytotoxicity has been described in which both the effector and target cells are primary cells (in contrast to using cell lines as target cells).[13] In this particular assay, autologous B cells isolated from PBMC are used as target cells and cocultured with antigen-specific CD8 effector cells. The killing of the fluorescently labeled target B cells (see NK cell section) is used to estimate cytotoxic activity. Antigen-specific effector cells are quantified in the total CD8 T-cell pool using MHC-peptide tetramers.[13]

A tetramer-based approach to quantifying and measuring function of CMV-specific CD8 T cells has been available in the clinical diagnostic immunology laboratory for over a decade (Fig. 93.7). However, the limitation of the tetramer approach in the clinical diagnostic setting is that it is constrained by the number of HLA-peptide tetramer (multimer) combinations that are available for use with a particular antigen (*e.g.*, CMV, EBV). It also requires *a priori* knowledge of the patient's HLA genotype, and very often it does not incorporate a comprehensive assessment of the CD8 and CD4 cytotoxic T-cell response. Although some laboratories use only a quantitative approach to determine immune competence to pathogens, such as CMV with the tetramer assay, more comprehensive assays are available that also provide a functional assessment of these antigen-specific CD8 T cells (Fig. 93.7).

An alternative approach is to evaluate for cytotoxic function by assessing degranulation by evaluating for CD107a expression by the effector T cell (described in further detail in the NK cell function section) and/or evaluating IFNγ production by activated

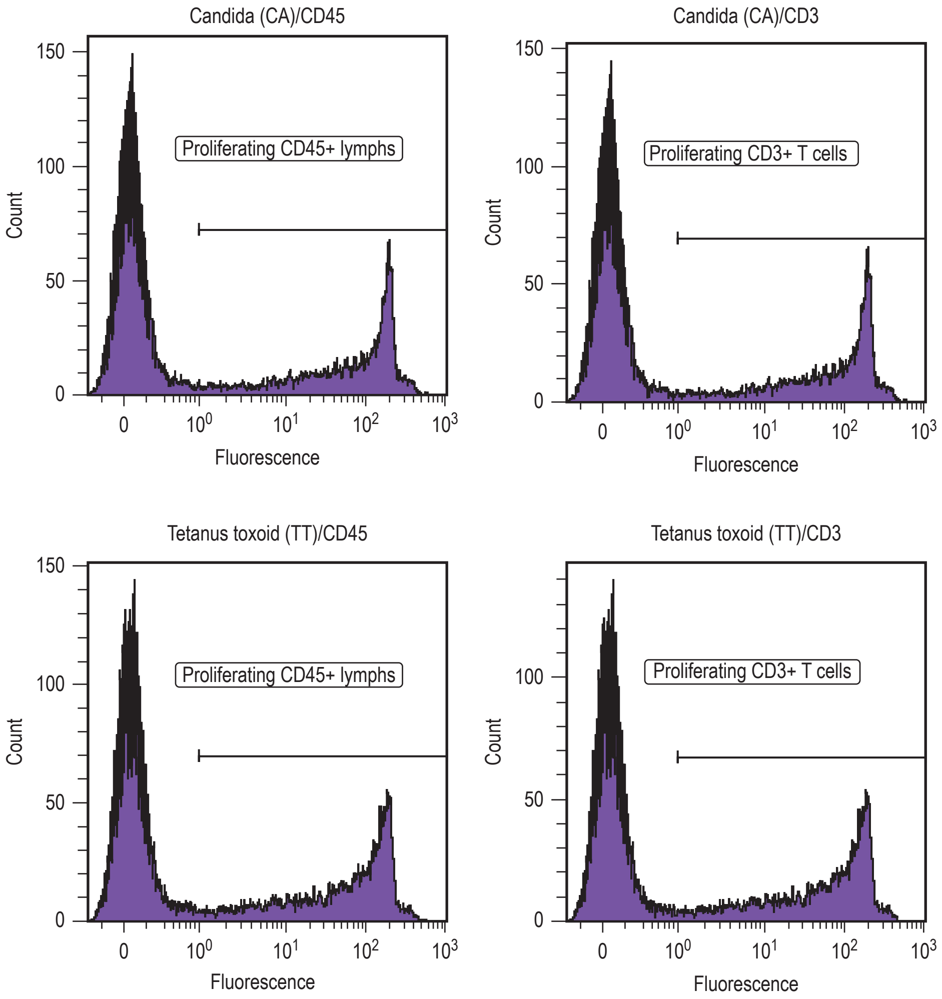


FIG 93.6 Assessment of Antigen-Specific T-Cell Proliferation. Antigens such as candida antigen (CA) and tetanus toxoid (TT) are frequently used to assess antigen-specific T-cell function. As with mitogens and anti-CD3 stimulation, proliferation is measured in total CD45$^+$ lymphocytes and CD3$^+$ T cells. An example in a healthy donor is shown.

cytotoxic CD8 T cells. As a positive control, CD8 T cells are polyclonally stimulated with PMA and ionomycin, and the same markers (CD107a expression and IFN-γ production) are measured. Approximately 10–60% of CD8 T cells in healthy adults are activated to express CD107a and to secrete IFN-γ under these conditions. An approach that avoids the use of tetramers involves the stimulation of patient-specific PBMC with overlapping peptide pools of the specific antigen and, subsequently, the evaluation of the CD8 and CD4 T cells that proliferate in response to antigenic stimulation. These T cells can also be assessed for cytotoxic potential by measuring intracellular perforin and granzyme expression as well as degranulation (CD107a expression) after stimulation.[8]

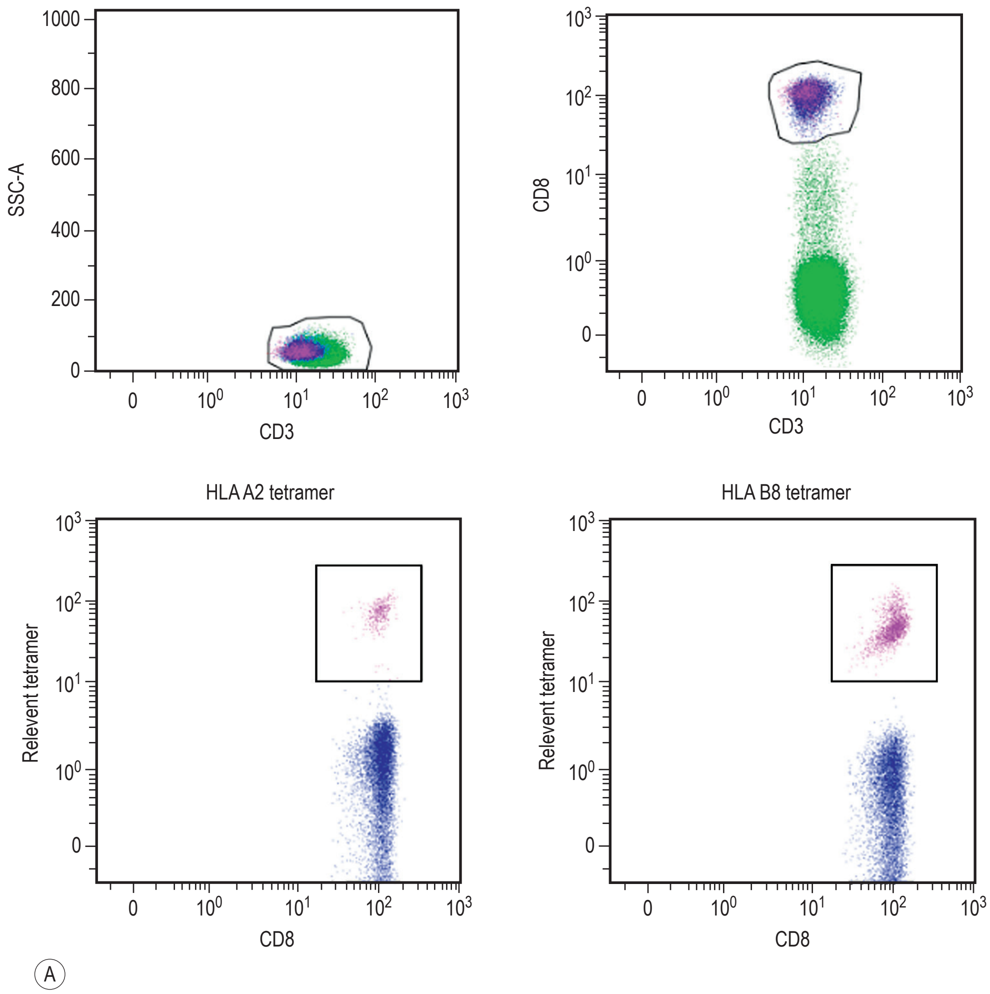
A CMV-specific assay has been developed that measures IFN-γ production in response to CD8 T-cell stimulation with a pool of MHC class I–restricted CMV peptides, the QuantiFERON-CMV (Cellestis Ltd., Melbourne, Australia).[14] This commercial assay is simpler to perform than the tetramer-based approach for quantification and functional analysis of antigen-specific CD8 T cells. It has been applied in the context of risk prediction for CMV in organ transplant and allogeneic hematopoietic cell transplant (HCT) patients. However, this assay has theoretical limitations in that assessing the production of a single cytokine is unlikely to represent the breadth of the immune response to a complex specific antigen such as CMV.

NK CELL ACTIVATION AND FUNCTION

NK cells are innate immune lymphocytes that provide the cellular basis for innate responses to specific viruses as well as to tumor cells[15,16] (Chapter 3). Unlike T and B cells of the adaptive immune system, NK cells do not have antigen-specific recognition and killing of target cells. NK cells participate directly in effector functions via cytotoxicity and production of cytokines (*e.g.*, IFN-γ) upon activation. Proinflammatory cytokines such as IL-12, IL-15, and IL-18 trigger NK cell proliferation as well as cytotoxicity and production of IFN-γ.[17] The negative regulation of NK cells is controlled by receptors that recognize MHC class I molecules preventing NK cell–mediated cytotoxicity. In contrast, virally infected or tumor cells typically downregulate MHC class I, making them appropriate targets of NK cell–mediated cytotoxicity in the presence of relevant ligands expressed by the target cell. NK cells, like cytotoxic T cells,[18] contain granules with cytotoxic proteins including perforin and granzymes, which are serine proteases that recognize different substrates.

NK CELL CYTOTOXICITY

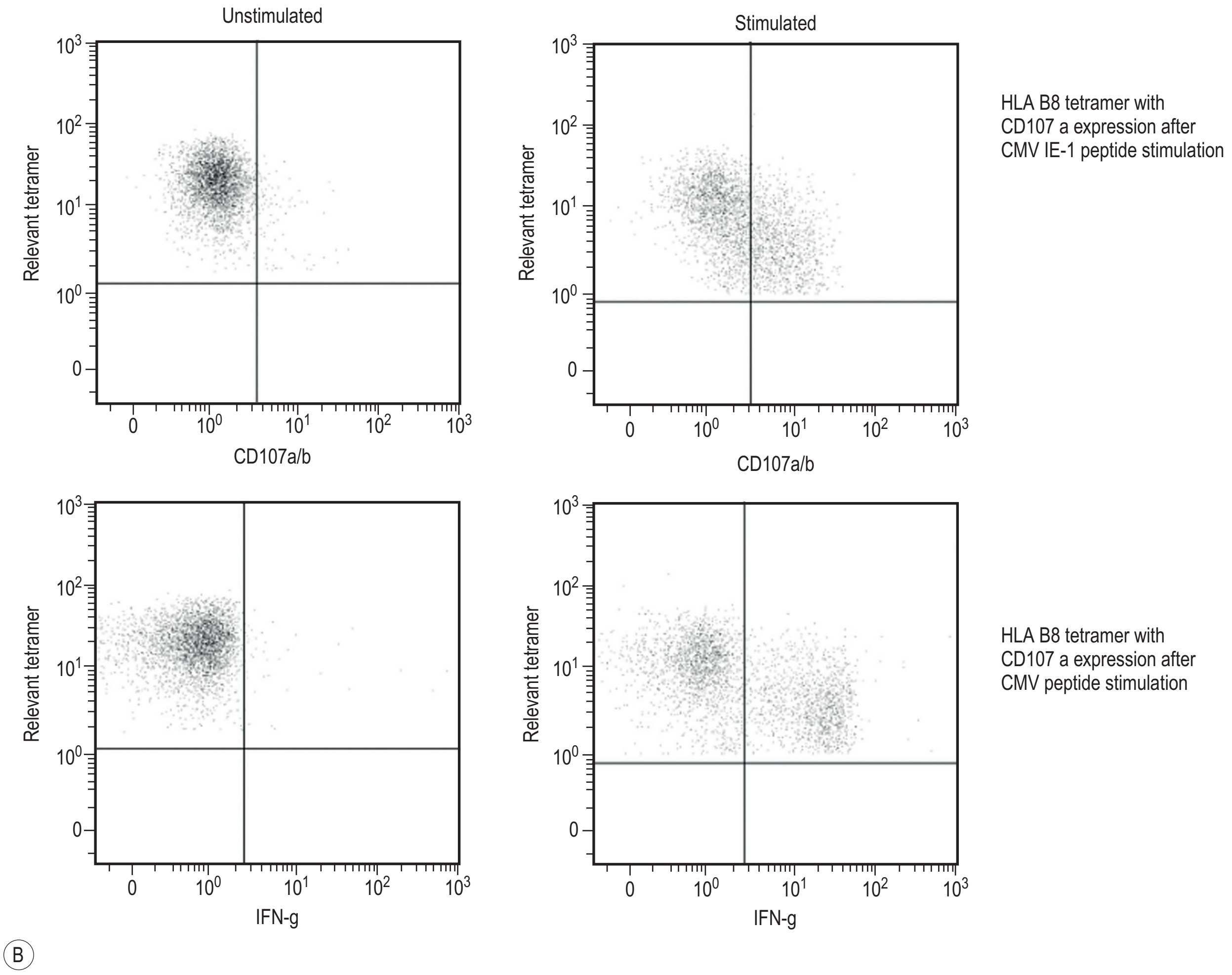
The majority of NK cells can be identified by a lack of CD3 on the cell surface in conjunction with the expression of CD56 (NCAM [neural cell adhesion molecule]) and CD16 (FcγRIII). NK cells can be subdivided into two major subsets based on their relative expression of CD56 and CD16: CD16$^{+++(\text{bright})}$ CD56$^{+/-(\text{dim})}$ NK cells, referred to as cytotoxic NK cells, and CD56$^{+++(\text{bright})}$ CD16 or CD16$^{+/-}$ NK cells, known as regulatory or cytokine-producing NK cells (Fig. 93.8). The majority (~90%) of circulating human NK cells belong to the cytotoxic category, while a minority (10%) represent cytokine-producing NK cells. Cytotoxicity can be subdivided into natural or spontaneous cytotoxicity directed largely toward virally infected cells or tumor cells, in the absence



HLA allele	CMV peptide
A1	pp50
A2	pp65
B7	pp65
B8	IE-1
B35	pp65

FIG 93.7A, B Quantitative of Antigen-Specific CD8 T Cells and Functional Evaluation Using Major Histocompatibility Complex (MHC) Class I Tetramers. The tetramer (multimer) technology has been useful for accurate quantitation of antigen-specific CD8 (or CD4 T cells, if MHC class II tetramers are used). The top panels show identification of CD3$^+$ T cells from total lymphocytes and subsequent segregation into CD3$^+$CD8$^+$ T cells. In this assay, cytomegalovirus (CMV)-specific CD8$^+$ T cells are quantitated using five MHC class I tetramers (human leukocyte antigen [HLA] A1, A2, B7, B8, and B35), each recognizing a unique CMV peptide from three major CMV antigenic proteins (pp50, pp65, and IE-1) as listed in the box. Based on the HLA class I haplotype of the individual, one or more tetramers are used for stimulation. The bottom panels show an example of a patient with HLA A2-specific CMV-CD8$^+$ T cells and another patient with HLA B8-specific CMV-CD8$^+$ T cells. These CMV-specific CD8$^+$ T cells can be assessed for functional capacity by gating on the tetramer-positive CD8$^+$ T cells and then measuring degranulation (CD107a expression) and inteferon (IFN)-γ production after *in vitro* stimulation with the specific CMV peptide (lower panel). The data are shown for CD107a expression and IFN-γ production in unstimulated CMV-CD8$^+$ T cells and peptide-stimulated CMV-CD8$^+$ T cells.

Continued



Unstimulated

Stimulated

HLA B8 tetramer with
CD107 a expression after
CMV IE-1 peptide stimulation

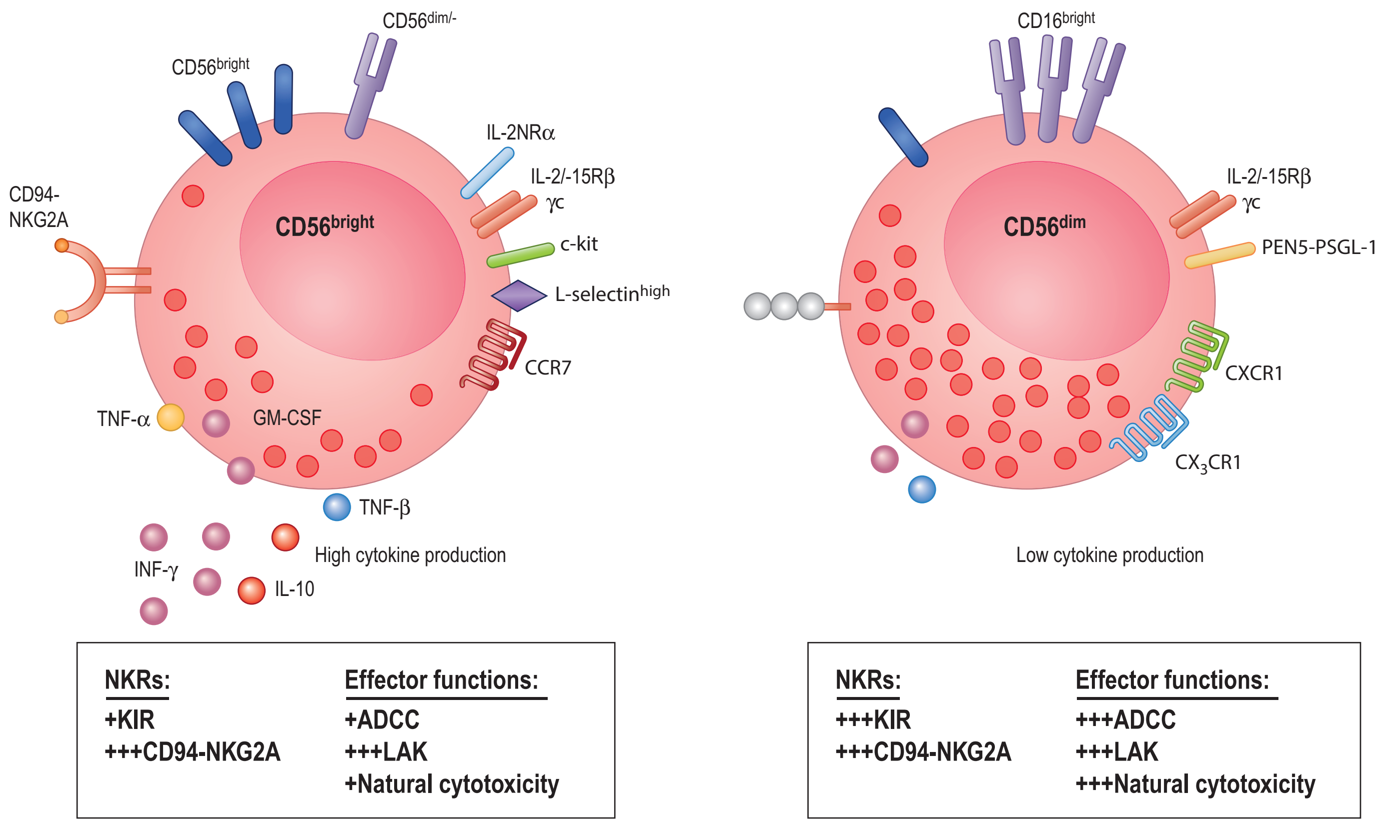
HLA B8 tetramer with
CD107 a expression after
CMV peptide stimulation

FIG 93.7B, cont'd

of prior stimulation or immunization, and antibody-dependent cellular cytotoxicity directed against antibody-coated target cells. NK cells go through a process of education or "licensing" whereby NK cells that express inhibitory receptors to self-MHC class I molecules are called licensed, which means they are more functionally responsive to stimulation, whereas unlicensed NK cells lack receptors for self-MHC class I and are hyporesponsive (Fig. 93.9).

NK cell function is measured in the clinical laboratory by assessment of spontaneous (natural) NK cell cytotoxicity using an MHC class I–deficient myelogenous leukemia cell line, K562. Traditional methods for measuring NK cell cytotoxicity are similar to those used for assessing cytotoxic T cell (CTL) function based on varying effector:target ratios in a 4- to 16-hr ^{51}Cr-release assay compared with the no-lysis and 100% lysis conditions as described for the T-cell cytotoxicity assay. However, there is interest in the use of flow cytometry to measure spontaneous or IL-2–activated (lymphokine-activated killer [LAK]) NK cell cytotoxicity in the clinical laboratory to avoid radioactivity. One flow cytometric assay employed in the clinical laboratory involves the use of fluorescently labeled (Cell Tracker® dyes) target cells (K562) incubated with effector cells (donor or patient PBMC) in the absence (spontaneous) or presence of IL-2 (LAK). Following coincubation, the lysis of target cells is measured by using 7-AAD (7-amino-actinomycin-D) (Fig. 93.10). IL-2 enhances cytotoxic function with increased lytic potential against a broad range of target cells. IL-2 has also been shown to induce IFN-γ secretion by NK cells with upregulation of activation markers, such as CD25 and CD69.[19] Treg cells control NK cell activation and cytotoxic function by limiting access to IL-2.[20] Other methods for measuring NK cell cytotoxic function that are primarily used in the research setting include the use of image cytometry,[21] microchip screening,[22] and flow-based assays using other dyes, such as calcein AM.[23]

The direct measurement of NK cell cytotoxic function is useful in a variety of clinical contexts, especially in patients with inherited immune defects affecting NK cells and function, including but not limited to recurrent/persistent herpesvirus infections and familial/primary hemophagocytic lymphohistiocytosis (FLH/HLH). There are other parameters that are widely used as surrogates of cytotoxicity, including flow cytometric measurement of granule exocytosis/degranulation. The membrane of cytotoxic granules in both NK cells and CD8 T cells is composed of several proteins, including CD107a (lysosomal-associated membrane protein 1 [LAMP-1]). On stimulation of NK cells and CD8 cytotoxic T cells, CD107a is upregulated and expressed on the cell surface (Fig. 93.11) concomitantly with cytokine secretion and target cell lysis, and therefore has been frequently used to


CD56dim/-
CD56bright
CD94-NKG2A
CD56bright
IL-2NRα
IL-2/-15Rβ
γc
c-kit
L-selectinhigh
CCR7
TNF-α
GM-CSF
TNF-β
INF-γ
IL-10
High cytokine production
CD16bright
CD56dim
IL-2/-15Rβ
γc
PEN5-PSGL-1
CXCR1
CX3CR1
Low cytokine production

NKRs:	Effector functions:
+KIR	+ADCC
+++CD94-NKG2A	+++LAK
	+Natural cytotoxicity

NKRs:	Effector functions:
+++KIR	+++ADCC
+++CD94-NKG2A	+++LAK
	+++Natural cytotoxicity

FIG 93.8 **NK Cell Subsets in Blood.** Natural killer (NK) cells are present as two major subsets in blood: cytotoxic NK cells (CD16^{++}CD56$^{+/-}$) and cytokine-producing NK cells (CD56^{++}). The former accounts for 90% of NK cells in circulation and the latter, 10%. The cytotoxic NK cells are very efficient in mediating target cell killing via spontaneous (natural), lymphokine-activated (LAK), or antibody-dependent cytotoxicity mechanisms. (Adapted from figure in Cooper MA, et al. *Trends Immunol* 2001;22:633–40.)

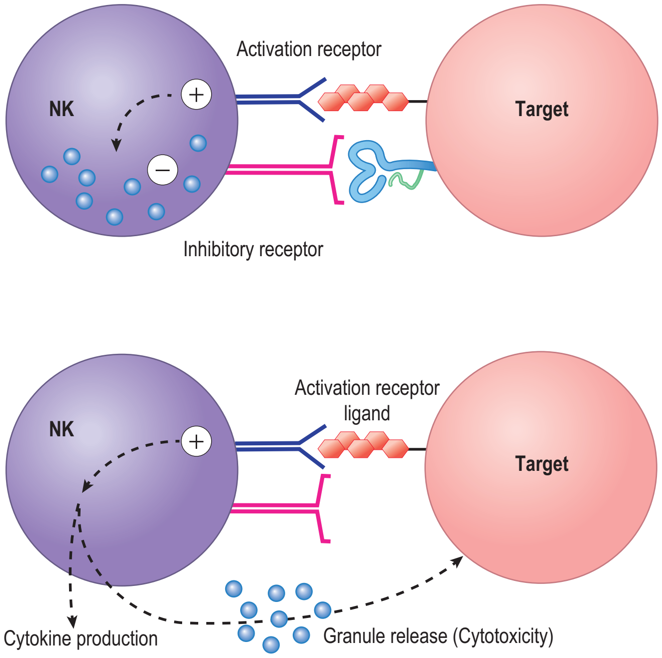

NK
Activation receptor
+
−
Target
Inhibitory receptor
NK
Activation receptor ligand
+
Target
Cytokine production
Granule release (Cytotoxicity)

FIG 93.9 **NK Cell Recognition of Target for Cytotoxic Function.** Natural killer (NK) cells have inhibitory receptors for self-major histocompatibility complex (MHC) class I. Therefore NK cells do not kill target cells expressing MHC class I (upper panel). However, when MHC class I expression is downregulated (viral infections, tumors), the target cell is primed for NK cell cytotoxicity, which is mediated by granule exocytosis and release of cytotoxic proteins, including perforin and granzymes (lower panel). (Adapted from figure in French AR and Yokoyama WM. *Arthritis Res Ther* 2004;6:8–14.)

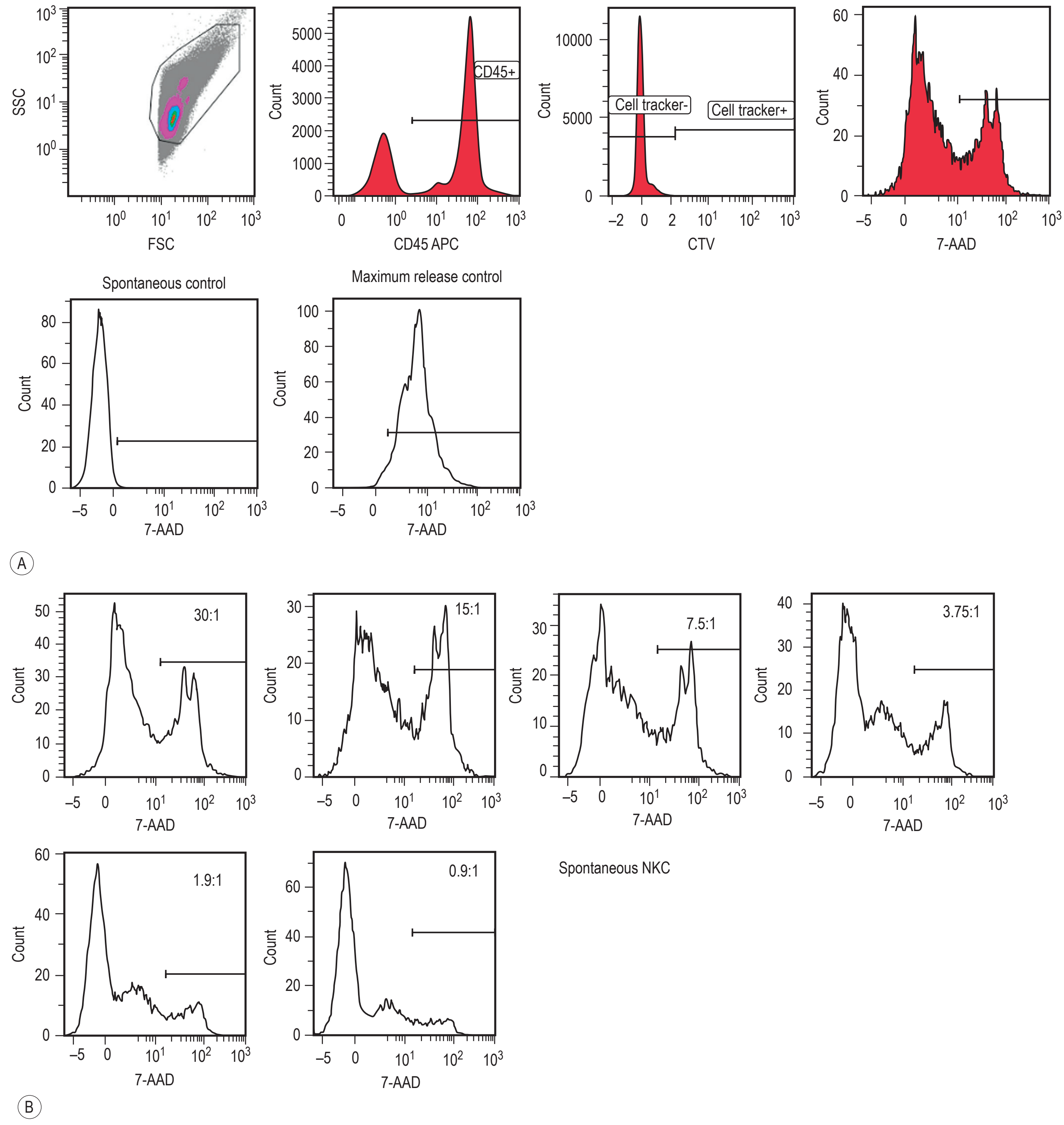


FIG 93.10A, B Flow Cytometric Assessment of Natural Killer (NK) Cell Cytotoxic Function. The "gold standard" for measuring natural killer cell cytotoxicity (NKC) uses a radioactive label for target cells. In the flow-based assay, target cells (K562, a major histocompatibility complex [MHC] class I–negative erythroleukemia cell line derived from a patient with CML, chronic myelogenous leukemia in blast crisis) are labeled with a fluorescent marker, Cell Tracker® (CTV). In the top panel, the gating strategy for analysis is shown. Peripheral blood mononuclear cells (PBMC) from healthy controls or patients are cocultured with labeled target cells, and the two cell populations are identified by forward scatter (FSC) and side scatter (SSC) parameters. Afterward, $CD45^+$ lymphocytes are identified because NK cells (effector cells) are within the lymphocyte subset. CTV^+ and CTV^- cells are separated; CTV^+ cells that are also 7-amino-actinomycin-D $(7\text{-AAD})^+$ are identified as killed target cells, and their frequency is estimated. A negative control for spontaneous cytotoxicity is achieved by incubation of labeled target cells without effector cells added (called spontaneous control). A positive control for maximal cytotoxicity is achieved by incubating labeled target cells for the same duration but using a permeabilization/fixation method to obtain maximal cell death and fluorescent-positive cells (called maximum release control). The middle panel depicts spontaneous NK cell cytotoxicity with different concentrations of effector cells (PBMC; E) to a single concentration of target cells (T) starting at 30:1. The data are expressed as delta (Δ)% cytotoxicity after the spontaneous control signal is subtracted from the E:T signal for each concentration (shown in the table).

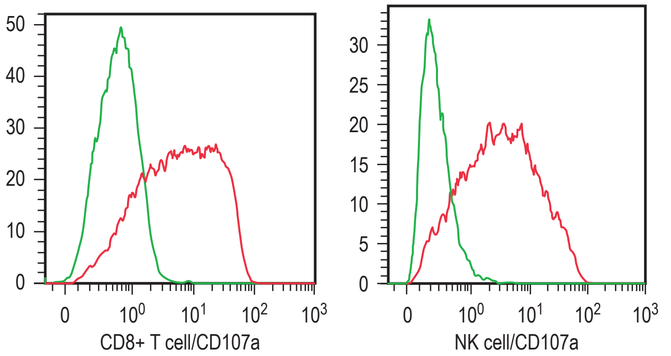


FIG 93.11 Cellular Degranulation Measurement. Cytotoxic lymphocytes undergo degranulation upon stimulation before mediating cytotoxic function. As a result of the degranulation process, CD107a, expressed intracellularly, is now expressed on the cell surface and can be assessed by flow cytometry. Degranulation can be determined in the CD8$^+$ T cells (left panel) and natural killer (NK) cells (right panel). In this analysis, peripheral blood mononuclear cells (PBMCs) were stimulated with phorbol myristate acetate (PMA) (10 ng/mL) and ionomycin (1 µg/mL) for 3 hours. Following stimulation, cells were stained for CD107a in CD8$^+$ T cells (left panel) and in NK cells (right panel). NK cells can also be stimulated with K562 cells as an alternative to mitogen stimulation. The degranulation assay is frequently used as a surrogate for, or as an adjunct to, measurement of cellular cytotoxicity.

extrapolate the magnitude of cytotoxic activity.[24,25] Degranulation (CD107a expression) assays have gained traction in the assessment of FHL.[26] The exception is in patients with mutations in the gene encoding perforin (*PRF1;* FHL type 2) where degranulation (CD107a expression) of cytotoxic cells is normal while cytotoxicity is abnormal. Therefore while degranulation assays may provide relevant information, they cannot substitute for direct measurement of cytotoxicity in all settings (Fig. 93.10B).

NK cells also mediate killing of target cells, via recognition of surface-bound immunoglobulin through the Fcγ receptors (Ig-Fc receptors, specifically CD16 or FcγRIIIa), through a mechanism called antibody-dependent cellular cytotoxicity (ADCC). ADCC has also traditionally used the ^{51}Cr method discussed above, although currently flow cytometry has largely replaced radioactive methods. ADCC assays are particularly relevant in the assessment of antitumor responses, especially of new biological immunomodulatory agents, as well as of the function of alloantibodies in allograft rejection. Recently a method for assessing ADCC by flow cytometry using cryopreserved PBMC has been described,[27] although most clinical laboratories use fresh samples in these functional assays.

EVALUATION OF REGULATORY T-CELL (TREG) FUNCTION

Regulatory T cells (FOXP3$^+$Treg) have been well described over the past several years as a distinct subset of T cells that are both developmentally and functionally unique and essential to maintaining immune homeostasis and self-tolerance.[28,29] The major subpopulations of Treg cells include natural Treg (nTreg) cells that are produced in the thymus and induced Treg (iTreg) cells that are generated in the periphery from conventional FOXP3$^-$CD4$^+$ T cells.[30] The dysfunction of Treg cells results in severe autoimmunity, with IPEX serving as the classic prototype[3,31] (Chapter 35). Besides their role in controlling development of autoimmunity, lack of Treg cells or abnormal Treg function has been implicated in the etiopathogenesis of graft-versus-host disease (GVHD)[32] and allograft rejection, while its presence and normal function have been shown to promote allograft tolerance in solid-organ transplantation.[33,34]

KEY CONCEPTS

Regulatory T Cells and B Cells

- Regulatory T cells (Tregs) and B cells (Bregs) are distinct subsets of cells with immune regulatory potential and importance in maintaining immune homeostasis and self-tolerance, preventing autoimmunity, and limiting inflammatory damage.
- FOXP3 transcription factor is a marker for natural regulatory T cells; while there are no specific cellular markers that define Breg, production of interleukin (IL)-10 is considered a hallmark.
- Both Treg and Breg have been shown to play important roles in autoimmunity, in graft-versus-host disease (GVHD), and in mediating transplant tolerance.
- Treg function is measured *in vitro* through different types of suppression assays, while Breg are characterized by their ability to produce IL-10 when stimulated via Toll-like receptor (TLR9)/B-cell receptor (BCR) or CD40L crosslinking.
- Natural, thymic-derived Treg are CD4$^+$25$^+$FOXP3$^+$ while Breg are CD19$^+$CD24^{hi}CD38^{hi}IL-10$^+$.
- Treg can also be induced in the periphery from conventional T cells through cytokine signals (iTreg).

Treg cell function is measured *in vitro* with Treg suppression assays, which utilize sorted CD4$^+$CD25$^+$ Treg cells in a coculture system with conventional effector T cells to assess suppression of proliferation.[35] However, there are limitations to this approach, including the issue of whether the Treg suppression assays *in vitro* reflect the biological process *in vivo.* In addition, antigen-specific Treg suppression cannot be adequately assessed due to the technical difficulties in obtaining sufficient numbers of antigen-specific cells, so the use of a polyclonal activation model with bulk PBMC represents the standard to study Treg suppression. Rapid tests for Treg cell function have been described using short-term (7–20 hr) flow-based assays to measure suppression of T-cell activation marker expression (CD40L and CD69). These assays utilize effector T cells (CD4$^+$25$^-$) activated with anti-CD3/anti-CD28 beads with and without the addition of freshly isolated Treg cells or *ex vivo* expanded Treg cells.[36,37] When implementing Treg suppression assays in the clinical diagnostic laboratory, a number of technical considerations can confound the interpretation of results, and these are well described by McMurchy *et al.*[38] Also, it remains unclear whether standard *in vitro* Treg suppression assays can adequately account for the significant complexity of Treg subsets that may have different functional properties in various clinical contexts. Therefore the newer techniques of mass cytometry may be required to address the phenotypic diversity[39] and to eventually be harnessed either to evolving applications or other technology to analyze the functional complexity of these populations.

ASSESSMENT OF SIGNALING AND DNA REPAIR PATHWAYS IN LYMPHOCYTES VIA PHOSPHOFLOW CYTOMETRY

A key aspect of studying lymphocyte responses is to assess signaling in appropriate lymphocyte subsets and its alteration in pathological conditions. A versatile tool called phosphoflow is unique in its ability to allow the study of multiple intracellular signaling molecules in specific lymphocyte populations at a single-cell level.[40] Phosphoflow assays for the signal transducer and activator of transcription (STAT) molecules[41,42] (Fig. 93.12) have been well described and are discussed in Chapter 92. The use of phosphoflow assays to assess Bruton's tyrosine kinase (BTK) phosphorylation in X-linked agammaglobulinemia (XLA) with leaky (hypomorphic) defects as well as to assess radiosensitivity and the DNA repair pathway will be covered briefly.

Mutations in *BTK* impair B-cell maturation and function, and patients with XLA can have either no peripheral B cells (null mutations) or reduced B cells (hypomorphic/leaky mutations), depending on the specific genetic defect (Chapter 34). Two regulatory tyrosine residues in BTK undergo rapid phosphorylation upon B-cell receptor (BCR) cross-linking (Y551 within the SH1 domain and Y223 in the SH3 domain).[43] Y551 is transphosphorylated by Src family kinases whereas Y223 is autophosphorylated. Phospho-BTK accounts for a small fraction (<5%) of the total BTK pool in BCR-activated B cells. BTK molecules are singly or doubly tyrosine-phosphorylated, and Y551 phosphorylation increases BTK activity, whereas Y223 phosphorylation is likely more relevant for protein-protein interactions due to its presence in the SH3 domain. Since the phosphorylation events are sequential (Y551 before Y223), the dephosphorylation event is also successive in the same order. In the flow assay, Y223 phosphorylation is measured after an anti-IgM antibody is used

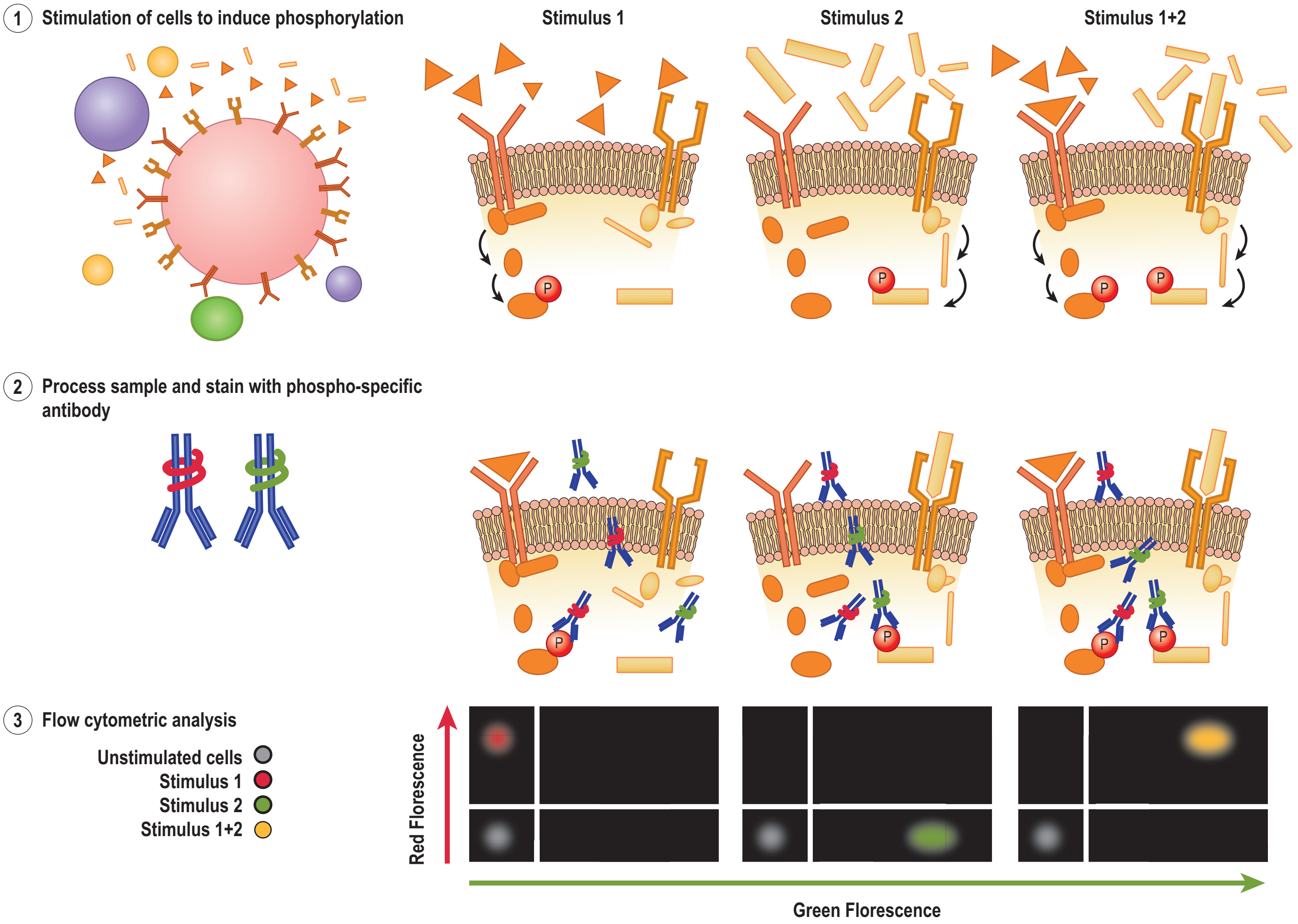

1 Stimulation of cells to induce phosphorylation
Stimulus 1
Stimulus 2
Stimulus 1+2
P
2 Process sample and stain with phospho-specific antibody
3 Flow cytometric analysis
Unstimulated cells
Stimulus 1
Stimulus 2
Stimulus 1+2
Red Florescence
Green Florescence

FIG 93.12 Analysis of Phosphoproteins in Cells by Flow Cytometry. Cells activated by specific stimuli can be assessed for activation and induction of specific signaling pathways by measuring phosphoproteins: A single stimulus or multiple stimuli can be used to phosphorylate different proteins, and intracellular staining can be performed with phospho-specific antibodies, which is subsequently analyzed by multicolor flow cytometry. See text for detailed explanation. (Figure modified from Figure 2 by Krutzik PO, et al. *Clin Immunol* 2004;110:206–21; published in previous edition of this book.)

KEY CONCEPTS

Assessment of DNA Repair Pathways by Phosphoflow

- DNA double-strand break (DSB) repair defects are relevant to VDJ recombination, isotype class switching, and lymphocyte maturation.
- Defects in this process can lead to immunodeficiency with susceptibility to infection and malignancy.
- Homologous recombination (HR), which is error-free, involves RAD50, RAD51, RAD52, and Mre11, while nonhomologous end-joining (NHEJ) is error-prone and utilizes Ku70/80, DNA-PKcs, Artemis, DNA Ligase IV, XRCC4, XLF/Cernunnos.
- ATM is a key regulator of cell-cycle checkpoints following irradiation-induced DSB and coordinates timing of phosphorylation of checkpoint proteins.
- Cell cycle analysis can help identify checkpoint defects and can be easily assessed by flow cytometry.
- Phosphorylation of H2AX (γH2AX) is a useful marker of DNA damage related to irradiation-induced DSB.
- However, several kinases phosphorylate H2AX, which can confound identification of specific defects; therefore assessing multiple proteins in the pathway allows identification of a broader group of DNA repair defects.
- Measurement of 53BP1 phosphorylation is a useful surrogate marker of the integrity of the chromatin ubiquitin ligase pathway.

to cross-link the BCR (for 3 minutes), because Y551 phosphorylation could not be detected in the time interval the assay was performed. In addition to PBMC, Ramos cell line (a B-cell line derived from a patient with Burkitt lymphoma) is used as a control for B cells, while pervanadate (complex of vanadate with hydrogen peroxide), an irreversible protein tyrosine-phosphatase inhibitor, is used in this assay as a positive control (Fig. 93.13). In the Ramos cell line, it is possible to visualize the Y551 phosphorylation (Fig. 93.13) with pervanadate treatment.

A number of genetic disorders, collectively classified as XCIND (x-ray [irradiation] sensitivity, cancer susceptibility, immunodeficiency, neurological involvement, and double-strand DNA breakage), cause impairment in cellular ability to repair DNA double-strand breaks (DSB). These defects have a significant impact on the ability of cells to grow, differentiate, and function normally; they include ataxia telangiectasia (AT) due to *ATM* gene mutations and radiosensitive severe combined immunodeficiencies (rs-SCID) due to mutations in the *DCLRE1C, LIG4, NHEJ1,* and *PRKDC* genes. These disorders make it very desirable to have an assay capable of rapidly assessing DNA repair in lymphocytes in response to radiation damage. A flow cytometry assay has been developed that is capable of measuring the function of several proteins in the DNA repair pathway after induction of DSBs via irradiation (Fig. 93.14A). The assay includes assessment of nonphosphorylated ATM protein and 53BP1[44,45] without irradiation to determine the amount of native protein present. Following low-dose irradiation, the function of ATM and ATR (ATM-Rad3–related kinase) pathways are assessed by analyzing the autophosphorylation of ATM at serine 1981. This step is required for ATM activation and is followed by phosphorylation of downstream targets SMC1, H2AX, 53BP1, and CHK1. H2AX is a histone that belongs to the H2A family and is a component of the histone octamer in nucleosomes. It is phosphorylated by both ATM and ATR and is the first step in the recruitment and localization of DNA repair proteins.[46] In patients with AT, there is a complete absence of phosphorylation of ATM, SMC1, and 53BP1 after irradiation-induced DSBs. Phosphorylation of H2AX appears normal, as it is also phosphorylated by ATR; however, the magnitude of phosphorylation is significantly decreased (Fig. 93.14B). The kinetics of phosphorylation indicates that maximal phosphorylation occurs 1 hr after induction of DSB and that there is dephosphorylation by 24 hr postirradiation (not shown). This flow-based assay enables a rapid assessment of radiation sensitivity as well as the ability to visualize the function of multiple proteins of the DNA repair pathway. It can also be used to characterize the DNA repair function of known and unknown genetic defects, and it can identify functional phenotypes in patients with atypical presentations that may be missed if only genetic information is used.

ASSESSMENT OF B-CELL FUNCTION

While T cells and NK cells form the foundation of the cellular immune response, B cells are the main driver of humoral immunity. B cells are multifaceted in their function; they produce antibodies via differentiation into plasma cells, they act as APCs for T cells, and they secrete potent immunomodulatory cytokines while downregulating immune responses via IL-10 production. B cells can proliferate in response to polyclonal mitogenic stimuli such as PWM (Fig. 93.15), albeit with a much weaker response than seen with T cells stimulated with PHA (Fig. 93.3). Clearly, the starting point for any evaluation of B-cell function involves measuring serum immunoglobulin levels followed by *in vivo* antibody responses to vaccination with both protein (*e.g.,* tetanus toxoid) and carbohydrate (*e.g.,* Pneumovax 23) antigens.[47]

A more recent area of focus in laboratory evaluation of B cells has been the regulatory function of B cells, specifically with regard to the subset classified as regulatory B (Breg) cells. B cells that secrete IL-10 have been described as Breg cells and are now recognized to be an important component of the host immune response that protects against autoimmunity and also limits inflammatory damage.[48,49] In humans, the phenotype of Breg cells has been described as $CD19^+CD27^-CD24^{hi}CD38^{hi}CD5^+CD1d^{hi}$. However, the ability to produce IL-10 appears to be the defining feature rather than any particular constellation of cell-surface markers.[50] Also, unlike Treg cells, no exclusive transcription factor equivalent to FOXP3 is expressed in Breg cells. In general, $IL\text{-}10^+$ Breg cells in blood and spleen appear to be enriched in the $CD19^+CD27^+CD24^{hi}$ memory B-cell subset. IL-10 exerts potent antiinflammatory effects and enhances survival, proliferation, differentiation, and isotype class-switching of B cells. Both naïve and memory human B cells have the ability to produce IL-10 in response to stimulation via TLR9 and the BCR, but only ~15% of B cells can produce IL-10 in response to stimulation. Stimulation through CD40L or IL-21 also promotes differentiation of IL-10–producing Breg cells. $CD19^+CD24^{hi}CD38^{hi}$ Breg cells play critical regulatory roles in autoimmune disease, chronic GvHD, and transplant tolerance.[51-53]

In the diagnostic laboratory, Breg cells can be assessed by isolating PBMC from blood and culturing with CpG-B (TLR9 stimulation) or CpG-B plus recombinant CD40L for 3 days *in vitro,* followed by the addition of PMA, ionomycin, and brefeldin A (BFA) for the last 5 hours of culture. Cells can then be harvested, washed, and stained with CD19 and IL-10 antibodies (the latter requires intracellular staining) and then analyzed with a flow cytometer. To assess the presence of Breg cells in blood without *in vitro* differentiation, blood or PBMC can be analyzed for $CD19^+CD24^{hi}CD38^{hi}$ B cells that are positive for intracellular

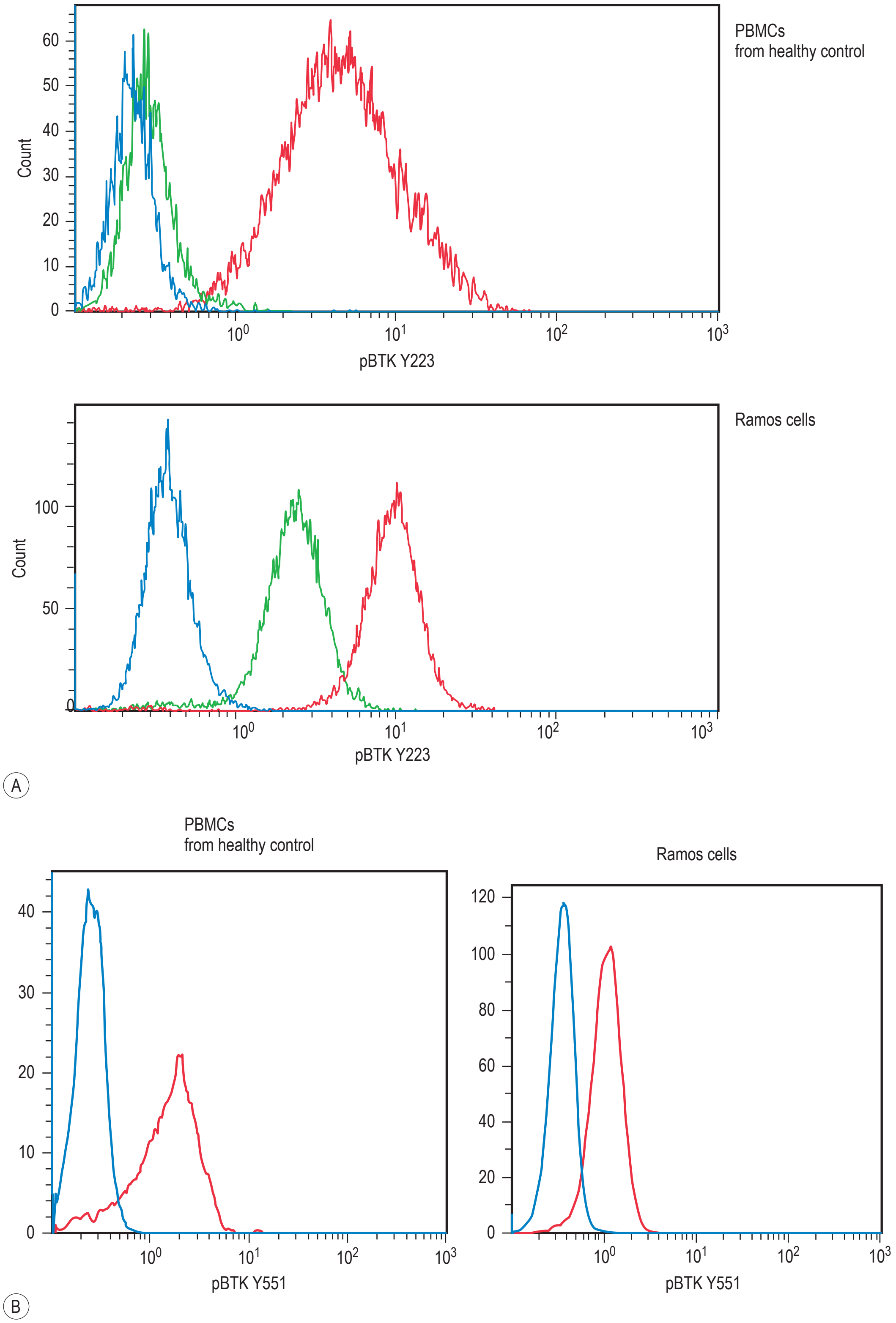


FIG 93.13A, B Analysis of BTK Phosphorylation in PBMC and Ramos B-Cell Line. Bruton's tyrosine kinase (BTK) phosphorylation can be assessed in B cells in X-linked agammaglobulinemia (XLA) patients with leaky defects who have some preserved B cells in blood. B cells are stimulated with either anti-immunoglobulin M (IgM) (B-cell receptor [BCR] cross-linking; green line) or pervanadate (positive control; red line). The example shown here is for BTK phosphorylation at the Y223 residue in peripheral blood mononuclear cells (PBMCs) from a healthy donor (top panel) and a B-cell line, Ramos cells. The blue line represents the unstimulated control. In the bottom panel, BTK phosphorylation of the Y551 residue is shown for PBMCs from a healthy donor (left) and Ramos cells (right). The phosphorylation of the Y551 residue cannot be visualized for anti-IgM stimulation, and therefore data are shown only for the positive control, pervanadate.

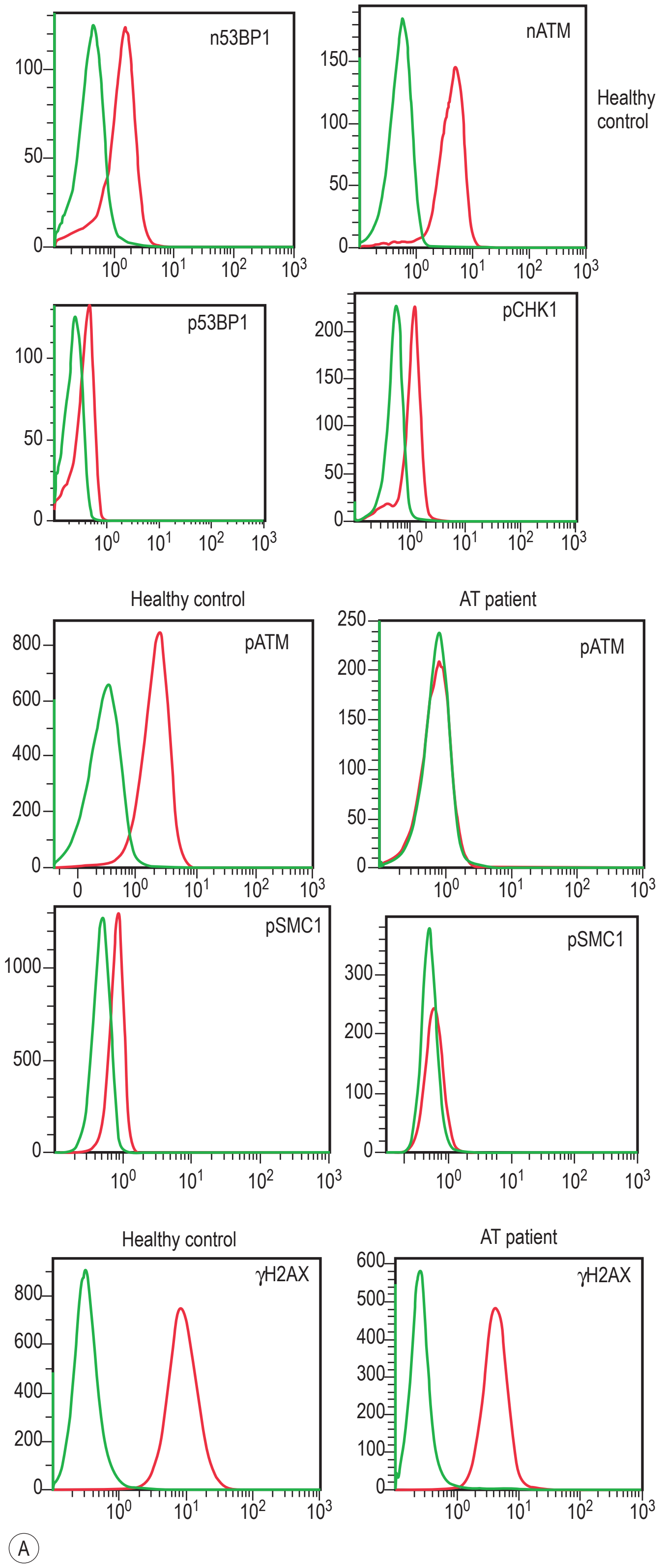
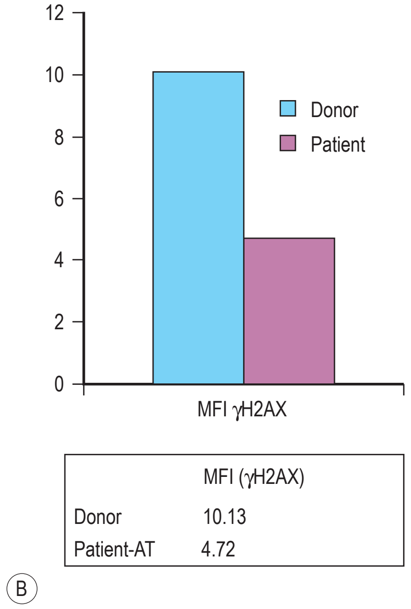


FIG 93.14A **Assessment of Radiosensitivity in Lymphocytes by Flow Cytometry.** DNA repair pathway defects can be rapidly and sensitively analyzed by flow cytometry. The top panel shows expression of native (nonphosphorylated) 53BP1 and ATM proteins and phosphorylated (p) 53BP1 and CHK1 proteins in total lymphocytes. The lower panels demonstrates normal ATM autophosphorylation in T cells of a healthy donor and absent (abnormal) phosphorylation of ATM in T cells of a patient with ataxia telangiectasia (AT) on exposure to low-dose (2 Gy) radiation. The green line represents the unirradiated sample, and the red line represents the data postirradiation. AT patients also show an inability of ATM to phosphorylate downstream targets, such as SMC1 in T cells. This is also true for other lymphocyte subsets (B cells and natural killer [NK] cells). Histone H2AX is phosphorylated (γH2AX) as a result of DNA double-strand breaks (DSBs) from exposure to radiation. No defect is apparent in γH2AX in AT when assessing the frequency of cells that express γH2AX.



	MFI (γH2AX)
Donor	10.13
Patient-AT	4.72

FIG 93.14B **The Relevance of Assessing Mean Fluorescence Intensity for Flow-Based Radiation Sensitivity Analyses.** Mean fluorescence intensity (MFI) can provide additional valuable information on DNA repair defects and should be used along with frequency analysis, as demonstrated in Fig. 93.14A. In ataxia telangiectasia (AT) patients, although the proportion of γH2AX appears normal, it is significantly decreased (by >50% in AT (purple bar) compared with healthy controls (blue bar).

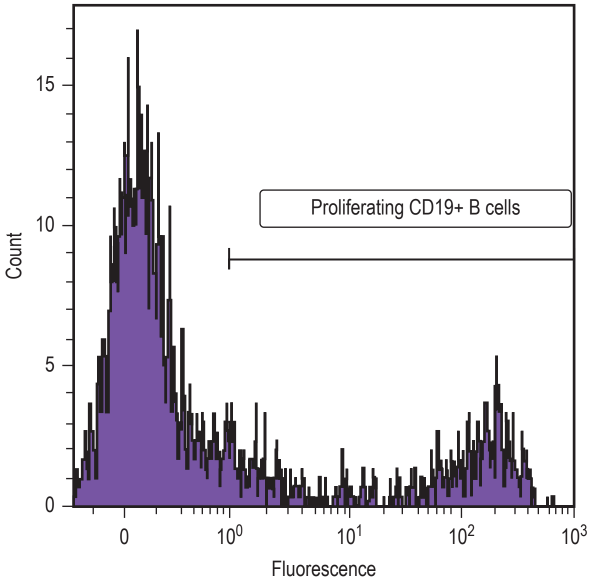


FIG 93.15 Assessment of B-Cell Function by Measuring Proliferation to Mitogen. B cells can proliferate on stimulation with certain mitogens, similar to T cells. B cells do not respond well to phytohemagglutinin (PHA) but can proliferate *in vitro* on stimulation with pokeweed mitogen (PWM), although it is a relatively weak B-cell stimulant. B-cell proliferation can be assessed by flow-based proliferation in CD19⁺ B cells.

IL-10 using a similar flow cytometry protocol, but the levels of these cells can be highly variable, and accuracy of detection and quantitation is likely to be dependent on the analytical method and clinical context.*

SUMMARY

ON THE HORIZON

- The advent of newer assays, including flow cytometry based as described herein, are likely to replace traditional and often less sensitive radioactive or other cumbersome methods for assessing lymphocyte function.
- Microchip and advanced mass cytometry techniques, along with imaging cytometry, may offer newer, multiplex approaches to the assessment of lymphocyte subsets, individually and in cellular interactions.
- Functional and phenotypic data need to be ideally correlated with relevant genomic, transcriptomic, proteomic, metabolomic, epigenomic, and microbiome analyses for effective characterization of the complex interactions that govern the immune response in the normal and dysregulated state.
- This will necessitate further refinements of "big data" analysis, which can include, but is not limited to, experimental studies in addition to computational modeling. Examples include antigen-specific characterization of individual B-cell and T-cell receptors in an individual as an extension of current "deep-sequencing" repertoire analysis or transcriptomic analysis during immune quiescence or activation on a multidimensional scale.

In conclusion, lymphocyte responses in humans can be assessed using a variety of analytical tools as described in this chapter and applied to many clinical contexts, including but not limited to primary immunodeficiencies, autoimmunity, transplantation, and immune dysregulation disorders. Most of these measurements use cells from blood and not from other lymphoid tissues due to accessibility and ease of generating control reference ranges. This chapter is not an exhaustive treatise on all aspects of lymphocyte function and the immune response; nor does it cover all areas of normal and abnormal pathology in this context. Rather, it is meant to take the reader on a tour of the immune landscape and to provide salient highlights of the diversity and relevance of lymphocyte function.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Puga I, Cols M, Barra CM, et al. B cell-helper neutrophils stimulate the diversification and production of immunoglobulin in the marginal zone of the spleen. Nat Immunol 2012;13(2):170–80.
2. Vinuesa CG, Chang PP. Innate B cell helpers reveal novel types of antibody responses. Nat Immunol 2013;14(2):119–26.
3. Torgerson TR, Ochs HD. Immune dysregulation, polyendocrinopathy, enteropathy, X-linked: forkhead box protein 3 mutations and lack of regulatory T cells. J Allergy Clin Immunol 2007;120(4):744–50, quiz 51–2.
4. Sakaguchi S, Wing K, Onishi Y, et al. Regulatory T cells: how do they suppress immune responses? Int Immunol 2009;21(10):1105–11.
5. Long EO, Kim HS, Liu D, et al. Controlling natural killer cell responses: integration of signals for activation and inhibition. Annu Rev Immunol 2013;31:227–58.
6. Buck MD, O'Sullivan D, Pearce EL. T cell metabolism drives immunity. J Exp Med 2015;212(9):1345–60.
7. Wherry EJ, Kurachi M. Molecular and cellular insights into T cell exhaustion. Nat Rev Immunol 2015;15(8):486–99.
8. Abraham RS. Lymphocyte Activation. In: Detrick B, Schmitz JL, Hamilton RG, editors. Manual of Molecular and Clinical Laboratory Immunology. 8th ed. Washington, DC: ASM Press; 2016. p. 269–79.
9. Abraham RS, Aubert G. Flow Cytometry, a Versatile Tool for Diagnosis and Monitoring of Primary Immunodeficiencies. Clin Vaccine Immunol 2016;23(4):254–71.
10. Appay V. The physiological role of cytotoxic CD4(+) T-cells: the holy grail? Clin Exp Immunol 2004;138(1):10–13.
11. van Leeuwen EM, Remmerswaal EB, Heemskerk MH, et al. Strong selection of virus-specific cytotoxic CD4+ T-cell clones during primary human cytomegalovirus infection. Blood 2006;108(9):3121–7.
12. Snyder JE, Bowers WJ, Livingstone AM, et al. Measuring the frequency of mouse and human cytotoxic T cells by the Lysispot assay: independent regulation of cytokine secretion and short-term killing. Nat Med 2003;9(2):231–5.
13. Mbitikon-Kobo FM, Bonneville M, Sekaly RP, et al. Ex vivo measurement of the cytotoxic capacity of human primary antigen-specific CD8 T cells. J Immunol Methods 2012;375(1-2):252–7.
14. Walker S, Fazou C, Crough T, et al. Ex vivo monitoring of human cytomegalovirus-specific CD8+ T-cell responses using QuantiFERON-CMV. Transpl Infect Dis 2007;9(2):165–70.
15. Colonna M, Jonjic S, Watzl C. Natural killer cells: fighting viruses and much more. Nat Immunol 2011;12(2):107–10.
16. Vivier E, Raulet DH, Moretta A, et al. Innate or adaptive immunity? The example of natural killer cells. Science 2011;331(6013):44–9.
17. Chan CJ, Smyth MJ, Martinet L. Molecular mechanisms of natural killer cell activation in response to cellular stress. Cell Death Differ 2014;21(1):5–14.

---

*Portions of this section have been reproduced with permission from ASM Press (Abraham RS, Lymphocyte Activation) (copyright permission obtained)

18. Sun JC, Lanier LL. NK cell development, homeostasis and function: parallels with CD8(+) T cells. Nat Rev Immunol 2011;11(10):645–57.
19. Yu TK, Caudell EG, Smid C, et al. IL-2 activation of NK cells: involvement of MKK1/2/ERK but not p38 kinase pathway. J Immunol 2000;164(12):6244–51.
20. Gasteiger G, Hemmers S, Firth MA, et al. IL-2-dependent tuning of NK cell sensitivity for target cells is controlled by regulatory T cells. J Exp Med 2013;210(6):1167–78.
21. Somanchi SS, McCulley KJ, Somanchi A, et al. A Novel Method for Assessment of Natural Killer Cell Cytotoxicity Using Image Cytometry. PLoS ONE 2015;10(10):e0141074.
22. Guldevall K, Brandt L, Forslund E, et al. Microchip Screening Platform for Single Cell Assessment of NK Cell Cytotoxicity. Front Immunol 2016;7:119.
23. Jang YY, Cho D, Kim SK, et al. An improved flow cytometry-based natural killer cytotoxicity assay involving calcein AM staining of effector cells. Ann Clin Lab Sci 2012;42(1):42–9.
24. Alter G, Malenfant JM, Altfeld M. CD107a as a functional marker for the identification of natural killer cell activity. J Immunol Methods 2004;294(1-2):15–22.
25. Betts MR, Brenchley JM, Price DA, et al. Sensitive and viable identification of antigen-specific CD8+ T cells by a flow cytometric assay for degranulation. J Immunol Methods 2003;281(1-2):65–78.
26. Bryceson YT, Pende D, Maul-Pavicic A, et al. A prospective evaluation of degranulation assays in the rapid diagnosis of familial hemophagocytic syndromes. Blood 2012;119(12):2754–63.
27. Yamashita M, Kitano S, Aikawa H, et al. A novel method for evaluating antibody-dependent cell-mediated cytotoxicity by flowcytometry using cryopreserved human peripheral blood mononuclear cells. Sci Rep 2016;6:19772.
28. Ohkura N, Kitagawa Y, Sakaguchi S. Development and maintenance of regulatory T cells. Immunity 2013;38(3):414–23.
29. Sakaguchi S, Miyara M, Costantino CM, et al. FOXP3+ regulatory T cells in the human immune system. Nat Rev Immunol 2010;10(7):490–500.
30. Schmitt EG, Williams CB. Generation and function of induced regulatory T cells. Front Immunol 2013;4:152.
31. Barzaghi F, Passerini L, Bacchetta R. Immune dysregulation, polyendocrinopathy, enteropathy, x-linked syndrome: a paradigm of immunodeficiency with autoimmunity. Front Immunol 2012;3:211.
32. Beres AJ, Drobyski WR. The role of regulatory T cells in the biology of graft versus host disease. Front Immunol 2013;4:163.
33. Li XC, Turka LA. An update on regulatory T cells in transplant tolerance and rejection. Nat Rev Nephrol 2010;6(10):577–83.
34. Waldmann H, Hilbrands R, Howie D, et al. Harnessing FOXP3+ regulatory T cells for transplantation tolerance. J Clin Invest 2014;124(4):1439–45.
35. Boks MA, Zwaginga JJ, van Ham SM, et al. An optimized CFSE-based T-cell suppression assay to evaluate the suppressive capacity of regulatory T-cells induced by human tolerogenic dendritic cells. Scand J Immunol 2010;72(2):158–68.
36. Canavan JB, Afzali B, Scotta C, et al. A rapid diagnostic test for human regulatory T-cell function to enable regulatory T-cell therapy. Blood 2012;119(8):e57–66.
37. Ruitenberg JJ, Boyce C, Hingorani R, et al. Rapid assessment of in vitro expanded human regulatory T cell function. J Immunol Methods 2011;372(1-2):95–106.
38. McMurchy AN, Levings MK. Suppression assays with human T regulatory cells: a technical guide. Eur J Immunol 2012;42(1):27–34.
39. Mason GM, Lowe K, Melchiotti R, et al. Phenotypic Complexity of the Human Regulatory T Cell Compartment Revealed by Mass Cytometry. J Immunol 2015;195(5):2030–7.
40. Krutzik PO, Trejo A, Schulz KR, et al. Phospho flow cytometry methods for the analysis of kinase signaling in cell lines and primary human blood samples. Methods Mol Biol 2011;699:179–202.
41. Fleisher TA, Dorman SE, Anderson JA, et al. Detection of intracellular phosphorylated STAT-1 by flow cytometry. Clin Immunol 1999;90(3):425–30.
42. Murphy J, Goldberg GL. Flow cytometric analysis of STAT phosphorylation. Methods Mol Biol 2013;967:161–5.
43. Wahl MI, Fluckiger AC, Kato RM, et al. Phosphorylation of two regulatory tyrosine residues in the activation of Bruton's tyrosine kinase via alternative receptors. Proc Natl Acad Sci USA 1997;94(21):11526–33.
44. Panier S, Boulton SJ. Double-strand break repair: 53BP1 comes into focus. Nat Rev Mol Cell Biol 2014;15(1):7–18.
45. Paull TT. Mechanisms of ATM Activation. Annu Rev Biochem 2015;84:711–38.
46. Kuo LJ, Yang LX. Gamma-H2AX - a novel biomarker for DNA double-strand breaks. In Vivo 2008;22(3):305–9.
47. Abraham RS. Relevance of antibody testing in patients with recurrent infections. J Allergy Clin Immunol 2012;130(2):558–9, e6.
48. Mauri C, Bosma A. Immune regulatory function of B cells. Annu Rev Immunol 2012;30:221–41.
49. Rosser EC, Mauri C. Regulatory B cells: origin, phenotype, and function. Immunity 2015;42(4):607–12.
50. Lin W, Cerny D, Chua E, et al. Human regulatory B cells combine phenotypic and genetic hallmarks with a distinct differentiation fate. J Immunol 2014;193(5):2258–66.
51. Flores-Borja F, Bosma A, Ng D, et al. CD19+CD24hiCD38hi B cells maintain regulatory T cells while limiting TH1 and TH17 differentiation. Sci Transl Med 2013;5(173):173ra23.
52. Nouel A, Simon Q, Jamin C, et al. Regulatory B cells: an exciting target for future therapeutics in transplantation. Front Immunol 2014;5:11.
53. Sarantopoulos S, Blazar BR, Cutler C, et al. B cells in chronic graft-versus-host disease. Biol Blood Marrow Transplant 2015;21(1):16–23.

MULTIPLE-CHOICE QUESTIONS

1. CD40L is an early activation marker present primarily on which lymphocyte subset?
   A. B cells
   B. Natural killer (NK) cells
   C. Regulatory T cells
   D. CD4 T cells

2. Natural killer (NK) cells recognize target cells in a cytotoxicity assay by which of the following mechanisms?
   A. Absence of CD16
   B. Absence of NKp46
   C. Absence of major histocompatability complex (MHC) class I
   D. Absence of MHC class II

3. The histone H2AX is phosphorylated (γH2AX) by which of the following kinases, in addition to ATM?
   A. BTK
   B. LCK
   C. ATR
   D. ITK

4. Which of the following cytokines is a hallmark of B cells that possess regulatory function (Breg)?
   A. Interleukin-6 (IL-6)
   B. IL-10
   C. IL-2
   D. IL-21

94

Assessment of Neutrophil Function

Debra Long Priel, Douglas B. Kuhns

Neutrophils, also known as polymorphonuclear neutrophils (PMNs; because of their multilobed nucleus) or granulocytes (because of the numerous granules found in the cytoplasm), are major contributors to innate host defense against invading microorganisms, particularly bacteria and fungi. Neutrophils are bone marrow–derived, terminally differentiated cells that are incapable of further cellular division but may be also be derived from progenitor populations in the spleen.[1] Early studies indicated that neutrophils have a short life span in the circulation ($t_{1/2}$= 6–8 hours) and then survive an additional 1–2 days in the surrounding tissue;[2] more recent data suggest that neutrophils may survive 10 times longer in circulation, up to 5.4 days.[3]

Neutrophils, with a diameter of 10–15 μm and a volume of 346 μm^3, have a unique morphology. The nucleus of the mature neutrophil is segmented into 3–5 lobes with chromosomes randomly distributed among the lobes. Neutrophils also have an extensive array of storage granules prepackaged with specific proteins defined by the differentiation stage during maturation.[4] Granules are classified into four distinct populations: azurophilic, specific, gelatinase, and secretory granules. Azurophilic granules contain myeloperoxidase, lysozyme, antimicrobial peptides, defensins, proteases, and the lysosomal acid hydrolases. The specific granules contain lactoferrin, lysozyme, and vitamin B_{12}–binding protein and also serve as storage pools for CD11b/CD18 and cytochrome b_{558} of the superoxide anion radical ($O_2^{\bullet-}$)–generating enzyme, nicotinamide adenine dinucleotide phosphate (NADPH) oxidase, or NOX2. The gelatinase granules are a subset of the specific granules that have a high content of gelatinase. The secretory granules are highly mobilizable intracellular vesicles that contain alkaline phosphatase and other surface antigens.

The primary function of neutrophils is the ingestion (phagocytosis) and subsequent killing of microorganisms. This process requires the assembly of NOX2, a multiphagocyte oxidase (phox) component enzyme complex consisting of at least three cytosolic components—p47^{phox}, p67^{phox}[5], and p40^{phox}[6] —and two membrane components—p22^{phox} and gp91^{phox}—that constitute cytochrome b_{558}.[7,8] This enzyme reduces molecular O_2 to $O_2^{\bullet-}$ using NADPH generated by the oxidation of glucose through the pentose phosphate pathway; $O_2^{\bullet-}$ either spontaneously or enzymatically converts to hydrogen peroxide (H_2O_2). In the presence of a metal, such as iron (Fe^{2+}), H_2O_2 and $O_2^{\bullet-}$ can react to form the highly reactive hydroxyl radical, $OH^{\bullet}$. Alternatively, the azurophilic granule constituent, myeloperoxidase, catalyzes the formation of hypochlorous acid from H_2O_2 and chloride (Cl^-). The combined activities of these reactive O_2 species (ROS), antimicrobial peptides, and lysosomal hydrolases result in the ultimate destruction of the ingested microorganism. Excess production of ROS and release of lysosomal hydrolases into the extracellular milieu can lead to tissue damage and inflammation.

Neutrophils can also contribute extracellular microbicidal activity through the formation of neutrophil extracellular traps (NETs), a matrix of DNA and granular enzymes that is purported to entrap bacteria and promote their killing.[9] During NET formation, the nucleus loses its lobular shape, the nuclear membrane disintegrates into a chain of vesicles surrounding the DNA, and granular integrity is lost. The nuclear material fills most of the cell, mixing with the granular contents. The cells round up and DNA is forcibly extruded from the cell, conveying with it granular enzymes trapped within the DNA matrix. Mitochondria in neutrophils are few in number and exhibit relatively little oxidative phosphorylation but may play an important role in NET formation. Activation of neutrophils by immune complexes results in marked depolarization of the mitochondria, increased mitochondrial ROS production, and redistribution of the mitochondria to the periphery of the neutrophil. A low-density subset of neutrophils may be more prone to NET formation.[10]

Neutrophils display a diverse array of cellular functions. Abnormalities in these functions can severely compromise host defense, leading to recurrent bacterial and fungal infections (Chapter 22). To localize specific deficiencies of neutrophil function, assays have been developed that mimic these functions both *in vivo* and *in vitro.* Often, a preliminary screening of several neutrophil functions is performed to localize deficits and then more vigorous testing of specific function is performed. Assays to assess neutrophil function should address several limitations—the number of cells required for the assay, the type of cell preparation needed (isolated neutrophils versus whole blood), the overall incubation time for the assay, the complexity of the assay, and the rapidity of data collection. These issues become more critical if multiple functional assays are planned concurrently. Since neutrophils cannot be stored or frozen and maintain viability, neutrophils from a normal subject are generally assayed in parallel to validate the results, doubling the number of assays to be performed. Additionally, isolation of neutrophils can take 1–2 hours, limiting the time available for functional assays. Fluorescent probes have increased the sensitivity of many of the assays and eliminated the need for radioactive probes. The use of multiwell microplates and microplate readers has reduced the number of cells required and has facilitated the collection of data. Experience in the handling of neutrophils and the time constraints of assays can limit the availability of this testing to laboratories that specialize in assessment of neutrophil function.

KEY CONCEPT

Criteria to Assess Neutrophil Function

Because of (1) the time required to isolate neutrophils and (2) the shortened life span of neutrophils after isolation, assays of neutrophil function should have minimal complexity and enable rapid data collection.

KEY CONCEPT

Estimated Yield from Whole Blood

In general, 1–2 × 10^6 neutrophils can be isolated per milliliter of whole blood from a normal subject with a normal white blood cell count.

ISOLATION OF NEUTROPHILS

Clinical Indications and Implications

Assays that avoid neutrophil isolation are preferred because of the artificial priming of neutrophils during isolation.[11] However, most assays require isolated neutrophils to eliminate any possible contributions of other leukocytes and blood components. In general, blood should be drawn using either citrate or heparin as anticoagulant and maintained at 20–25°C in polypropylene containers. Most isolation protocols require 1–2 hours to obtain purified neutrophils.

Principles and Interpretation of Laboratory Assessment

Most neutrophil isolation protocols use differences in the cell density as the basis for the separation. The relative densities of blood cells are as follows: erythrocytes > neutrophils and eosinophils > monocytes, lymphocytes, and basophils > platelets. Ficoll-Paque is a solution of sodium diatrizoate (a dense, triiodinated compound) and Ficoll (a polysaccharide) with a density (1.077 g/cc) that falls between the density of neutrophils and that of the mononuclear cells. To isolate neutrophils,[12] whole blood is diluted with saline and underlaid with Ficoll-Paque solution. After centrifugation for 30 minutes at 500 g, the less-dense monocytes, lymphocytes, basophils, and platelets remain at the upper interface of the Ficoll-Paque solution, whereas the denser erythrocytes and neutrophils pass through the solution and pellet at the bottom. The mononuclear cells are carefully harvested and the remaining Ficoll-Paque solution aspirated. The erythrocyte/neutrophil pellet is resuspended with saline and mixed with 3% dextran. Dextran promotes the formation of rouleaux by the erythrocytes, causing them to sediment more rapidly than the neutrophils at 1 g. The neutrophil-enriched supernatant fluid is harvested from the bulk of the erythrocytes. Contaminating erythrocytes are removed by a brief (30-second) hypotonic lysis with 0.2% saline. The isotonicity is quickly restored with an equal volume of 1.6% saline. A second hypotonic lysis removes many of the red blood cell (RBC) ghosts. In general, 1–2 × 10^6 neutrophils can be isolated per milliliter of whole blood from a normal subject with a normal white blood cell (WBC) count. All procedures are performed at room temperature, and the isolated cells are maintained in a balanced salt solution without divalent cations. The most common cell contaminants of the neutrophil preparation are eosinophils. Further purification of a standard neutrophil preparation with anti-CD16 magnetic immunobeads results in a neutrophil preparation that is generally ≥99% neutrophils. A second neutrophil isolation protocol that uses a discontinuous gradient of plasma/Percoll has often been used to minimize exposure of neutrophils to trace contamination by bacterial lipopolysaccharide (LPS) and reduce neutrophil priming.[13]

Isolated neutrophils are routinely frozen in aliquots of 5 × 10^6 cells/vial. For Western blot studies, neutrophils (1 × 10^6 cells/mL of buffer) are pretreated for 20 minutes with the cell permeant, irreversible serine protease inhibitor, diisopropylfluorophosphate (DFP, 1–5 mM). DFP is a volatile, potent neurotoxin that can irreversibly bind to and inactivate acetylcholinesterase and should be used with extreme caution. The cell suspension is then spun and the supernatant fluid removed from the cell pellet before freezing. Waste solutions and disposable laboratory items should be flushed with sodium hydroxide to inactivate any residual DFP. These frozen neutrophil pellets, though not viable, can also be a source of DNA for genetic analyses.

HISTOCHEMICAL ANALYSIS OF NEUTROPHILS

Clinical Indications and Implications

Because of their unique morphology, microscope examination of neutrophil preparations with a differential stain, such as Wright stain, or a histochemical stain, such as Kaplow stain, for myeloperoxidase[14] remains an essential element of neutrophil study and can provide valuable insight into some genetic immunodeficiencies.

Principle and Interpretation of Laboratory Assessment

In Wright's stain, the nucleus of a segmented neutrophil is normally multilobed (usually 2–5 lobes), and each lobe is connected by a narrow filament (Fig. 94.1A). The nuclear chromatin is coarsely clumped with purple staining. Nucleoli are generally not present. In a band neutrophil, the nucleus is horseshoe shaped, with no indication of constriction into lobes. The pink-violet staining of the cytosol is associated with numerous, evenly distributed, specific granules; occasionally a dark-staining primary granule may be present. Kaplow stain[14] identifies the myeloperoxidase-containing primary granules as dark blue granules uniformly distributed throughout the cytosol (Fig. 94.1B). Neutrophils (and platelets) from patients with Chédiak-Higashi syndrome have giant primary granules that are generally considered pathognomonic for the disease (see Fig. 94.1C).[15] Myeloperoxidase staining of Chédiak-Higashi neutrophils is very distinctive, with staining localized to the discrete giant primary granules (see Fig. 94.1D). Neutrophils from a patient with specific granule deficiency exhibit primarily bilobed nuclei (pseudo–Pelger-Huët anomaly) with a paucity of specific granule staining in the cytosol (see Fig. 94.1E).[16] Staining of the myeloperoxidase granules of neutrophils from a patient with specific granule deficiency appears normal, since the defect is primarily associated with the specific granules (see Fig. 94.1F). Neutrophils from a patient with myeloperoxidase deficiency fail to stain for myeloperoxidase in the neutrophils. However, the peroxidase in the eosinophilic granules continues to stain positive (see Fig. 94.1G).[17]

ANALYSIS OF GRANULE CONSTITUENTS

Clinical Indications and Implications

The granules of the neutrophils can be distinguished by their specific contents. Deficiency of only one granule constituent can be associated with a specific genetic defect, such as myeloperoxidase deficiency; alternatively, deficiency of multiple constituents

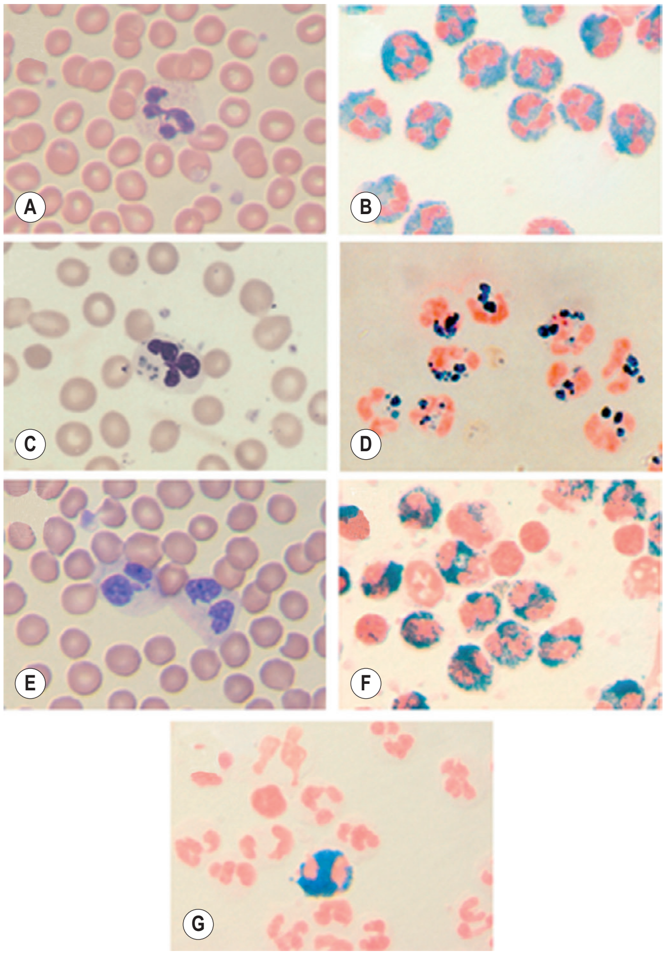


FIG 94.1 **Neutrophils Stained With Wright Stain and Kaplow Stain.** (A, C, and E) Blood smears stained with Wright stain. (B, D, F, and G) Neutrophil preparations stained with Kaplow stain. (A, B) Neutrophils from a normal subject. (C, D) Neutrophils from a patient with Chédiak-Higashi syndrome. (E, F) Neutrophils from a patient with specific granule deficiency. (G) Neutrophils from a patient with myeloperoxidase deficiency. The blue positive-staining cell is an eosinophil.

of a certain granule can be associated with deficiency of an entire pool of granules, such as specific granule deficiency. Both enzymatic assays and immunoassays are available to determine the cellular content of many of these granule constituents.

Principles and Interpretation of Laboratory Assessment

The cellular content of neutrophils can be determined by solubilization of a neutrophil pellet with 0.2% Triton X-100, followed by sonication to disrupt the cells and generate a homogeneous lysate. Analysis of the lysate using commercial immunoassays can identify deficiencies of certain granule contents. Diagnosis of myeloperoxidase deficiency can be confirmed by analysis of neutrophil lysates. Similarly, deficiency of both lactoferrin and neutrophil gelatinase (matrix metalloprotein-9 [MMP-9]) is indicative of specific granule deficiency, although interestingly, specific granule deficiency is also associated with a deficiency of the α-defensins (human neutrophil peptides 1–3), stored in the azurophilic granules.[18]

NEUTROPHIL ADHERENCE

Clinical Indications and Implications

Adherence of neutrophils to the endothelium is a prerequisite step to the migration of neutrophils into the tissues. Neutrophils isolated from patients with leukocyte adhesion defect-1 (LAD-1) who lack the common β_2 integrin subunit CD18 exhibit abnormal adherence to the endothelium[19] and therefore are not able to migrate efficiently into the surrounding tissues, often resulting in marked granulocytosis.[20] LAD-2 is a milder form of the disease, in which patients exhibit a defect in fucose metabolism and glycoprotein biosynthesis.[21] Neutrophils from patients with LAD-2 exhibit abnormal expression of the glycoprotein L-selectin and fail to roll along the endothelium. However, they do exhibit normal β_2 integrin–mediated adherence.

Principles and Interpretation of Laboratory Assessment

Adherence of neutrophils can be assessed by measuring binding to plastic using a 96-well plate either uncoated or coated with fetal bovine serum or a specific extracellular matrix (ECM) protein, such as fibrinogen or fibronectin. Endothelial cell monolayers harvested from human umbilical veins may serve as a more physiological substrate for the measurement of cell adhesion. Isolated neutrophils are preloaded with the cell permeant acetoxymethyl ester derivative of the fluorescent dye calcein (calcein-AM). Nonspecific esterases in the cytosol cleave the ester linkage, trapping the fluorescent probe in the cytosol. The labeled neutrophils are added to each well and incubated in the absence or presence of phorbol myristate acetate (PMA) to promote adherence through activation of the integrins. At the end of the incubation, the wells are washed three times to remove nonadherent cells. The fluorescence of each well is determined with a fluorescent microplate reader and compared with the fluorescence of a control well with a fixed number of fluorescent cells. As shown in the left panel of Fig. 94.2, under control conditions, fewer than 10% of the neutrophils adhere to plastic or to plastic coated with fetal bovine serum or fibrinogen. Treatment of normal neutrophils with PMA for 30 minutes results in the adherence of >90% of the neutrophils under all conditions. This adherence assay is valuable in the diagnosis of patients with leukocyte adhesion deficiency. As shown in the right panel of Fig. 94.2, neutrophils isolated from patients with LAD-1 generally exhibit <5% adherence under control conditions and do not increase adherence after treatment with PMA.

NEUTROPHIL CHEMOTAXIS

Clinical Indications and Implications

Neutrophil migration is a prerequisite for neutrophil accumulation at sites of inflammation. Patients with leukocyte chemotactic defects usually show recurrent skin abscesses and occasional life-threatening, invasive infections.

Principles and Interpretation of Laboratory Assessment

Chemotaxis *in vitro* is generally measured using a Boyden chamber. The Boyden chamber includes three components: a lower (chemoattractant) chamber, a nitrocellulose or polycarbonate filter layer, and an upper cell chamber. The lower compartments of the Boyden chamber are filled with a chemoattractant, such as formyl-methionyl-leucyl phenylalanine (fMLF; 10^{-8} M) or interleukin-8 (IL-8; 10^{-8} M). Alternatively, a rapid fluorescence-based measurement of neutrophil chemotaxis that uses a 96-well

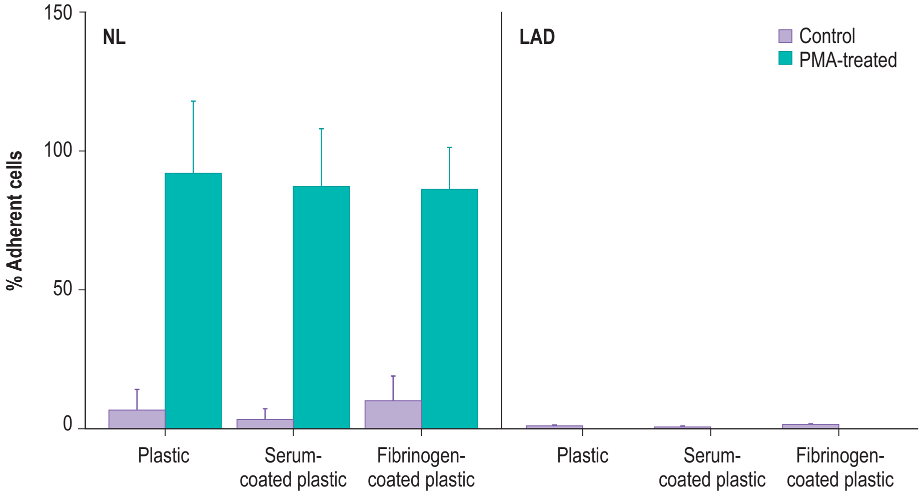


FIG 94.2 Adherence of Neutrophils to Plastic: Normal Versus Leukocyte Adhesion Deficiency. Neutrophils (1 × 10^7 cells/mL Hanks balanced salt solution [HBSS] without divalent cations) were preloaded with acetoxymethyl ester derivative of calcein (calcein-AM: 5 μg/mL) for 15 minutes at 37°C. The cells were washed twice and resuspended in HBSS/4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) with 2% bovine serum albumin (BSA) at a cell concentration of 2 × 10^6/mL. The wells of a 96-well plate were coated for 1 hour at 37°C with 32 μL of either buffer alone, fetal bovine serum, or fibrinogen (2.5 mg/mL). The wells were washed three times, and then cells were added to each well (160 μL/well, 3.2 × 10^5/well). After a 10-minute preincubation at 37°C, PMA (100 ng/mL) was added and the plate was incubated for 30 minutes at 37°C. The wells were then washed three times with HBSS/HEPES to remove nonadherent neutrophils. The percentage of adherent cells was determined by the ratio of the fluorescence of the well compared with the fluorescence of a known standard well. The panel (NL) on the left represents the data (mean ± standard deviation [SD]) from 22 normal subjects and the panel on the right represents the data from three patients with leukocyte adhesion deficiency.

disposable chemotaxis chamber,[22] that can be read in a fluorescence microplate reader, can be used. The lower chamber contains the chemoattractant and is separated from the cellular compartment by a filter. However, instead of a top chamber, the filter has a hydrophobic mask around each filter site that creates surface tension in the cell suspension and aligns the suspension on the hydrophilic filter located directly above the chemoattractant chamber. Calcein-labeled neutrophils are placed on top of the filter. The chemotaxis chamber is incubated for up to 60 minutes at 37°C. Nonmigrating neutrophils atop the filter are rinsed off with buffer, and then the plate is read in a fluorescence microplate reader. The number of migrating neutrophils can be determined by comparison of the fluorescence to a standard well with a known number of fluorescent neutrophils. Less than 5 × 10^6 fluorescent neutrophils are needed to determine neutrophil chemotaxis using several doses of the chemoattractants, fMLF, IL-8, C5a, and leukotriene B_4. The advantages of this assay are high sensitivity, rapid acquisition and analysis of data, and reduced labor in loading the cell suspension. The 96-well format also allows for multiple comparisons to be made under identical conditions.

Instrumentation, such as the EZ-TAXIScan, is now available to monitor chemotaxis temporally. By acquiring digital images over time and analyzing them using imaging software, the coordinates of individual cells can be determined. Changes in the distance (and velocity) in the direction of the chemoattractant (directed migration) and orthogonal to the direction of the chemoattractant (random migration) can be determined. Tracks of multiple cells can be anchored at the origin and displayed graphically (Fig. 94.3, top panels). Adding time as a dimension in the analysis of chemotaxis provides a mechanism to evaluate simultaneously both chemotactic and chemokinetic responses in neutrophils and to detect more subtle defects. Using buffer as a chemoattractant, the average cellular velocity vector in the direction of the chemoattractant and the average cellular velocity vector orthogonal to the direction of the chemoattractant are typically equivalent (Fig. 94.3, bottom panel). When using a chemoattractant (*e.g.,* fMLF, IL-8), typically there is marked increase in the average cellular velocity vector in the direction of the chemoattractant, but little change in the average cellular velocity vector orthogonal to the direction of the chemoattractant (see Fig. 94.3, bottom panel).

EXPRESSION OF SURFACE ANTIGENS

Clinical Indications and Implications

The expression of neutrophil membrane antigens is altered *in vivo* during exudation or after challenge with intravenous endotoxin. Flow cytometric analysis of adhesion molecules on the neutrophil cell surface can indirectly reflect neutrophil adhesion function. Patients with LAD-1 exhibit a deficit in the expression of the common β_2 integrin, CD18, resulting in deficiency in CD11a, CD11b, CD11c, and CD18[23] (Chapter 22).

Principles and Interpretation of Laboratory Assessment

The expression of cell surface antigens is determined on neutrophils stained with specific fluorescent monoclonal antibodies (mAbs) and analyzed by flow cytometric analysis. Neutrophils

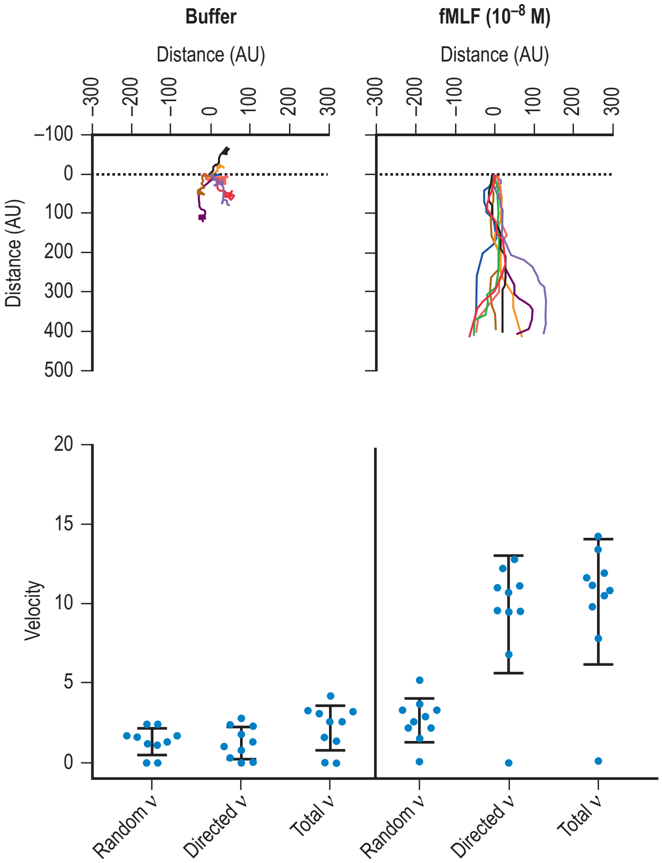


FIG 94.3 Analysis of Chemotaxis. In the top panel, neutrophils (1 µl of 2×10^6 cells/mL in Hanks balanced salt solution [HBSS] with divalent cations) were added to the "Cell" well of EZ-TAXIScan and either buffer (left column) or formyl-methionyl-leucyl-phenylalanine (fMLF; 1×10^{-8} M, right column) was added to the "Chemoattractant" well. The cells were incubated for 60 minutes and images were collected every 2.5 minutes. Using the acquired images, 10 randomly chosen cells were electronically tracked and the paths of the cells plotted with their position at t = 0 anchored at the origin. Presented in the bottom panel are scattergrams of the average velocities of the individual cells that were tracked in the top panel.

stained with nonspecific isotype antibodies are used to determine the nonspecific background staining. To determine the expression of circulating neutrophils and avoid artifacts induced by neutrophil isolation, an aliquot of whole blood can be stained with the appropriate antibody before lysis of the erythrocytes. During flow cytometric analysis, the neutrophils are easily differentiated using their forward light scatter and right angle light scatter to gate on the neutrophil population. Since very little blood is needed (100 µL) for each antigen, neutrophils can be stained with a panel of antibodies to many relevant surface antigens so that a more complete representation of surface antigen expression on neutrophils can be obtained. The panel should include the β_2 integrins (CD11a, CD11b, CD11c, and CD18); selectins (CD62L); Fcγ receptors I, II, and III (CD64, CD32 and CD16); leukosialin (CD43); the common leukocyte antigen (CD45); and distinct surface markers for the granules—carcinoembryonic antigen-related cell adhesion molecule 8 (CEACAM8, or CD66b), a GPI-anchored glycoprotein family member stored in the specific granules, and lysosomal-associated membrane protein 3 (LAMP-3, or CD63), stored in the azurophilic granules. During exudation, the expression of CD11b and CD18 is increased over that observed in peripheral neutrophils, whereas the expression of CD43 (leukosialin) and CD62L is markedly reduced.

The antibody 7D5 recognizes an extracellular epitope of gp91*phox* and can be used to identify surface expression of gp91*phox* as well as mobilization of latent pools of gp91*phox* stored in the specific granules. Flow cytometric analysis of neutrophils stained with 7D5 can often be used to identify patients with X-linked chronic granulomatous disease (CGD) (no 7D5 staining) and X-linked chronic carriers of CGD (mosaic pattern of staining), particularly in patients where the number of cells available for testing is limited.

The expression of surface antigens can also be used to assess the responsiveness of neutrophils to particular ligands, such as fMLF and LPS. As shown in Fig. 94.4, neutrophils isolated from a patient who has a genetic defect in IL-1 receptor–associated kinase-4 (IRAK-4)[24] exhibit abnormal regulation of surface antigen expression to LPS but exhibit normal regulation of surface antigen expression to fMLF. Antigen expression can be upregulated because of translocation of latent antigen to the plasma membrane or downregulated because of either internalization or shedding of the antigen.

NEUTROPHIL DEGRANULATION

Clinical Indications and Implications

The proteases, acid hydrolases, and inflammatory mediators released from storage granules in the neutrophils can mediate bacterial killing, tissue damage, healing, and immune regulation. Lactoferrin that is released from specific granules can chelate iron, resulting in a bactericidal or bacteriostatic effect. Elevation of plasma lactoferrin is an indication of intravascular activation and degranulation of neutrophils.

Principles and Interpretation of Laboratory Assessment

Stimulation of neutrophils with various secretagogues can result in the release of granular enzymes into the extracellular fluid. Treatment of the neutrophils with cytochalasin b (5 µg/mL) disrupts microfilament assembly and facilitates the release of both specific and azurophilic enzymes. Since stimulation of neutrophil degranulation is often accompanied by ROS generation and oxidative inactivation of enzymes, both the cell supernatant fluid and the cell pellet should be analyzed to determine the percentage of enzyme released. To differentiate degranulation from cell lysis, release of the cytosolic enzyme lactate dehydrogenase should be monitored simultaneously.

The release of azurophilic granules can be assessed by determining the levels of myeloperoxidase or elastase. CD63 is also found in the membrane of azurophilic granules and migrates to the neutrophil surface after stimulation with fMLF in the presence of cytochalasin b. The release of specific granules can be assessed by determination of lactoferrin levels by using an enzyme-linked immunoassay. The carcinoembryonic antigen CD66b is found on the neutrophil surface and the specific granules, and its expression on the surface of the neutrophils is increased after stimulation with fMLF or LPS. The secretory granules usually contain proteins that are translocated into the membrane from cytosol during degranulation. Detection of the constituents of secretory granules can be assessed by flow cytometric analysis of the change in expression of surface proteins, such as adhesion molecules, and cytochrome b_{558} of the NADPH oxidase.

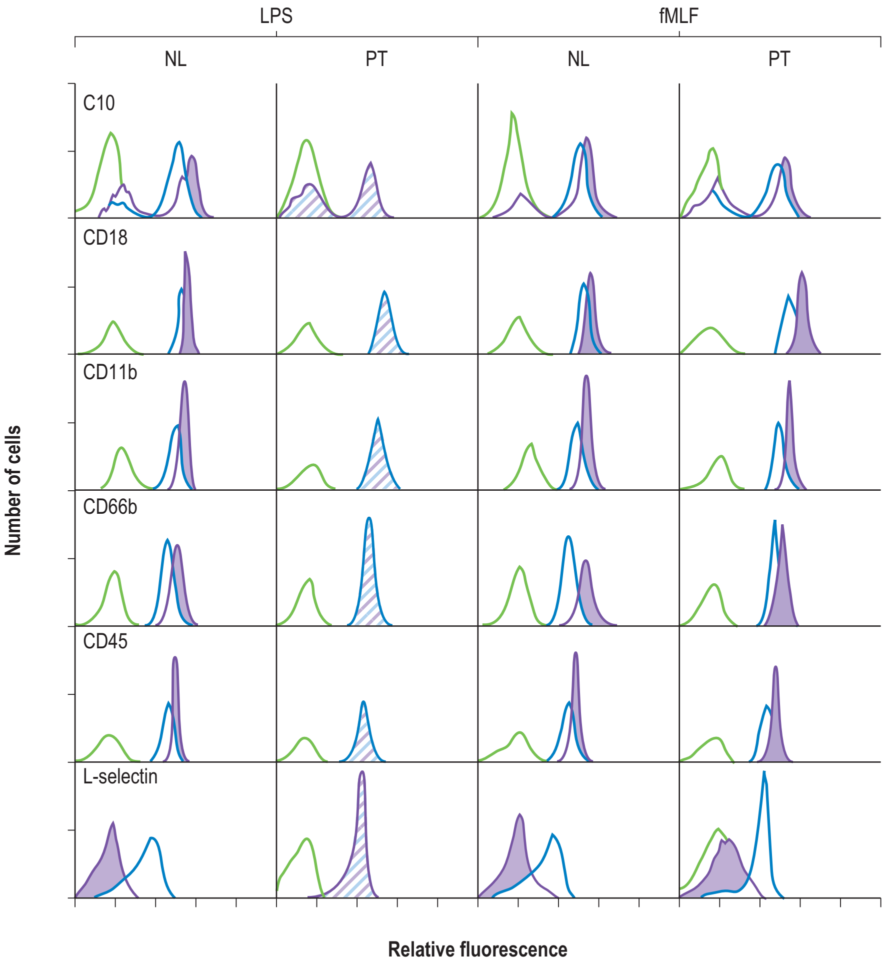


FIG 94.4 **Upregulation of Neutrophil Surface Antigen Expression.** Neutrophils (2.5 × 10^6/mL Hanks balanced salt solution [HBSS] + 10% AB sera) isolated from a normal subject (NL) or from a patient with an interleukin-1 receptor-associated kinase-4 (IRAK-4) mutation (PT) were treated with either lipopolysaccharide (LPS; 100 ng/mL) or formyl-methionyl-leucyl-phenylalanine (fMLF; 0.1 μM) for 30 minutes at 37°C. The cells were washed and stained with C10 (an antibody that demonstrates neutrophil heterogeneity), CD18, CD11b (antibodies to β2 integrins), CD66b (a specific granule marker), CD45 (the common leukocyte antigen), and L-selectin. The green lines represent the isotype control, blue lines represent control neutrophils, and purple lines represent stimulated cells. Differences between control and stimulated cells have been shaded. (From Kuhns DB, Long Priel DA, Gallin JI. Endotoxin and IL-1 hyporesponsiveness in a patient with recurrent bacterial infections. J Immunol 1997;158:3959, with permission of the American Association of Immunologists, Inc.)

GENERATION OF REACTIVE OXYGEN SPECIES

Clinical Indications and Implications

The release of ROS, such as $O_2^{\bullet}$ and H_2O_2, is an important component of the bactericidal machinery of a neutrophil. Neutrophils isolated from patients with CGD have a defect in the NADPH oxidase and are unable to generate ROS, resulting in an O_2-dependent bactericidal defect. The production of ROS has become an important tool to perform risk assessment in patients with CGD. Patients with the lowest ROS generation (<1% of normal generation) have lower survival than patients with higher ROS generation (3–10% of normal). Moreover, survival in CGD is a continuous function of ROS production, suggesting that therapeutic interventions that result in an increase in ROS generation should incur a survival benefit to patients with CGD.[25]

CLINICAL PEARLS

Reactive Oxygen Species (ROS) in Chronic Granulomatous Disease (CGD)

- Neutrophil ROS production, the primary determinate in diagnosis of patients with CGD, ranges from 0.1% to 27% of that observed in normal subjects.
- In addition, survival in CGD is strongly associated with residual ROS production as a continuous variable, and is independent of the specific protein defect.
- ROS production is an important, early indicator of overall risk in CGD.
- In addition, small increases (as little as 3–5% of normal) in residual neutrophil ROS production may confer a survival benefit.
- Careful monitoring with detection of even small increases in ROS may be an important indicator of clinical efficacy during therapeutic intervention.

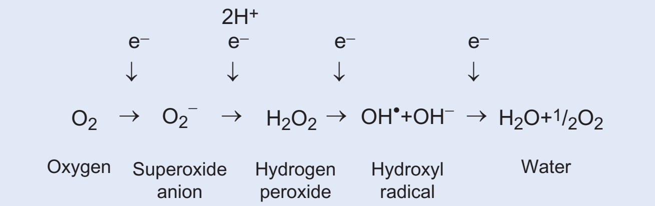
Principles and Interpretation of Laboratory Assessments

The nitroblue tetrazolium (NBT) test is a qualitative assay of ROS production. Either whole blood or isolated neutrophils are mixed with NBT in a chamber slide and stimulated with PMA for 15–30 minutes at 37°C. The neutrophils are allowed to settle on the slide. The slide is air-dried, counterstained with 0.1% safranin, and examined under a microscope. The NBT test yields a visual record of the reduction of the NBT dye to the insoluble, blue-black deposits of formazan. Normal neutrophils, but not neutrophils from patients with CGD, reduce the yellow dye to black-brown-blue aggregates in the cells. Because of the random inactivation of the X chromosome, X-linked carriers of CGD exhibit both NBT^+ and NBT^- neutrophils. The percentage of NBT^+ neutrophils in X-linked carriers of CGD ranges from 5 to 95%. The drawback of the NBT test is the need for manual counting to obtain an accurate reflection of the percentage of positive cells. An alternative to the NBT test is a flow cytometric assay using the dye dihydrorhodamine–123 (DHR-123).[26] Neutrophils are loaded with the nonfluorescent dye and then stimulated with PMA for 15 minutes at 37°C. The H_2O_2 produced oxidizes the dye and results in increased fluorescence, which is detectable with a flow cytometer. The assay is dependent on endogenous myeloperoxidase in the primary granules. Catalase is added to prevent cell-to-cell diffusion of H_2O_2. Since dye is localized to the cytoplasm and catalase is present in the extracellular fluid, the DHR-123 assay detects the intracellular production of ROS. As shown in Fig. 94.5, stimulation of normal neutrophils (Fig. 94.5A) with PMA results in a two-log shift in the fluorescence intensity. Neutrophils from an X-linked carrier of CGD (see Fig. 94.5B) exhibit mosaicism with a negatively stained (abnormal) population and a brightly stained positive population. Neutrophils from a patient with X-linked CGD that lack $gp91^{phox}$ (Fig. 94.5C) express little increase in fluorescence while neutrophils from a patient with a deficiency in $p47^{phox}$ (see Fig. 94.5D) exhibit a slight increase in fluorescence. The major advantages of the DHR-123 assay are the sensitivity, the signal-to-noise ratio, and the ease of counting a large number of cells. Moreover, it has been shown that the DHR-123 assay yields reliable results on ethylenediaminetetraacetic acid (EDTA) or heparin-treated blood samples that have been stored overnight. In general, more than 90% of the neutrophils from the control blood samples will exhibit increased DHR-123 fluorescence. For this same reason, however, overnight samples should not be used to rule out X-linked heterozygosity, since a highly lyonized CGD carrier (>90%) could yield similar results.

The production of $O_2^{\bar{\bullet}}$ can be detected using the reduction of cytochrome c. Because $O_2^{\bar{\bullet}}$ causes a one-to-one stoichiometric reduction of ferricytochrome c to ferrocytochrome c, the resultant increase in the absorption spectrum at 550 nM can be used to quantitate the production of $O_2^{\bar{\bullet}}$. Superoxide dismutase is added to an identical tube to control for the nonspecific reduction of cytochrome c. However, since cytochrome is not permeable to the cells, the detection of $O_2^{\bar{\bullet}}$ is limited to that released into the extracellular milieu. Neutrophils isolated from normal volunteers produce 0.42 ± 0.67 nmol/10^6 neutrophils/10 min under resting conditions; treatment with PMA results in 35.92 ± 11.92 nmol/10^6 neutrophils/10 min (Fig. 94.6). An estimate of normal $O_2^{\bar{\bullet}}$ production over 60 minutes can be obtained by reducing the number of neutrophils in the assay to 2×10^5. Neutrophils isolated from patients with CGD produce little, if any, $O_2^{\bar{\bullet}}$ in response to PMA in 10 minutes (Fig. 94.6). However, some patients with autosomal forms of CGD have low, but detectable $O_2^{\bar{\bullet}}$ production in 60 minutes. Neutrophils isolated from X-linked heterozygous carriers of CGD can yield a full spectrum of $O_2^{\bar{\bullet}}$ production, whereas neutrophils from autosomal recessive carriers of CGD generally yield a normal response (see Fig. 94.6). Although the detection of $O_2^{\bar{\bullet}}$ by reduction of cytochrome c is useful in the diagnosis of patients with CGD, it cannot be used in the diagnosis of carriers because of the wide spectrum of responses that result from the degree of X chromosome lyonization.

KEY CONCEPT

Reactive Oxygen Species



The stepwise reduction of oxygen (O_2) leads to reactive oxygen species and finally to water (H_2O).

Studies have shown that $O_2^{\bar{\bullet}}$ determinations sufficiently reliable to diagnose CGD can be obtained from neutrophils isolated from heparinized whole blood that has been stored overnight. Hence, analyses can be performed on blood samples shipped by overnight express. A normal control blood sample should accompany the sample to ensure adequate shipment handling. By 48 hours of storage, however, there are marked reductions in the PMA response, and the data are no longer valid.

An alternative assay to measure ROS production is luminol-enhanced chemiluminescence. This versatile assay, in addition to its quick and easy setup, offers the capability to test several different subjects and stimuli on the same plate using reduced numbers of cells while providing high sensitivity. The luminescence readings are taken every 1–5 min for up to 2 hours, and data are expressed as relative light units (RLUs). Different stimuli exhibit different kinetics (*e.g.,* fMLF induces a rapid respiratory burst that decays quickly, whereas PMA [20–100 ng/mL] induces a peak of luminescence within 5–15 minutes that slowly decays by 2 hours). Results from normal subjects and patients can be assessed simultaneously and monitored kinetically or using the area under the curve (AUC). Luminol-enhanced chemiluminescence is a measure of both intracellular and extracellular ROS production, although it may not detect them with equivalent efficiency. Addition of superoxide dismutase significantly reduces both peak height and AUC after stimulation with PMA, suggesting that at least a portion of the response can be attributed to $O_2^{\bar{\bullet}}$. Typically, patients with CGD who are deficient in $gp91^{phox}$ have little to no detectable luminescence in this assay; however, as observed with the ferricytochrome c assay, patients with CGD who are deficient in $p47^{phox}$ have detectable luminescence that becomes most evident at later readings (40–80 minutes) (Fig. 94.7).

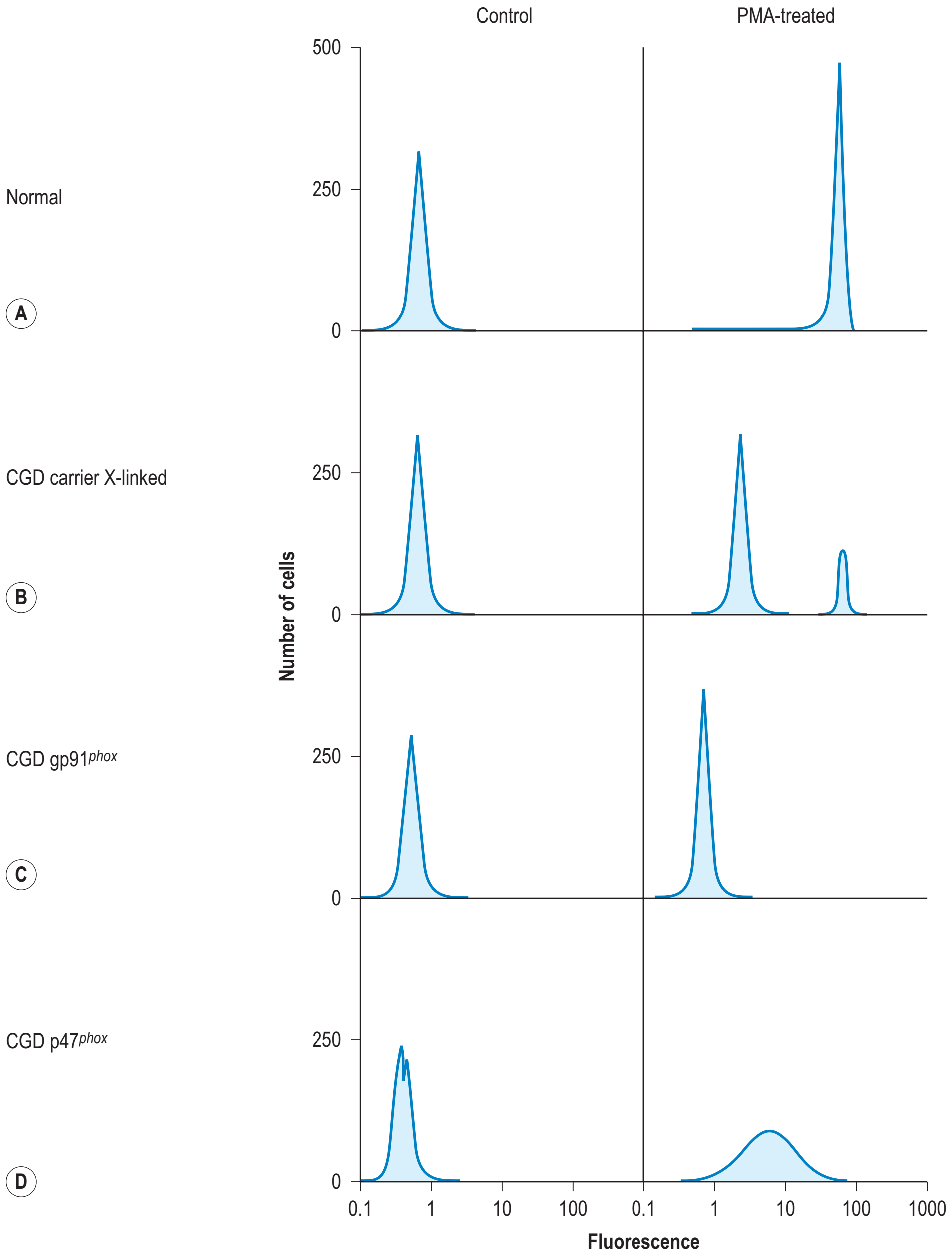


FIG 94.5 Flow Cytometric Analysis of Dihydrorhodamine (DHR) Staining. Whole blood (1.2 mL with ethylenediaminetetraacetic acid [EDTA] as anticoagulant) collected from a normal subject (A), an X-linked chronic granulomatous disease (CGD) carrier (B), a gp91*phox* CGD patient (C), and a p47*phox* CGD patient (D) was lysed by using an ammonium chloride–potassium bicarbonate solution. The remaining leukocytes were resuspended in HBSS and incubated with DHR-123 (100 μM) and catalase (50 μg/ml) for 5 minutes at 37°C. The cells were then incubated an additional 15 minutes at 37°C with either buffer (control) or phorbol myristate acetate (PMA; 400 ng/mL). The cells were immediately analyzed by flow cytometric analysis. Neutrophils were gated using forward light scattering and right-angle light scattering. The analyses presented represent 5000 events within the gated area.

WESTERN BLOT ANALYSIS OF NADPH OXIDASE PROTEIN SUBUNITS

Clinical Indications and Implications

The NADPH oxidase consists of two membrane components (p22*phox* and gp91*phox*), three cytosolic components (p47*phox*, p67*phox*, and p40*phox*), and several guanosine triphosphate (GTP)–binding proteins. CGD is characterized by defects in any one of four of these components—p22*phox* (≈5% of patients with CGD), p47*phox* (≈25% of patients with CGD), p67*phox* (≈5% of patients with CGD), and gp91*phox* (remaining 65% of patients with CGD).

Principles and Interpretation of Laboratory Assessments

The severity of CGD can be related to the specific protein defect. Determination of the specific protein defect in CGD by Western blot analysis also provides direction for determination of the

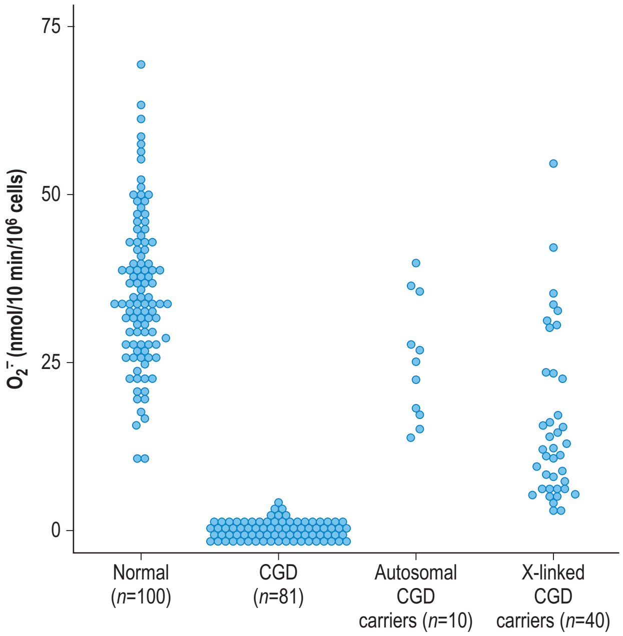


FIG 94.6 $O_2^{\bullet -}$ Generation From Normal Subjects, Patients With Chronic Granulomatous Disease (CGD), and Carriers of CGD. Neutrophils (1 × 10^6/mL Hanks balanced salt solution [HBSS]) were incubated in the presence of 100 µM cytochrome c with phorbol myristate acetate (PMA; 100 ng/mL) for 10 minutes at 37°C. The reaction was terminated by centrifugation at 4°C. Reduction of cytochrome c was monitored at an analytical wavelength of 549.5 nm and a micromolar extinction coefficient of 0.0211. An identical tube containing superoxide dismutase (100 µg/mL) served as a blank.

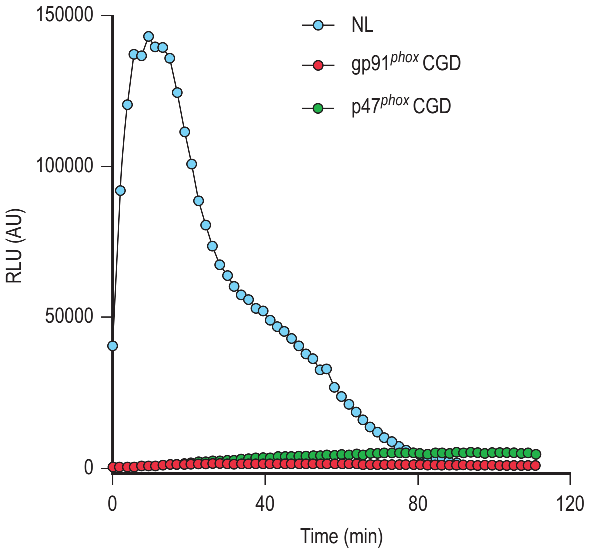


FIG 94.7 Luminol-Enhanced Chemiluminescence. Neutrophils (1 x 10^5/200 µL) were preloaded with luminol (100 µM) for 10 min at 37°C. At t = 0, either buffer or phorbol myristate acetate (PMA; 100 ng/mL) was added and luminescence was monitored for 2 hours, with readings recorded every 2 minutes. Note that at later readings, neutrophils from a patient with p47*phox* deficiency had increased luminescence compared with neutrophils from a patient with chronic granulomatous disease (CGD) gp91*phox* deficiency.

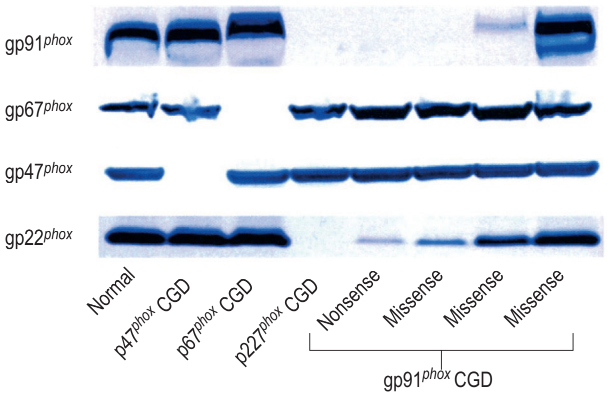


FIG 94.8 Determination of Nicotinamide Adenine Dinucleotide Phosphate (NADPH) Oxidase Protein Defect by Western Blot Analysis. Frozen, diisopropylflurophosphate (DFP)–treated neutrophil pellets (5 × 10^6 cells) were resuspended in polyacrylamide gel electrophoresis (PAGE) sample buffer and sonicated to break up the DNA. One million cell equivalents were loaded into each lane (10% PAGE gels for p47*phox* and p67*phox*, 4–12% gradient PAGE gels for p22*phox* and gp91*phox*). A validated normal control was run on each gel for quality control. The gels were transferred to nitrocellulose, blocked with 5% powdered milk, and incubated overnight with specific antibodies to the phox proteins. The blots were washed, incubated with peroxidase-labeled secondary antibody, and developed with a color reagent. The blots were scanned for permanent storage, and a composite figure was created by using the relevant bands from each blot. The lanes are identified by the specific protein defect and, for gp91*phox*, the type of mutation. CGD, chronic granulomatous disease.

specific genetic defect and enables appropriate genetic counseling for the extended family. A validated normal control is included on each gel for band identification and intensity comparisons. In addition, a control sample from a patient with a known mutation in gp91*phox* CGD is included on each blot to ensure adequate development of p22*phox*. Typical phox protein band patterns are presented in Fig. 94.8. Patients with CGD who have mutations in p47*phox* are Western blot negative. Patients with CGD who have mutations in p67*phox* are generally Western blot negative; however, we have analyzed one patient with a missense mutation in p67*phox* who was Western blot positive. Because p22*phox* and gp91*phox* exist as a membrane complex, patients with a defect in p22*phox* are generally Western blot negative for both p22*phox* and gp91*phox*. In contrast, defects in gp91*phox* yield more variable results. In general, patients with nonsense defects in gp91*phox* exhibit low but detectable levels of p22*phox*. Patients with missense mutations in gp91*phox* that yield detectable gp91*phox* protein exhibit proportionately higher levels of p22*phox*. Neutrophils isolated from overnight samples can be used to diagnose p47*phox* deficiency because of the stability of the protein. However, detection of other phox protein defects in overnight samples can be more problematic because of proteolysis of p67*phox* and the gp91*phox*–p22*phox* complex.

CONCLUSIONS

Our understanding of neutrophil biology is undergoing significant changes. Long-standing axioms are being challenged. Chakravarti et al. suggested that treatment of neutrophils with a cocktail of cytokines can reprogram a phenotypically distinct subset of neutrophils with increased survival and altered function.[27] Rodriguez et al. have shown that patients with renal cancer have a subset of lighter density neutrophils that colocalize with mononuclear cells atop the density cushion and appear to be phenotypically "activated";[28] moreover, these lighter-density neutrophils are capable of suppressing T-cell proliferation through a mechanism dependent on arginase-1 stored in the gelatinase granules, a novel function of neutrophils.

ON THE HORIZON

There is growing evidence that neutrophils may be derived from organs other than bone marrow, such as the lung, liver, and, particularly, spleen. Recently, it has been suggested that the spleen may serve as a reserve site for "emergency myelopoiesis" during infection. Other studies suggest that c-kit$^+$ myeloid precursors from blood migrate to a site of cutaneous infection, where they proliferate and form mature polymorphonuclear neutrophils (PMNs). These findings are based on murine studies and raise the following questions:

- Are these extramedullary neutrophils found in humans?
- Are there specific biomarkers that distinguish them from their bone marrow–derived circulating counterparts?
- Do they have functionality that is different from circulating neutrophils?

Much of our current understanding of neutrophil biology has developed from the study of genetic immunodeficiencies. Whether there are immunodeficiencies that result from defects in these recently discovered pathways has yet to be determined.

ACKNOWLEDGMENT

This project has been funded in whole or in part with federal funds from the National Cancer Institute, National Institutes of Health, under Contract HHSN261200800001E. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the US Government.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Jhunjhunwala S, Alvarez D, Aresta-DaSilva S, et al. Splenic progenitors aid in maintaining high neutrophil numbers at sites of sterile chronic inflammation. J Leukocyte Biol 2016;100:253–60.
2. Bainton DF. The cells of inflammation: a general view. In: Weissman G, editor. The cell biology of inflammation. 2nd ed. New York: Elsevier/North-Holland; 1980.
3. Pillay JP, den Braber I, Vrisekoop N, et al. *In vivo* labeling with ^2H_2O reveals a human neutrophil lifespan of 5.4 days. Blood 2010;116(4):625–7.
4. Borregaard N, Cowland JB. Granules of the human neutrophilic polymorphonuclear leukocyte. Blood 1997;89:3503–21.
5. Volpp BD, Nauseef WM, Clark RA. Two cytosolic neutrophil oxidase components absent in autosomal chronic granulomatous disease. Science 1988;242:1295–7.
6. Matute JD, Arias AA, Wright NAM, et al. A new genetic subgroup of chronic granulomatous disease with autosomal recessive mutations in p40^{phox} and selective defects in neutrophil NADPH oxidase activity. Blood 2009;114:3309–15.
7. Parkos CA, Dinauer MC, Walker LE, et al. Primary structure and unique expression of the 22-kilodalton light chain of human neutrophil cytochrome *b*. Proc Natl Acad Sci USA 1988;85:3319–23.
8. Segal AW, Cross AR, Garcia RC, et al. Absence of cytochrome b$_{-245}$ in chronic granulomatous disease—a multicenter European evaluation of its incidence and relevance. N Engl J Med 1983;308:245–51.
9. Brinkmann V, Reichard U, Goosmann C, et al. Neutrophil extracellular traps kill bacteria. Science 2004;303:1532–5.
10. Lood C, Blanco LP, Purmalek MM, et al. Neutrophil extracellular traps enriched in oxidized mitochondrial DNA are interferogenic and contribute to lupus-like disease. Nat Med 2016;27:146–53.
11. Kuijpers TW, Tool ATJ, van der Schoot CE, et al. Membrane surface antigen expression on neutrophils: a reappraisal of the use of surface markers for neutrophil activation. Blood 1991;78:1105–11.
12. Böyum A. Isolation of mononuclear cells and granulocytes from human blood. Isolation of monuclear cells by one centrifugation, and of granulocytes by combining centrifugation and sedimentation at 1 g. Scand J Clin Invest Suppl 1968;97:77–89.
13. Haslett C, Guthrie LA, Kopaniak MM, et al. Modulation of multiple neutrophil functions by preparative methods or trace concentrations of bacterial lipopolysaccharide. Am J Pathol 1985;119:101–10.
14. Kaplow LS. Simplified myeloperoxidase stain using benzidine dihydrochloride. Blood 1965;26:215–19.
15. Introne W, Boissy REB, Gahl WA. Clinical, molecular, and cell biological aspects of Chediak–Higashi syndrome. Mol Genet Metab 1999;68:283–303.
16. Strauss RG, Bove KE, Jones JF, et al. An anomaly of neutrophil morphology with impaired function. NEJM 1974;290:478–84.
17. Nauseef WM. Diagnostic assays for myeloperoxidase deficiency. In: Quinn MT, DeLeo FR, Bokoch GM, editors. Neutrophil methods and protocols: methods in molecular biology. Totowa, NJ: Humana Press; 2007. p. 525–30.
18. Parmley RT, Gilbert CS, Boxer LA. Abnormal peroxidase-positive granules in "specific granule" deficiency. Blood 1989;73:838–44.
19. Buescher ES, Gaither T, Nath J, et al. Abnormal adherence-related functions of neutrophils, monocytes, and Epstein-Barr virus-transformed B cells in a patient with C3bi receptor deficiency. Blood 1985;65:1382–90.
20. Anderson DC, Schmalsteig FC, Finegold MJ, et al. The severe and moderate phenotypes of heritable Mac-1, LFA-1 deficiency: their quantitative definition and relation to leukocyte dysfunction and clinical features. J Infect Dis 1985;152:668–89.
21. Etzioni A, Frydman M, Pollack S, et al. Recurrent severe infections caused by a novel leukocyte adhesion deficiency. N Engl J Med 1992;327:1789–92.
22. Frevert CW, Wong VA, Goodman RB, et al. Rapid fluorescence-based measurement of neutrophil migration *in vitro*. J Immunol Methods 1998;213:41–52.
23. Anderson DC, Springer TA. Leukocyte adhesion deficiency: an inherited defect in the Mac-1, LFA-1, and p150,95 glycoproteins. Annu Rev Med 1987;38:175–94.
24. Kuhns DB, Long Priel DA, Gallin JI. Endotoxin and IL-1 hyporesponsiveness in a patient with recurrent bacterial infections. J Immunol 1997;158:3959–64.
25. Kuhns DB, Alvord WG, Heller T, et al. Residual NADPH oxidase and survival in chronic granulomatous disease. N Engl J Med 2010;363:2600–10.
26. Emmendörffer A, Hecht M, Lohmann-Matthes M-L, et al. A fast and easy method to determine the production of reactive oxygen intermediates by human and murine phagocytes using dihydrorhodamine 123. J Immunol Methods 1990;131:269–75.
27. Chakravarti A, Rusu D, Flamand N, et al. Reprogramming of a subpopulation of human blood neutrophils by prolonged exposure to cytokines. Lab Investig 2009;89:1084–99.
28. Rodriguez PC, Ernstoff MS, Hernandez C, et al. Arginase I-producing myeloid-derived suppressor cells in renal cell carcinoma are a subpopulation of activated granulocytes. Cancer Res 2009;15:1553–60.

MULTIPLE-CHOICE QUESTIONS

1. Results from laboratory testing show that neutrophils from a patient exhibit a negative dihydrorhodamine (DHR) analysis but exhibit normal/supranormal production of $O_2^{\bar{\bullet}}$ by cytochrome c reduction. What is the MOST likely reason(s) for these results?
   A. The patient probably has myeloperoxidase (MPO) deficiency; the discrepancy in these test results should be resolved by Kaplow peroxidase staining of the polymorphonuclear neutrophil (PMN) preparation and/or MPO enzyme-linked immunosorbent assay (ELISA).
   B. The patient has no immune defect since the $O_2^{\bar{\bullet}}$-dependent ferricytochrome c assay falls with the normal range.
   C. The patient has chronic granulomatous disease (CGD).
   D. None of the above.

2. Why are neutrophils treated with the fluorescent dye, calcein-AM, in the adherence assay?
   A. Calcein is needed to activate the cells.
   B. Calcein is needed to activate the integrins.
   C. The calcein fluorescence of the adherent cells can be quantified with the use of a fluorescent plate reader.
   D. None of the above.

3. The DHR analysis of neutrophils stimulated with phorbol myristate acetate (PMA) from a patient yielded a broad peak of fluorescence with a moderate shift to the right that did not fall within the normal range. What is the MOST likely interpretation of this finding?
   A. The subject most likely has CGD with a defect in p47*phox*.
   B. The subject is most likely a carrier of gp91*phox* deficient CGD.
   C. The subject most likely has CGD with a defect in gp91*phox*.
   D. The subject is most likely normal.

4. When performing Western blot analyses for the detection of components of nicotinamide adenine dinucleotide phosphate (NADPH) oxidase, what is the MOST important reason to run a control from a patient with a known mutation in gp91*phox*?
   A. It is performed to serve as validation of the assay.
   B. It is performed to ensure adequate development of p22*phox* for the correct identification of the impaired protein of the membrane component.
   C. It is only important to run a control from a normal subject for band identification and intensity comparisons
   D. Only molecular weight markers are needed to run to serve as band identification.

95

Assessment of Human Allergic Diseases

Robert G. Hamilton

Human allergic disease comprises a spectrum of immunoglobulin E (IgE)–mediated immediate-type hypersensitivity reactions that manifest as reactions in skin (urticaria, dermatitis), the respiratory tract (asthma, rhinitis, or sinusitis), eyes (conjunctivitis), the gastrointestinal (GI) tract (abdominal pain, bloating, vomiting, diarrhea), and, in their most extreme condition, systemic anaphylaxis. These reactions are precipitated by exposure of a genetically predisposed and sensitized (IgE antibody-positive) individual to a variety of environmental substances that are ubiquitous and usually well tolerated by most healthy individuals. This chapter reviews the principles and performance characteristics of analytical methods used in the diagnosis and management of individuals with allergic disease. It examines *in vivo* and *in vitro* methods for the quantification of total and allergen-specific IgE.

BIOLOGICAL PROPERTIES OF IGE

In 1921, Prausnitz and Küstner (PK)[1] reported that an intradermal (ID) injection of serum from an allergic individual into the skin of an unsensitized (nonallergic) individual, followed 24 hours later by injection of specific antigen into the same skin site, induced local itching and swelling surrounded by a zone of erythema. This passively transferred allergic reaction, or *PK reaction,* reached a maximum within 10 minutes, persisted for about 20 minutes, and gradually disappeared. The antibody mediating this reaction was shown to be thermolabile, losing its sensitizing activity after heating serum at 56°C for several hours. In 1967, this antibody was identified as a fifth human immunoglobulin isotype and designated as IgE.[2-4]

Total serum IgE concentrations are the lowest of the five human immunoglobulin isotypes (0–0.0001 g/L; 0.004% of the total adult serum immunoglobulin).[5] Approximately 50% of IgE is localized in the extravascular space. Its short biological half-life of 1–5 days in peripheral blood is primarily the result of a relatively high fractional catabolic rate (71% of the intravascular pool catabolized/day). IgE does not pass the placenta or activate the classical complement pathway. Its reaginic (mast-cell sensitizing) activity is dependent on its ability to bind to the α chain of the high-affinity IgE Fc-ε receptor (α-FcεRI) that resides on the membrane surface of basophils and mast cells. The interaction between IgE Fc and the α-FcεRI is blocked by the therapeutic subcutaneous administration of anti-IgE therapeutics, such as omalizumab (a humanized IgG1-κ anti-IgE Fc), as discussed below.

CLINICAL IMPORTANCE OF TOTAL SERUM IGE

The concentration of IgE in the serum is highly age dependent.[5] Cord serum IgE concentrations are low, usually <2 kU/L (<4.88 µg/L; 1 kU = 2.44 µg and 1 IU = 2.44 ng). Serum IgE levels progressively increase in healthy children up to the age of 10–15 years and gradually decline in an age-dependent manner from the second to the eighth decades of life. Infants with atopia have an earlier and steeper rise in serum IgE levels during their early years compared with age-matched controls without atopia.[6]

Clinically, a patient's total serum IgE level should be evaluated against reference intervals established with sera from an age-stratified, healthy skin test-negative (nonatopic) population.[6] Many clinical laboratories define a total serum IgE >100 kU/L (240 ng/mL) as a general demarcation into the atopic region.[7] After age 14 years, serum IgE levels >333 kU/l (>800 µg/L) are abnormally elevated and strongly associated with atopic disorders, such as allergic rhinitis, extrinsic or allergic asthma, and atopic dermatitis. Extreme elevations in serum IgE are common in parasitic infections and are necessary for the diagnosis of the hyper-IgE (Job's) syndrome. Low total IgE levels <25 kU/L in individuals with asthma suggest that IgE-mediated mechanisms play an insignificant role in the pathogenesis of their condition. Low total serum IgE levels thus support the diagnosis of nonallergic (intrinsic) asthma and help exclude allergic bronchopulmonary aspergillosis. There is extensive overlap between IgE levels in atopic and nonatopic populations,[6-9] which means that an elevated serum IgE can be useful in confirming the clinical diagnosis of allergic respiratory or skin diseases, but a low or normal value does not eliminate the possibility of an IgE-mediated mechanism. For instance, a group of adults with allergic asthma had a mean serum IgE level of 1589 ng/mL (range 55–12 750 ng/mL). In contrast, approximately 90% of patients with atopic dermatitis had elevated total serum IgE levels (mean 978 kU/L; range 1.3–65,208 kU/L). Parasitic infections, selected immunodeficiency states, cancer (Hodgkin disease, bronchial carcinoma), rheumatoid arthritis, liver disease, and atopic dermatitis (eczema) are other disease states that have been associated with a dysregulation of total serum IgE levels. The total serum IgE must therefore be interpreted within the relevant clinical context for each patient.

Because of the overlap between individuals with atopia and those without, total serum IgE measurements have been largely replaced in the routine diagnosis of allergic disease by the quantification of allergen-specific IgE antibody. However, quantification of total serum IgE has remained important for computing the therapeutic dose of anti-IgE. Omalizumab is a recombinant, humanized IgG1-κ monoclonal antihuman IgE Fc drug that specifically binds to the region on the ε heavy chain that interacts with α-FcεR1.[10,11] It is used to treat moderate to severe persistent allergic asthma and chronic idiopathic urticaria by blocking IgE binding to the α-FcεR1. The binding of omalizumab to IgE *in vivo* reduces both the number of free IgE molecules able to interact with the α-FcεR1 and the number of α-FcεR1 receptors on the surface of effector cells. The consequence is a reduction in mediator release and allergy symptoms following allergen exposure.[10,11]

CLINICAL IMPORTANCE OF ALLERGEN-SPECIFIC IGE

In contrast to total serum IgE, the presence of allergen-specific IgE antibody on the surface of circulating basophils or skin mast cells or in the serum is highly predictive of an individual's propensity to exhibit an allergic response following reexposure to that allergen. Before its identification as a novel immunoglobulin, IgE was only detectable with *in vivo* bioassays (skin test, bronchial or nasal provocation tests). Purification of IgE myeloma protein and the subsequent production of antisera specific for IgE led to the development of the first *in vitro* assay (radioallergosorbent test [RAST]) for the detection of allergen-specific IgE antibody in serum.[4,12,13] Since then, nonisotopic autoanalyzer variants based on the original noncompetitive cellulose paper disc solid-phase RAST design have been widely used in clinical immunology laboratories throughout the world.[13,14] These assays, discussed below, involve an immobilized allergen containing reagent that binds allergen-specific antibody from human serum and an enzyme conjugated anti-IgE the detects bound IgE antibody.

Historical studies have compared the diagnostic performance (sensitivity and specificity) of *in vivo* and the *in vitro* assays in the diagnosis of human allergic disease. These intermethod comparisons have shown that the presence of IgE antibody as measured by serological immunoassay methods usually agrees well with the presence of IgE detected in leukocyte and mast-cell histamine release assays, and provocation tests, such as the skin test, food challenge, and inhalation provocation test.[15,16] However, these early studies emphasize that the presence of IgE antibody as detected either *in vivo* or *in vitro* is at best a confirmatory measurement for sensitization.[17] IgE antibody is necessary but not sufficient for induction of an allergic symptom. It is considered an important risk factor in the diagnosis of allergic disease that supports a patient's medical, family, and environmental histories of a temporal association between allergic symptoms and allergen exposure. The clinical importance of differences in diagnostic sensitivity between skin test and serological detection of IgE antibody may be less important for patients with allergies to inhaled (pollen, dust mite, and epidermal) allergens than in those facing life-threatening anaphylactic reactions caused by *Hymenoptera* stings and certain drugs. In these latter cases, skin tests are preferable to *in vitro* immunoassay analyses for the detection of allergen-specific IgE antibodies.[16,18] Immunoassays of IgE antibody in serum can, however, be helpful in cases where the patient has taken antihistamines, β-receptor stimulants, or high-dose steroids, which can reduce the *in vivo* provocation test's measured response in children, pregnant women, and older patients, in whom skin testing may not be well tolerated and when dealing with allergens (*e.g.*, foods, molds) where commercial extracts can be highly variable or labile.[13,16]

CLINICAL HISTORY

The diagnosis of human allergic disease is driven principally by the patient's clinical history in which objective evidence is collected that an allergic reaction has occurred following exposure to a known or suspected allergen source. During the history, a number of factors need to be considered.[16] These include the patient's symptom characteristics (location, reproducibility, severity, duration, delay time from appearance following allergen exposure), atopic factors (personal and family histories, age of onset, infantile atopic dermatitis), opportunity for sensitization (geography, seasons, duration of exposure, prior exposures, employment, and hobbies), specificity of allergen triggers, and comorbidities (*e.g.*, nasal polyps, recurrent sinusitis, chronic obstructive pulmonary disease [COPD]). From the history, an *a priori* or pretest probability or likelihood of allergic disease is derived, and this determines whether or not confirmatory IgE antibody testing for sensitization is warranted.

DIAGNOSTIC METHODS

A combination of *in vivo* provocation and *in vitro* laboratory tests may be used to confirm sensitization and provide support for the clinical diagnosis of allergic disease. The actual tests selected depend on the nature of the disease process (*e.g.*, allergic asthma, urticaria/angioedema, rhinitis/sinusitis, or anaphylaxis) and the suspected allergen triggers (*e.g.*, aeroallergens, venoms, drugs, foods) under investigation. In making a decision about which diagnostic test to perform, clinicians start by searching for information on the diagnostic sensitivity and specificity of each confirmatory test. The true *diagnostic sensitivity* or ability of each test to detect IgE antibody in the presence of allergic disease and *diagnostic specificity* or ability of the test to not detect IgE antibody in health (absence of allergic disease) are difficult to determine definitively. Not only are there a myriad of assay methods, diverse techniques (*e.g.*, skin testing), reagents (extracts and molecules) and grading, interpolation, and interpretation methods, but most importantly, there is a general absence of gold standard methods for defining the presence of allergic disease. For this reason, results of the confirmatory tests need to be viewed as additional risk factors and tests for sensitization, rather than definitive indicators of disease.[17] In the end, the choice of the confirmatory test is a matter of clinical judgment.

INITIAL CLINICAL LABORATORY TESTS

Following the collection of a medical history and performance of a physical examination, the patient who is suspected of having allergic disease may undergo several preliminary blood tests. A complete blood count (CBC), and/or a total blood eosinophil count, if performed, should be obtained before any systemic corticosteroids or epinephrine is administered. A normal whole blood eosinophil level ranges from 0 to 500 cells/mm^3. Children generally have higher normal levels (mean 240 cells/mm^3 95% confidence interval [CI] = 0–740 cells/mm^3), with peak levels occurring at 4–8 years of age. Most clinical laboratories consider a differential white blood cell (WBC) count with an eosinophil proportion >5–10% of the total WBC count to be abnormal. Blood, sputum, and nasal secretion eosinophilia is characteristic of asthma, whether or not IgE-mediated allergic processes are present. In a bronchitic sputum specimen, neutrophils predominate. A neutrophilic nasal discharge is characteristic of sinusitis. Other laboratory tests that may be ordered, as indicated, include pulmonary function tests and a chest X-ray or sinus computer tomography (CT) scan.

IN VIVO PROVOCATION TESTING

Both the skin test and nasal/bronchial/GI provocation tests are useful *in vivo* diagnostic tools for the confirmation of immediate-type hypersensitivity reactions associated with allergic disease. They can also allow the identification of offending allergens in an allergy patient's workup for avoidance, or management with pharmacotherapy, immunotherapy or anti-IgE therapy.

Skin Tests

Historically, Guerin and Watson[19] described a three-phase skin response during an immediate-type skin test reaction following the administration of a stimulus (allergen or histamine-positive control). First, a bluish-white area appears that involves the constriction of capillaries and that typically disappears within minutes. Second, an erythematous peripheral halo or flare appears as a result of arteriole dilatation. Finally, a circular urticarial papule or wheal is observed, as a result of extravasation of plasma into the skin. The response is generally maximal by 15–20 minutes. The immediate "wheal and flare" reaction can be followed by a late-phase reaction 5–6 hours later that appears as a poorly defined edema-like reaction and that usually disappears by 24 hours. An allergen extract can be administered either by a prick or puncture or by ID injection.[20]

Puncture skin testing was first described by Lewis and Grant in 1924[21] and modified by Pepys.[22] It involves placing a drop of each test allergen extract or control solutions (histamine and saline) on the skin of the forearm or back and the introduction of allergen into the epidermis by a needle puncture. Importantly, drops are spaced at least 2 cm apart to prevent cross-over bleeding that can produce false-positive reactions or difficulty reading each discrete test site because of overlapping erythema. A variety of single-point (23–26 gauge), multipoint, and bifurcated needles have been used.[20,23] After the prick or puncture, the excess allergen is removed by blotting with tissue paper or gauze approximately 1 minute later. With single devices, separate needles are used to insure no cross-contamination. An immediate reaction (wheal and erythema) is read at 15–20 minutes as it reaches its maximum size. Because of the direct skin irritation with some crude allergen extracts, bleeding can produce false-positive results.

The ID skin test, first reported by Mantoux[24] in 1908, is 1000–30,000 times more analytically sensitive by concentration than the puncture skin test. A 0.02- to 0.05-mL volume of diluted allergen extract or controls (histamine or saline) in a 0.5- to 1.0-mL tuberculin syringe is injected intracutaneously through a 26- to 27-gauge needle. Importantly, the bevel of the needle needs to face up and injection should be no deeper than the superficial layers of the skin. A 0.02-mL injection will initially produce a superficial 2- to 4-mm-diameter bleb. Like the puncture test, the ID skin test is read at 15–20 minutes, when the reaction is maximal. Dilutions of extract >1 : 1000 weight/volume (w/v) are commonly used to minimize false-positive reactions because of irritation and the potential for systemic reactions, which range from 0.02% of 1.4% of patients tested.[20] Subcutaneous administration of the allergen may lead to a false-negative result. The volume of allergen extract that is injected only slightly influences the size of the wheal-and-flare reaction, whereas concentration is the most important determinant of the final ID skin test result. ID testing allows an investigator to perform a skin test titration to quantitatively determine the patient's skin sensitivity. For serial titration, the same volume (*e.g.,* 0.02 mL) of 3–10-fold serial dilutions of allergen extract are injected into different skin sites and the concentration of allergen required to produce a wheal or erythema of a defined mean diameter (*e.g.,* 8-mm wheal) is interpolated.[25] The higher the concentration of allergen required to induce the defined size of wheal or erythema, the less sensitive is the patient to that allergen preparation and/or the lower is the allergenic potency of the extract.

Variables That Influence Skin Test Responses

The quality (composition, potency, heterogeneity) of the allergen extract is the single most important variable that affects skin test performance. Most skin test extracts are nonstandardized, and their potency is reported in biological or weight per volume units.[26] Allergen extracts used in puncture skin testing contain 50% glycerin, which is present to enhance stability. However, the glycerin causes skin irritation and false-positive skin test results if used intradermally without dilution. Other factors that influence the skin test response include the area of the body that is tested (back vs forearm), age of patient (skin wheals increase in size from infancy to adulthood), race (clarity of reading on dark vs light skin pigmentation) and preadministered drugs (*e.g.,* antihistamines, tranquillizers, corticosteroids).

In terms of controls, the saline (negative control) can identify dermatographic patients and trauma-induced reactivity produced by the puncturing device. The histamine positive control (for prick or puncture at 5.43 mmol/L or 1 mg/mL of histamine base) is useful in detecting medication- or disease-induced suppression of the skin test response. It is also used as a quality control reagent to document the reproducibility of technician performance.[27]

Relationship Between Puncture and Intradermal Skin Test Responses

Fig. 95.1 shows the relationship between the ng/mL level of *Dermatophagoides pteronyssinus* (*Dpt,* dust mite) specific IgE antibody in sera from 30 subjects with dust-mite allergy, as measured in serum by an *in vitro* assay, as well as the ID skin test midpoint *Dpt* allergen extract titer required to produce an 8 mm wheal in the same individual. Using the same *Dpt* extract in both tests, a higher degree of skin sensitivity (*i.e.,* lower titer of antigen required to induce an 8-mm wheal) was strongly correlated ($r^2 = 0.77$; $p < 0.001$) with higher serum IgE antibody levels in those with the higher levels of skin sensitivity (<10 ng/mL midpoint). Fig. 95.2 shows the strong correlation between the wheal size that is observed in the same patients with *Dpt* allergy receiving the same dust-mite extract by a single puncture skin test and a midpoint ID skin test titration. Both the puncture and the ID skin test procedures produce a maximal wheal and flare size by 15–20 minutes, which is measured with a millimeter ruler or caliper. The maximal diameter and the midpoint perpendicular diameter are averaged to generate an index. A permanent record of the skin reaction can be made by applying adhesive cellulose tape over the wheal-and-flare skin area, which has previously been outlined with a felt-tip or ballpoint pen. Using a single concentration of allergen, the ID skin test can be graded according to one of several reported systems (Table 95.1).[20,23] Alternatively,

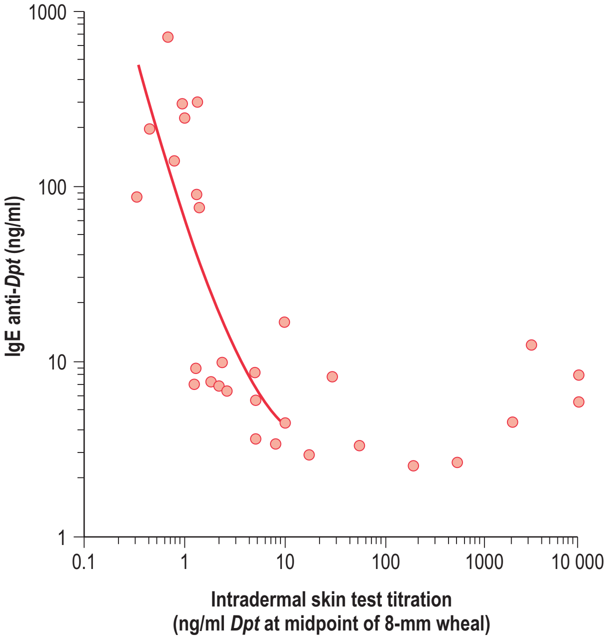


FIG 95.1 Relationship between immunoglobulin E (IgE) anti-*Dermatophagoides pteronyssinus* (*Dpt*) measurements in skin (*x*-axis, intradermal skin test titration; ng/mL of *Dpt* required to produce an 8-mm wheal) and in the serum (*y*-axis, ng/mL of IgE anti-*Dpt* as measured by IgE antibody serology; sensitivity = 2 ng/mL). These results were obtained by testing the skin and serum of 30 individuals with dust-mite allergy and varying degrees of clinical sensitivity by using the same *Dpt* extract in both IgE antibody serology assay and the intradermal (ID) skin test titration study. A lower "titer" of antigen required to induce an 8-mm wheal (*e.g.,* higher degree of skin sensitivity) was strongly correlated ($r = 0.77$; $p < 0.001$) with a higher serum IgE antibody level in individuals with the higher level of skin sensitivity (<10 ng/mL midpoint). Less sensitive patients (titers >10 ng/mL *Dpt*) had lower levels of serum antibody (2–15 ng/mL) that did not relate well with skin sensitivity.

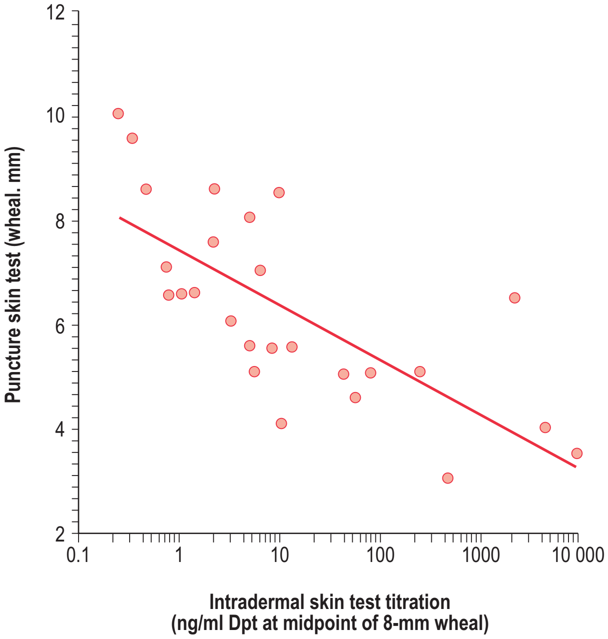


FIG 95.2 The wheal size in millimeters at a single dose of *Dermatophagoides pteronyssinus* (*Dpt*) allergen administered in a puncture skin test compared with the titer or nanogram per milliliter (ng/mL) of the same *Dpt* allergen obtained in an intradermal (ID) skin test titration on the same 26 dust mite-allergic patients to produce an 8-mm wheal. These data indicate that the wheal size obtained with a single dose of allergen by the less labor-intensive puncture skin test is as predictive ($r^2 = 0.72$; $p < 0.001$) of relative patient sensitivity as the more technically complex intradermal skin test titration study, which involves the administration of seven increasing concentrations of the same allergen into different skin sites.

a midpoint titer can be interpolated from a skin test titration, including 3–10-fold serial dilutions of the allergen extract. Some investigators prefer to use the erythema (flare) size rather than the wheal size obtained during titration studies because the slope of the flare's regression line with dose is steeper.[25] The strong relationship between the size of the intradermal erythema and wheal observed with the mean of 304 duplicate skin tests is shown in Fig. 95.3. This relationship is useful to know because the erythema is difficult to assess in many dark-skinned subjects.

Conjunctival, bronchial and nasal provocation tests

Conjunctival, bronchial, and nasal provocation challenges are performed primarily as research procedures to identify a relationship between allergen exposure and a change in the patient's ocular, bronchial, or nasal physiology. Along with puncture skin tests, they are more specific but less sensitive than ID skin tests.[28] Bronchoprovocation studies with methacholine or histamine are the most commonly performed, particularly in the diagnosis of difficult cases of asthma. The bronchoprovocation procedure involves the administration of either methacholine or histamine via a calibrated nebulizer, starting at doses of 0.03–0.1 mg/mL and doubling the concentration up to 10–25 mg/mL. Alternatively, the administration of increasing doses of allergen extracts is performed principally as a research tool to investigate humoral and cellular mechanisms involved in the pathogenesis of allergen-induced airway responses. Pulmonary function is monitored after each dose. A positive response is typically defined as the concentration of agonist that results in a drop in the forced expiratory volume per second (FEV_1) of 20% or more from the baseline. A value >70% of predicted needs to occur to obtain a valid interpretation. More extensive details regarding the methods and interpretation of bronchial challenges are presented elsewhere.[29,30]

Nasal provocation in its simplest form involves the controlled administration of buffer (human serum albumin–saline) or increasing concentrations of allergen into one or both sites of a washed nasal passage in individuals suspected or known to be allergic. The symptoms (*e.g.,* itching, number of sneezes, rhinorrhea, and nasal conjugation) induced and/or degree of mast cell and basophil degranulation markers (*e.g.,* histamine, tryptase, prostaglandin D_2, cysteinyl leukotrienes) and cytokines (interleukin-4 [IL-4], IL-5, IL-10, IL-13) released into nasal lavage fluids after each concentration of allergen indicates the relative level of sensitivity to that allergen. Nasal airway resistance is a less satisfactory endpoint because of high intrinsic variations. Details of the procedures and applications can be found elsewhere.[31]

TABLE 95.1 Grading System for Puncture and Intradermal Skin Testing Using Histamine as a Reference[a]

Grade or Class	Wheel Size (mm)	Erythema Size (mm)
Skin Testing Grading System		
0	No discernible wheal	
1+	≤3 histamine wheal	
2+	>3 histamine and <13 histamine wheal	
3+	= size of histamine wheal ± 1 mm	
4+	>13 histamine wheal and <23 histamine wheal	
5+	>23 histamine wheal	
Alternative Skin Test Grading System for Intradermal Skin Testing Only Involving Interpretation of Wheal and Erythema Responses[†]		
0	<5	<5
+/–	5–10	5–10
1+	5–10	11–20
2+	5–10	21–30
3+	10–15	31–40
4+	>15 with pseudopods	41–50

[a] Prick/puncture histamine (3–10 mg/mL); ID histamine (100 μg/mL).

(Modified from Norman PS. Skin testing. In: Middleton E, Ellis EF, Reed CE, editors. Allergy: principles and practice. 2nd ed. St. Louis: CV Mosby; 1982, with permission from Elsevier.)

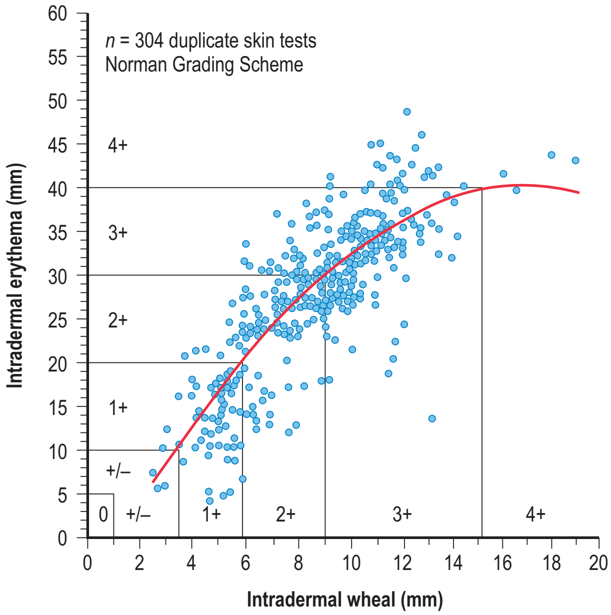


FIG 95.3 Correlation plot of the mean wheal (*x*-axis) and erythema (*y*-axis) in millimeters for the mean of 304 duplicate intradermal skin tests to dust mite (*Dermatophagoides pteronyssinus*) obtained in a population of individuals with sensitivity to dust-mites. The relationship is highly correlated ($r = 0.82$; $p < 0.001$) in grades 0 to 3+, indicating that either can be used to judge the degree of intradermal skin sensitivity. In highly allergic individuals (>35 mm erythema), however, the slope declines dramatically, indicating that wheal size may be more discriminating than erythema.

Oral Food Challenge Tests

Oral food challenges remain the definitive test for food hypersensitivity, since *in vitro* and other *in vivo* diagnostic methods suffer from suboptimal diagnostic specificity.[32] They are performed to establish an accurate diagnosis of food allergy or when skin tests are not possible because of atopic dermatitis or allergic GI disease or to assess the development of tolerance to a food. Passing a direct food challenge ideally permits reintroduction of the food into the patient's diet. Food-induced GI reactions (*e.g.*, nausea, colic, vomiting, and diarrhea) can occur minutes to hours after the consumption of food allergens by a sensitized individual.[33] Commonly eaten foods that contain potent allergens include cow milk (caseins, β-lactoglobulin, α-lactalbumin, lactoferrin, and bovine serum albumin), chicken egg white (ovalbumin, ovomucoid, ovotransferrin), cereal grains (wheat, rye, barley, oats), legumes (peanut, soybean, white bean), and fish and seafood (shrimp, crab, lobster, oyster).

Diagnosis of a food allergy begins with a medical history, which includes an assessment of diet diaries and elimination diets, as necessary. Skin tests and serological tests for IgE antibody to extracted food allergens and open food challenges with fresh and cooked foods can be used to confirm sensitization to suspected foods. No evidence is available to indicate that food-specific IgG or IgG4 antibody levels have any diagnostic value.[34] The double-blind placebo-controlled food challenge (DBPCFC) is considered the definitive diagnostic test for food allergy.[32] The open food challenge has its utility in clinical practice as a first-line challenge procedure as it is easier than DBPCFCs. Moreover, they are especially useful when the probability of a negative challenge is high, since a negative open challenge result can obviate the need for a DBPCFC and it reduces anxiety and thus reduces the risk of bias in infants and young adults. An extensive discussion of open challenges, DBPCFCs, and variables that influence their outcome is presented elsewhere.[32,33]

In 2001, prospectively collected sera from 100 children and adolescents who had been previously evaluated by a skin test and DBPCFC were analyzed in the ImmunoCAP autoanalyzer (Thermo Fisher Scientific/Phadia, Uppsala, Sweden) for food-specific IgE antibody.[35] A diagnostic level of IgE antibody specific for egg (7 kU/L), milk (17 kU/L), peanut (14 kU/L), and fish (20 kU/L) was identified, above which clinical reactivity could be predicted with more than 95% certainty. The study concluded that by measuring the concentration of food-specific IgE antibody with the ImmunoCAP system, a subset of children who were highly likely (>95% probability) to experience a clinical reaction to egg, milk, peanut, or fish could be identified. Importantly, this study and others subsequently have shown that judicious use of quantitative serological measurements for food-specific IgE antibody as measured by ImmunoCAP may be able to eliminate the need for time-consuming DBPCFCs in a significant number of children who are suspected of having food allergy. The relationship between IgE antichicken egg, cow's milk, and peanut as measured in the ImmunoCAP and Immulite (Siemens Healthcare Diagnostics, Tarrytown, NY, USA) has been investigated in an effort to extend published ImmunoCAP predictive values to reported Immulite measurements of IgE antibody to these foods.[36]

IN VITRO TESTING

Clinical immunology laboratories worldwide offer serological tests that are useful in the diagnosis and management of human

KEY CONCEPTS

Immunoglobulin E (IgE) (Reaginic) Antibody Detection

- Allergen-specific IgE can be detected by skin test using a puncture or intradermal administration of allergen or in the serum by laboratory-based immunoassays.
- In general, the intradermal (ID) skin test is more analytically sensitive than a puncture skin test, which is roughly comparable to the best *in vitro* methods for IgE antibody detection in serum.
- ID skin tests are the diagnostic procedure of choice in the workup of patients with suspected *Hymenoptera* venom and drug allergy, while both *in vitro* an *in vivo* assay methods are complementary for evaluating aeroallergen-related disease.
- Serological analyses of IgE antibody–specific for food allergens are often favored over extract-based skin test analyses in part because of more enhanced reagent quality control; however, the double-blind placebo-controlled food challenge (DBPCFC) remains the gold standard for definitive diagnosis of food allergies.
- IgE antibody to allergenic molecules (components) can in some cases (*e.g.,* peanut, hazelnut) provide clarity in terms of the specificity of the patient's sensitization profile (genuine vs cross-reactivity) and relative risk for mild versus serious systemic reactions.

allergic disease. Analytes commonly measured in these laboratories include the total serum IgE, IgE antibodies to hundreds of allergen specificities, *Hymenoptera* venom-specific IgG, the IgE antivenom-inhibition test, and mast cell tryptase. IgG antibody measurements to allergens other than *Hymenoptera* venom have not been shown to be clinically useful.[34] Basophil mediator and activation tests, although rarely offered as clinical tests because of the requirement for fresh blood, are useful investigational methods that are also reviewed in this section.

Total Serum IgE

Of the diagnostic allergy tests that are performed in the clinical immunology laboratory, total serum IgE is currently the only diagnostic allergy analyte regulated under the US Clinical Laboratory Improvement Amendment of 1988 (CLIA-88). Current commercial assays to measure the total level of IgE in serum employ nonisotopic labels, such as enzymes (horseradish peroxidase, alkaline phosphatase, β-galactosidase) or fluorophores, and have been cleared by the US Food and Drug Administration (FDA). The minimum detectable concentration of the commercial total serum IgE assays is between 0.5 and 1 μg/L. The intermethod agreement of the different commercial IgE assays is excellent (*e.g.,* intermethod coefficients of variation [CVs] typically <15%).[13,37] Nonatopic age-adjusted reference intervals for total serum IgE must be used for normative interpretation.[5]

Total IgE Measurements After Therapeutic Anti-IgE Administration

Omalizumab (anti-IgE) is currently used as a fourth therapeutic modality to supplement avoidance, pharmacotherapy, and immunotherapy in the management of persistent asthma and urticaria, and off-label for other IgE-mediated states (*e.g.,* allergic bronchopulmonary aspergillosis, pretreatment of food allergy patients receiving immunotherapy). Since its conception,[38] clinicians have desired to quantify the level of total and "free" (uncomplexed) IgE in anti-IgE treated patients as a rationale for treatment failures or to justify modification of a patient's dosing regimen to maximize treatment success. A systematic evaluation of the impact of therapeutic anti-IgE on the performance of clinically used total IgE assays showed variable interference that resulted in a 1.9–51.9% reduction in accuracy, depending on the assay.[39] Accurate quantitation of the level of uncomplexed or "free" IgE in the serum of treated patients has been more technically difficult. The performance of the existing free IgE assays has been heavily criticized,[40] leading some investigators to a conclusion that there may be no simple and analytically accurate clinical method for quantifying the level of free IgE in the serum of patients who have received omalizumab. This, however, may change with a new anti-IgE therapeutic that has been engineered so it is removed from circulation once it complexes IgE.[41]

Allergen-Specific IgE

Laboratories in the United States that perform clinical diagnostic allergy testing must be federally licensed, use a United States FDA-cleared assay method, and perform successfully in an external diagnostic allergy proficiency survey (*e.g.,* College of American Pathologists [CAP] SE Survey).[13,37] These assays, which have been through rigorous validation, have achieved unsurpassed intraassay precision, interassay reproducibility, and a high degree of quantification.[13] Their basic design can be traced to the first IgE antibody assay, the RAST, reported by Wide et al. in 1967.[12]

Allergen

The most highly variable component of the IgE antibody assay is the allergen-containing reagent. Allergens are mixtures of molecules, typically proteins, glycoproteins, lipoproteins, or protein-conjugate chemicals or drugs that have been solubilized from a defined (usually biological) source, a portion of which can elicit an IgE antibody response in exposed and genetically predisposed individuals. They possess common properties of stability to processing (*e.g.,* heat) and digestion because of multiple cysteine linkages. They tend to be abundant in nature, form aggregates or polymers, commonly interact with lipid structures, and serve to defend their biological source. Cross-tabulation of the protein family (PFAM) database (n = 16 230 proteins families) with the Structural Database of Allergenic Proteins (SDAP) identified 130 PFAMs in the Allergenic Family database of allergenic proteins. Thus importantly, allergens comprise a small fraction of protein families with particular structures and biological functions.[42] The Clinical Laboratory Standards Institute (CLSI) has an established international guidance document that defines the expected performance characteristics of allergenic materials used in immunological assays for human IgE antibodies.[13] It provides a compendium of the genus and species of all the allergen specificities of clinical interest, subdividing them into extract and component allergens and indicating whether they are well-documented or rare. They are categorized on the basis of their sources, into weed pollen, grass pollen, tree pollen, animal dander, mold, house dust-mite fecal material, parasites, insect venoms, occupational allergens, foods, and drugs. Except for drugs, these extracts are complex heterogeneous mixtures that contain both nonallergenic and allergenic proteins. Some allergens share structural similarity or cross-reactive epitopes, and others possess unique IgE antibody-binding determinants. Among the different species of a genus, such as ragweed (*e.g.,* Canyon, Desert, Giant, Short, Silver, Southern, Western), there is extensive allergenic cross-reactivity resulting from structural similarity. Extensive allergenic cross-reactivity has also been documented within other pollen groups, such as the grasses (June, Brome, Timothy, Perennial Rye, Fescue, Orchard, Red Top, Salt, Sweet Vernal, Velvet).

TABLE 95.2 **Principal Allergen Families and Their Associated Biological Functions**

Family	Function	Diagnostic Utility and Clinical Features	Examples
Profilins	Actin-binding proteins involved in the dynamic turnover and restructuring of the actin cytoskeleton. Highly conserved, extremely cross-reactive and ubiquitous proteins in pollen and plant foods	Positive skin test and immunoglobulin E (IgE) responses to (nonrelated) pollen species (often including grasses) and plant food extracts are indicative for IgE cross-reactivity to profilins. After proof of sensitization to one profilin pollen and plant food extracts are of no further use because of their subsequent lack of analytical specificity	Bet v 2 (birch); Phl p 12 (Timothy Grass), Hev b 8 (Natural Rubber Latex), Mal d 4 (apple),
Serum albumins	Highly cross-reactive plasma protein carriers involved in transport of hormones, enzymes, hemin, and fatty acids. Also maintains oncotic pressure	Positive skin test and IgE responses to furred animals indicate IgE cross-reactivity to mammalian serum albumins. Proof of sensitizations to one serum albumin can explain clinical symptoms to rarely or uncooked meat ("cat-pork-syndrome")	Fel d 2 (cat); Can f 3 (dog),
Pathogenesis-related proteins family 10 (PR-10), Bet v 1- homolog	Plant defense proteins; Bet v 1 is a quercetin-3-O-sophoroside-binding molecule (17 kDa); inflammation response proteins	Positive Bet v 1-specific IgE reflects sensitization to fagales tree pollen (*i.e.,* hazel, alder, birch, beech, oak). Thermo and digestion labile Bet v 1-homologs in fruits, legumes, and vegetables. Bet v 1-sensitized individuals can cause predominantly oropharyngeal symptoms after consumption of raw food items (*i.e.,* pip fruits, stone fruits, tree nuts, carrots, soy).	Bet v 1 (birch); Cor a 1 (hazelnut); Mal d 1 (apple); Gly m 4 (soy)
Polcalcin	Cross-reactive, ubiquitous calcium binding proteins in pollen; involved in calcium regulation	Positive skin test and IgE responses to (nonrelated) pollen species are indicative for IgE cross-reactivity to polcalcins. After proof of sensitization to one, polcalcin pollen extracts are of no further use because of their subsequent lack of analytical specificity	Phl p 7 (Timothy Grass); Bet v 4 (birch); Amb a 10 (Short Ragweed)
Nonspecific lipid transfer proteins (nsLTP, PR-14)	Inflammation response proteins; responsible for shuttling phospholipids and other fatty acids between cell membranes	Primary food allergen after sensitization to peach lipid transfer protein (LTP) (possibly through the skin?); predominantly found in the Mediterranean subjects; variable degree of cross-reactivity between the thermally stable LTPs in fruits and vegetables causing frequently oropharyngeal and sometimes systemic symptoms (*i.e.,* exercise-induced)	Pru p 3 (peach); Ara h 9 (peanut); Cor a 8 (hazelnut)
Lipocalins	A highly heterogeneous group of extracellular proteins within various subfamilies involved in transport of small hydrophobic molecules, such as steroids, bilins, retinoids, and lipids	Positive skin test and IgE responses to more than one or many furred animals are indicative for IgE sensitization and subsequent serological cross-reactivity to lipocalins of a certain subfamily. As a consequence animal extracts are of no further use because of their subsequent lack of analytical specificity	Fel d 4, 7 (cat); Can f 1, 2, 4, 6 (dog)
Parvalbumins	Calcium binding proteins; localized in fast-contracting muscles and being involved in calcium signaling	Thermostable and digestion-stable major fish allergen with fairly high, but not complete cross-reactivity and high abundance in almost all fish species. Only limited sequence homology and no cross-reactivity between fish and shellfish Ca^{++}-binding proteins	Gad c 1 (cod); Cra c 4, 6 (shrimp)
Tropomyosins	Integral components of actin filament that play a role in regulating muscle contraction. Also regulate actin filament stability in nonmuscle cells	Thermostable and digestion-stable major shellfish allergen with broad cross-reactivity and high abundance in all shellfish species. IgE sensitization is associated to allergic reactions to various panels of shellfish species	Der p 10 (dust mite); Pen m 1 (shrimp)
Seed storage proteins	Heterodimeric, stable and highly abundant proteins involved in nutrient storage: *e.g.,* 2S albumins, 7/8S globulins (vicilins), 11S globulin (legumins)	Thermostable and digestion-stable important primary food allergens in legumes (peanut, soy), tree nuts (hazelnut), capsule fruits, and seeds with limited cross-reactivity within the subfamilies of different species; primary sensitization starts in early childhood and can persist lifelong; high allergen-specific IgE levels are associated with systemic allergic reactions	Ara h 2 and 1 and 3 (peanut); Cor a 14 and 11 and 9 (hazelnut)

Modified from the I/LA20 Guidance Document from the Clinical Laboratory Standard's Institute. Hamilton RG, Matsson PNJ, Chan S, et al. Analytical performance characteristics, quality assurance and clinical utility of immunological assays for human immunoglobulin E (IgE) antibodies of defined allergen specificities. 3rd ed. I/LA20-A3, International CLSI-Guideline. Wayne, PA: Clinical Laboratory Standards Institute; 2016.

In contrast, other grass pollens, such as those produced by Bermuda Grass, Johnson Grass, and cultivated corn, oat, and wheat, are minimally cross-reactive (allergenically distinct). Variations in the allergenic content of extracted source materials, the extraction process from the raw source material, allergen-reagent manufacturing methods, differential binding to various allergosorbent supports, instability during storage, heterogeneity of internal reference allergen standards, and differences in characterization procedures (antisera, assays) make the production of reproducible allergens for *in vitro* use a challenge.

Cross-reactivity has also been shown at the allergen component level. There are 10 principal allergen families that show structural similarity and extensive cross-reactivity (Table 95.2).[42,43] The most prominent is illustrated by the Pathogenesis-Related (PR)–10 family of allergens, also known as the *Bet v 1 homologues*. These are small (17-kDa) proteins in many plant species that transport steroids and exhibit low stability at high pH and in the presence of digestive enzymes. The first allergen in this family was identified from birch pollen (Bet v 1). Others with high amino acid sequence homology include Cor a 1-hazelnut, Mal d 1-apple, Pru p 1-peach, Gly m 4-soybean, Ara h 8-peanut, Aln g 1-alder, Act d 8-kiwi, Api g 1-celery, and Dau c 1-carrot. Other component based cross-reactivity groups include the profilins, nonspecific lipid transfer proteins tropomyosins, serum albumins, polcalcins, lipocalins, parvalbumins, storage-binding proteins, and carbohydrate cross-reactive determinants (CCDs) (see Table 95.2). Each of these cross-reactive allergen families is extensively discussed in the *Handbook on Molecular Allergology.*[44]

Calibration

The second attribute that varies widely among commercially available IgE antibody assays is calibration algorithm and methodology. There exists no internationally recognized polyclonal human IgE antibody reference preparation. Thus all three autoanalyzers currently in commercial use in the United States have found a heterologous interpolation procedure to be optimal in which allergen-specific IgE antibody response data are interpolated from a total serum IgE calibration curve. This procedure is valid as long as the total IgE calibrators and the patient's allergen-specific IgE antibody levels dilute out in parallel so that parallelism is maintained. The assays report IgE antibody levels in the same units (kUa/L), using internal total IgE calibrators that are cross-verified and traceable to the World Health Organization (WHO) 11/234 IgE international reference preparation.[45] This calibration system allows interpolation of IgE antibody results from a limit of quantitation of 0.1 kUa/L to 100 kUa/L levels of IgE antibody. In terms of quantitation, at least one of the IgE antibody autoanalyzers (ImmunoCAP) has demonstrated equivalence in which 1 kUa/L of chimeric allergen-specific IgE antibody was shown to be equivalent to 1 kU/L (2.4 ng/mL) of total serum IgE.[46]

Single-Plex, Multiallergen, and Multiplex Assays

The autoanalyzers in common clinical use are single-plex (or monoplex) assays in which one analyte (*e.g.,* IgE anti–cat dander) is measured in a single analysis. In contrast, a multiplex antibody assay allows many specificities of a single antibody isotype (*e.g.,* IgE) to be individually detected and semiquantified in a single analysis. These are distinguished from a multiallergen screening assay, which is a form of single-plex analysis that involves the use of a mix of multiple allergens that is immobilized on the same allergosorbent. The purpose of the multiallergen assay is to simultaneously screen a serum for IgE antibody to a concise number (*e.g.,* 10–15) allergens either of the same allergen source type (*e.g.,* foods: chicken egg, cow's milk, peanut, soybean, cod fish) or diverse sources (*e.g.,* respiratory allergens as an aeroallergen mix: pollen from select trees, grasses, weeds, pet epidermals, dust mites, molds).[47] A qualitative result (positive or negative) is generated, and it serves as an inexpensive and efficient single analysis to assess the general atopic status (IgE positivity) of an individual. One multiaeroallergen screening assay is being used to assess the atopic status (genetic predisposition to produce IgE antibody) of research subjects enrolling in clinical studies of asthma.[48] Clinically, the multiaeroallergen screening assay has a high negative predictive power and thus is used to rule out IgE-mediated allergic disease where the suspicion based on the clinical history is weak.

The availability of unlimited quantities of molecular allergens has made it possible to develop multiplex chip based microarrays for diagnostic allergy confirmatory testing. The first microarray for semiquantification of IgE antibody involved a preactivated glass slide or chip on which 94 purified allergens were each immobilized in triplicate microdot arrays.[49] The present day version of this assay is the immune solid phase allergen chip or ISAC (Thermo Fisher Scientific/Phadia, Uppsala, Sweden) in which 40 µL of serum is applied to a chip that has 112 individual allergenic molecules immobilized.[50] IgE anti–cow's milk components as measured in the multiplex ISAC have agreed well with those obtained in a single-plex autoanalyzer using the same allergen specificities.[51] Other groups have tried alternative multiplexing technologies to detect IgE antibodies that have included immobilizing allergen extracts on chips by using Luminex bead–based suspension arrays, employing nanotechnology biosensors, detecting surface plasmon resonance, and using plates equipped to produce electrical pulse generated chemiluminescence.[51,52] When selecting a single-plex or multiplex assay format, one must consider their strengths and weaknesses. Single-plex systems have the advantage of greater analytical sensitivity or a lower limit of quantitation, greater precision and accuracy, more established internal and external quality control, and wider global availability of technology. In contrast, multiplex systems provide increased speed of analysis with reduced turnaround times, conservation of sample volume, greater simplicity, and reduced technical and reagent costs.[52]

The quality of allergen-specific IgE antibody results reported from clinical diagnostic allergy laboratories is not uniformly equivalent. In addition to the variability of results for a given serum between assays from different manufacturers as a result of allergen and calibration variance, the positive thresholds used for the same assay by different laboratories varies (*e.g.,* 0.1, 0.35, and 0.7 kUa/L).[53] For this reason, physicians requesting IgE antibody testing bear some responsibility for determining the quality of the results they receive. In the United States, testing should be performed in a clinical laboratory that is federally licensed for highly complex immunology clinical testing under CLIA-88 (verified by requesting a copy of the federal laboratory license).[54] The requesting physician should inquire about the assay method used, the source of its reagents, and how assays are quality controlled by the laboratory. As part of the formal record, the assay method used in patient analysis should be indicated on the final report.

Competitive IgE Antibody Inhibition Assay

The competitive inhibition format of the IgE antibody assay has been used by researchers, allergen manufacturers, and regulators to determine the relative potency of allergen extracts.[55] One practical research application of the IgE antibody inhibition assay has been as a tool for monitoring the concentration of allergens released into environments (*e.g.,* latex allergen in hospital air[56]; air sampling of airplanes for aerosolized peanut allergen). In the diagnostic allergy laboratory, the competitive IgE antibody inhibition assay has been used analytically to confirm an IgE antibody assay's minimum detectable dose (sensitivity) and nonspecific binding and to document the extent of cross-reactivity by determining the allergen specificity of IgE antibody.[13]

The one clinical application of the IgE antibody inhibition assay has been as an adjunct to define the appropriate therapeutic composition of venoms for patients who have insect-sting allergy with multiple potentially cross-reactive sensitivities and who have elected to receive immunotherapy.[57] Indications for this test include a strong skin reactivity; a high level of serum IgE antibody to yellow jacket venom (YJV) and a weak skin reactivity; or a low level of serum IgE antibody specific for *Polistes* wasp venom (PWV). The structural similarity between *Vespid* and

Polistes wasp phospholipase A1/B (Ves g 1; Pol a 1) and Antigen 5 (Ves g 5; Pol a 5) frequently produces IgE antibody cross-reactivity. Sera from 305 patients with *Hymenoptera* venom allergy with >2 ng/mL of IgE antibody to YJV and PWV were evaluated in the IgE antibody inhibition assay to document its clinical utility. The diagnostic question for these patients is whether PWV should be included in the venom immunotherapy together with yellow jacket or mixed vespid venom. Using this procedure, the venom-specific IgE antibody inhibition assay identified one-third (36.4%) of subjects with a primary YJV sensitivity who were candidates for exclusion of PWV from their immunotherapy regimen because their IgE anti-PWV was >95% cross-inhibitable with soluble YJV.[57]

More recently, molecular venom allergens may be particularly useful in clarifying whether patients with venom allergies with IgE antibody positivity for two or more distinct venoms (YJV, HBV, and/or PWV) reflects true multiple venom sensitization or whether the observed multiple positivity results from either clinically irrelevant cross-reactive carbohydrate-specific IgE antibodies or common protein epitopes of homologous venom antigens.[58] The important carbohydrate cross-reactive determinants (CCD) in insect venoms involves an α1,3-linked fucose residue that is attached to the innermost N-acetylglucosamine carbohydrate on venom proteins.[59] For YJV and HBV, the known cross-reactive proteins include hyaluronidase (Api m 2, Ves v 2), dipeptidyl peptidase (Api m 5, Ves v 3), and vitellogenin (Api m 12, Ves v 6). IgE antibody cross-reactivity between YJV and PWV is independent of CCD and can occur as a result of common epitopes on homologous antigens: Antigen 5 (Ves v 5, Pol d 5) and phospholipase A 1 (Ves v 1 and Pol d 1).[60] Specific IgE antibody measurements against rVes v 5 and rPol d 5 alone were able to aid in differentiating sensitivity against *Vespula* and *Polistes* venoms in some patients. However, Savi et al.[61] concluded that the venom competitive inhibition analysis remains the optimal serological test at present for differentiating between true and cross-reactivity–related dual *Vespula* and *Polistes* venom sensitization. Moreover, when rPol d 1 becomes available for testing in parallel with currently available Ves v 1, Ves v 5, and Pol d 5, the effectiveness of molecular allergen based differentiation of dual *Vespula* and *Polistes* venom sensitization will need to be reassessed.

CLINICAL PEARLS

*Clinical Utility of Immunoglobulin E (IgE) Anti-***Hymenoptera** *Venom Inhibition Test*

- Useful in identifying venom cross-reactive IgE antibodies and selecting appropriate venoms for immunotherapy.
- One-third of venom-allergic patients with concomitant yellow jacket venom (YJV) and *Polistes* wasp venom (PWV) sensitivity can be treated with YJV alone or mixed vespid venoms, owing to >95% cross-reactive IgE anti-PWV with soluble YJV.

Allergen-Specific IgG

Allergen immunotherapy is known to enhance the production of specific IgG "blocking" antibodies.[62,63] Quantitative measurements of allergen-specific total IgG or IgG subclass antibodies in studies of allergic rhinitis have not generally correlated with the control of clinical symptoms in individual patients. However, clinically successful immunotherapy is almost always accompanied by high serum levels of allergen-specific IgG (typically IgG1 and IgG4) blocking antibodies.[63] Interestingly, despite a decrease to preimmunotherapy baseline levels after 2 years of immunotherapy discontinuation, overall functional inhibitory activity, as measured by an IgE-dependent facilitated allergen binding assay, appears to be maintained.[63] For patients *with Hymenoptera* venom allergy, specific IgG antibody measurements have been used as an indicator of effective immunotherapy. Quantitative venom-specific IgG antibody levels may be useful in individualizing the dose and frequency of injections while maximizing the protective effects. However, their clinical utility may be restricted to the first 4 years of venom immunotherapy.[64] In contrast, the presence or levels of IgG antibodies specific for food antigens have shown no correlation with the results of positive DBPCFCs, and they are not indicated in the diagnostic workup of a patient with suspected food allergy.[34]

CLINICAL PEARLS

Immunoglobulin G (IgG) Antibody Measurements

- Clinically successful aeroallergen immunotherapy is almost always accompanied by high serum levels of allergen-specific IgG (predominantly of the IgG1 and IgG4 subclasses).
- Quantitative venom-specific IgG antibody levels can be of value in individualizing venom doses and frequencies for patients on venom immunotherapy for up to 4 years.
- Food-specific IgG and IgG4 assay results do not correlate with the results of double-blind placebo-controlled food challenges and are not clinically indicated.

Mast-Cell Tryptase

Mast cells have been identified in skin, respiratory, and digestive tract connective tissues and distinguished on the basis of the neutral proteases present in their secretory granules. One group contains only tryptase, whereas the other contains both tryptase and chymase.[65] Mast-cell tryptase (MW 134 kDa) is a serine esterase with four subunits, each having an enzymatically active site. A resting mast cell contains 10–35 picograms (pg) of tryptase that is stored attached to heparin. When dissociated from heparin, it rapidly degrades into its monomers and loses enzymatic activity. As basophils have ~500-fold less tryptase compared with mast cells, elevated tryptase levels in serum are considered a relatively specific indicator of mast-cell involvement in a clinical reaction. Unstimulated tissue mast cells continually secrete immature protryptase into the tissue, and it diffuses into the circulation to provide a measure of total mast-cell number. α-Protryptase and β-protryptase represent the bulk of the immature tryptase in nonanaphylactic sera. α-Protryptase remains enzymatically inactive, whereas some of β-protryptase is autoprocessed from the proform within the mast cell into the mature enzyme by a dipeptidase where it is stored in granules.[66] Only upon activation of mast cells are both the pro and mature forms of tryptase secreted in parallel with prestored histamine and newly generated vasoactive mediators. Total tryptase in serum from healthy humans ranges from 1 to 11.4 μg/L (average 3–5 μg/L). Mature tryptase is normally undetectable (<1 μg/L) in serum from healthy individuals who have no history of anaphylaxis during the preceding hours. A noncompetitive two-site fluorescent enzyme immunoassay (Phadia, ImmunoCAP) is available to measure total tryptase in serum. Elevated levels of total tryptase (>11.4 μg/L) can be detected in serum 1–4 hours after the onset of systemic anaphylaxis with hypotension. Baseline levels of >20 μg/L are detected in most individuals with systemic mastocytosis.[66]

Recommended serum collection times for tryptase quantification range from 30 minutes to 4 hours after the onset of an acute event. Because serological tests for mature tryptase are not widely available, it is important to compare an acute event total tryptase level (within 4 hours) with a baseline total tryptase 24 hours after all signs and symptoms of the event have subsided.[66] Postmortem specimens are often difficult to analyze for tryptase because of high viscosity related to gross hemolysis. However, reported mature tryptase levels in postmortem cases of fatal anaphylaxis have ranged from 12 to 150 μg/L in all nine fatalities caused by *Hymenoptera* venom and in six of eight food-induced fatalities.[67] The frequency of a severe reaction to venom from a *Hymenoptera* sting increases significantly with higher serum baseline total tryptase concentrations.[68,69] The clinical indications for and diagnostic value of tryptase as a marker of anaphylaxis and mastocytosis are reviewed elsewhere.[68]

Basophil Mediator Release Assays

There are a number of diagnostic assays for sensitization that use the basophil as the indicator modality. Some assays monitor the magnitude of mediator release from basophils, and others assess the degree of cell surface marker expression following exposure to allergen.

Histamine Release Assay

The potent vasoactive mediator histamine is stored in cytoplasmic granules of basophilic leukocytes and mast cells. It is released along with other mediators of inflammation in response to both immunological and nonimmunological stimuli.[70] The basophil histamine release (BHR) assay has been particularly useful as a quantitative assay of allergen potency and as an *in vitro* model for the study of triggering mechanisms of mediator release from basophils. In its most basic form, peripheral blood leukocytes are isolated from a donor and incubated with varying concentrations (*e.g.*, 3–10-fold dilutions) of allergen extract or antihuman IgE as a positive control. Histamine release is complete within 30 minutes, and then histamine in the supernatant is measured by enzymatic, radiometric, or spectrophotofluorometric techniques.[71] Details of the BHR assay are given elsewhere.[70,72]

Patient sensitivity for a given allergen can be determined with a positive BHR test. The results are highly correlated with those determined by skin testing[71] and bronchoprovocation.[73] Although the BHR test has been almost exclusively used in research laboratories because of its expense, time-consuming nature, and the need for fresh blood (<24 hours old), it can be successfully applied to the clinical diagnosis of allergic disease in selected cases. Its results parallel those of other IgE antibody tests. BHR has also been a useful tool for clarifying discrepancies between skin test and serological IgE antibody test results.

Leukotriene C4 (LTC4) Release Assay

An assay method for measuring LTC4 released from allergen-activated basophils has been reported as the cellular antigen stimulation test (CAST)–enzyme-linked immunosorbent assay (ELISA).[74] The LTC4 assay is designed for use with either whole blood preparations or washed leukocytes. Using dust-mite, food, *Hymenoptera* venoms[75] and drugs as challenge allergens, the observed diagnostic sensitivity of the CAST compared with the combination of a clinical history and skin test ranged from 18% with aspirin and 50–85% for selected food allergens.[76] The reported diagnostic specificity of the CAST in the same studies ranged from 67% to 100%. These data indicate that the CAST assay is not sufficiently sensitive for effective clinical use in the diagnosis of IgE-mediated sensitivities to β-lactam or nonsteroidal antiinflammatory drugs (NSAIDs). Its utility in the diagnosis of sensitization to other allergen specificities appears more promising, but further documentation involving clinical studies is required.

Utility of Mediator Release Assays as Diagnostic Tests

Despite their unquestioned value as research methods, the basophil histamine and LTC4 release assays are rarely used clinically in the routine diagnosis of human allergic disease, and it is unlikely that this trend will change in the foreseeable future.

Flow Cytometry Basophil Activation Assays

In the late 1990s, basophils were shown to upregulate the expression of a number of surface proteins (*e.g.*, CD45, CD63, CD69, and CD203c) when activated by allergen.[76-78] CD63 is a member of the transmembrane-4 superfamily that is expressed on basophils, mast cells, macrophages, and platelets.[76,77] In resting basophils from individuals without atopia and allergies, CD63 is attached to intracytoplasmic granules. Activated IgE-sensitized basophils express a high density of CD63 on their surface. By quantifying ~500 basophils with gated flow cytometry, the percentage of activated basophils can be identified, after correcting for spontaneous CD63 expression. A variety of criteria are used to identify an allergen-induced positive response, such as a stimulation index (allergen-induced/basal ratio) >2. One confounding variable in this analysis is the adherence of CD63-expressing platelets in basophil preparations that can sometimes generate false-positive results.

The flow assay stimulation test (FAST), which is also referred to as the basophil activation test (BAT) or BasoTest, uses CD63 as a marker for activation. Its performance was evaluated by using minor determinant mixture, benzyl penicilloyl polylysine, penicillin, ampicillin, amoxicillin, and cephalosporin as stimulating antigens.[78,83] Subjects (n = 58) with a history of immediate-type reactions to β-lactam antibiotics or nonallergic controls (n = 30) were evaluated in the CD63 FAST/BAT and Phadia ImmunoCAP system for drug-specific IgE levels. Relative to the clinical history, the CD63 FAST/BAT displayed a diagnostic sensitivity and specificity of 50% and 93%, respectively. Diagnostic sensitivity was marginally increased to 66% by simultaneously using the drug-specific IgE antibody serology result. The study concluded that the CD63 basophil activation test could be helpful in supporting the diagnosis of IgE-mediated allergy to β-lactam drugs when used in conjunction with an additional diagnostic test, such as IgE antibody serology.[79]

CD203c (also known as *neural cell surface differentiation antigen, E-NIPP3 PD-1β, 97A6, B10,* and *gp130rb13–6*) is a member of the ectonucleotide pyrophosphate phosphodiesterase (E-NPP) family.[80] It is expressed only on IgE-bearing basophils, mast cells, and their progenitors and is upregulated after activation of IgE-sensitized basophils with allergen or anti-IgE in a manner similar to CD63. The CD203c-based flow assay is performed in a manner analogous to the CD63-based assay. Its advantage over CD63 is its restricted expression on basophils in peripheral blood, which minimizes concern about the adherence of platelets that may produce false-positive CD63 results. Additionally, there is no need for the use of an additional fluorescent anti-IgE reagent to gate the basophils. Basophil activation as determined by CD203c expression was studied in patients with allergy to *Hymenoptera* venom and healthy controls in which their basophils were activated with anti-IgE or bee/*Vespid* venoms.[81] Fifteen minutes after stimulation, basophils from 15 wasp-sensitized patients upregulated CD203c expression from 4.2-fold to 13.5-fold. Moreover, basophils from six patients with honeybee venom allergy upregulated CD203c by a mean of 8.3-fold. The study concluded that the CD203c-based flow assay could confirm the presence of venom-specific IgE antibody on basophils in 20 of 22 of patients (91%) with a clinical history of and skin test positivity to venom allergy. One false-positive result was observed among 13 healthy controls with a negative result on the skin test.

Diagnostic Utility of Basophil Activation Flow Cytometric Assays

The utility of basophil activation flow cytometric assays is limited by a number of technical concerns.[79,82] First, endotoxin-free whole blood must be received within 24 hours by the laboratory, where it can processed by skilled technologists. Second, the blood is preincubated with stimulation buffer often containing IL-3, which primes basophils. It is unclear what effect IL-3 has on altering mediator release or CD63/CD203c upregulation, and thus false-positive mediator release and/or biomarker surface expression are potential concerns. Third, varying concentrations of crude allergen extract or recombinant allergen or anti-IgE are incubated with the cell preparation. Unfortunately, crude extracts lack potency estimates and can often be cytotoxic. Thus each stimulating allergen needs to be prequalified with basophils from well-characterized patients before use. Fourth, criteria for defining positive assay results vary among the laboratories performing the test, and stimulating allergen preparations from even the same manufacturer can have variable potency. Fifth, there is a concern about false-positive results related to platelet adherence on basophils in the CD63 based BAT assay. Possibly most important, when the performance characteristics were directly compared in clinical studies to puncture skin test results, the diagnostic sensitivity and specificity of the flow-cytometric analysis of *in vitro* allergen-activated basophils was less than optimal. For these reasons, cell-based cytometric methods will most likely remain useful research techniques for the investigation of allergic disease, but they will have limited application to clinical diagnosis.

ON THE HORIZON

New trends in laboratory methodology for the assessment of human allergic diseases include the following:

- Although there is a transition from the use of crude allergen extracts to allergenic components in immunoglobulin E (IgE) antibody serological assays, extracts will remain the principal allergen reagent source for the foreseeable future because of their comprehensive nature.
- Multiplexing platforms are being increasingly developed to allow rapid simultaneous detection of IgE antibodies to many allergenic components using microliter quantities of serum; however, due to their semiquantitative nature, fixed panel of specificities, and lower analytical sensitivity, single-plex assays will remain the most cost-effective and widely used assay format for the foreseeable future.
- Qualitative "point of care" IgE antibody assays to rapidly assess sensitization to aeroallergens during the patient's visit will be slow to be adopted because of their limited/fixed allergen menus, lower diagnostic sensitivity, and concern about potential patient misinterpretation of results.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Prausnitz C, Kustner H. Studine uber die Ueberemfindlichkeil. Zentralbl Bakteriol Mikrobiol Hyg 1921;86:160, [in German].
2. Ishizaka K, Ishizaka T. Physiochemical properties of reaginic antibody. I. Association of reaginic activity with an immunoglobulin other than gamma A or gamma G globulin. J Allergy 1967;37:169.
3. Johansson SGO. Raised levels of a new immunoglobulin class (IgND) in asthma. Lancet 1967;2:951.
4. Hamilton RG. The science behind the discovery of IgE. J Allergy Clin Immunol 2005;115:648.
5. Hamilton RG. Human immunoglobulins. In: O'Gorman MRG, Donnenberg AD, editors. Handbook of human immunology. 2nd ed. Boca Raton, FL: CRC Press; 2008. p. 63–106.
6. Barbee RA, Halomen M, Lebowitz M, et al. Distribution of IgE in a community population sample: correlations with age, sex and allergen skin test reactivity. J Allergy Clin Immunol 1981;68:106.
7. Kleine-Tebbe J, Poulsen LK, Hamilton RG. Quality management in IgE-based allergy diagnostics. J Lab Med 2016;40:81–96.
8. Wittig HJ, Belloit J, DeFillippi I, et al. Age-related serum IgE levels in healthy subjects and in patients with allergic disease. J Allergy Clin Immunol 1980;66:305.
9. Dati F, Ringel KP. Reference values for serum IgE in healthy nonatopic children and adults. Clin Chem 1982;28:1556.
10. Lin H, Boesel KM, Griffith DT, et al. Omalizumab rapidly decreases nasal allergic response and FcεR1 on basophils. J Allergy Clin Immunol 2004;113:297.
11. Saini SS, Bindslev-Jensen C, Maurer M, et al. Efficacy and safety of omalizumab in patients with chronic idiopathic/spontaneous urticaria who remain symptomatic on H1 antihistamines: a randomized, placebo-controlled study. J Invest Dermatol 2015;135:67–75.
12. Wide L, Bennich H, Johansson SGO. Diagnosis by an in vitro test for allergen specific antibodies. Lancet 1967;2:1105.
13. Hamilton RG, Matsson PNJ, Chan S, et al. Analytical performance characteristics, quality assurance and clinical utility of immunological assays for human immunoglobulin E (IgE) antibodies of defined allergen specificities. 3rd ed. I/LA20-A3, International CLSI-Guideline. Wayne, PA: Clinical Laboratory Standards Institute; 2016.
14. Hamilton RG, Oppenheimer J. Serological IgE analyses in the diagnostic algorithm for allergic disease. J Allergy Clin Immunol Pract 2015;3:833–40.
15. Yunginger JW, Ahlstedt S, Eggleston PA, et al. Quantitative IgE antibody assays in allergic diseases. J Allergy Clin Immunol 2000;105:1077.
16. Adkinson NF Jr, Hamilton RG. Clinical history-driven diagnosis of allergic diseases: utilizing in vitro IgE testing. J Allergy Clin Immunol Pract 2015;3:871–6.
17. Hamilton RG. Allergic sensitization is a key risk factor for but not synonymous with allergic disease. J Allergy Clin Immunol 2014;134:360–1.
18. Hamilton RG. Diagnostic in vivo and in vitro methods in insect allergy. In: Freeman T, Tracy J, editors. Stinging insect allergy: a clinician's guide. NY: Springer Science; 2017. p. 85–99.
19. Guerin B, Watson RD. Skin tests. Clin Rev Allergy 1988;6:211.
20. Mirela Chiriac A, Bousquet J, Demoly P, et al. In vivo methods for study and diagnosis of allergy: skin tests, techniques and interpretation. In: Adkinson NF Jr, Bochner BS, Burks AW, et al, editors. Middleton's allergy: principles and practice. 8th ed. Philadelphia, PA: Elsevier Saunders; 2014. p. 1119.
21. Lewis T, Grant R. Vascular reactions of the skin to injury. Part II. The liberation of a histamine-like substance in injured skin, the underlying cause of factitious urticaria and of wheals produced by burning; and observations upon the nervous control of certain skin reactions. Heart 1924;209:1924.
22. Pepys J. Skin testing. Br J Hosp Med 1975;14:412–25.
23. Tversky JR, Chelladurai Y, McGready J, et al. Performance and pain tolerability of current diagnostic allergy skin prick test devices. J Allergy Clin Immunol Pract 2015;3:888–93.

24. Mantoux C. Intradermoreaction de la tuberculose. CR Acad Sci 1908;147:355.
25. Turkeltaub PC, Rastogi SC, Baer H, et al. A standardized quantitative skin test assay of allergen potency and stability. Studies on the allergen dose-response curve and effect of wheal, erythema and patient selection on assay results. J Allergy Clin Immunol 1982;70:343.
26. Slater JE, Menzies SL, Bridgewater J, et al. The US Food and Drug Administration review of the safety and effectiveness of nonstandardized allergen extracts. J Allergy Clin Immunol 2012;129:1014–19.
27. McCann WA, Ownby DR. The reproducibility of the allergy skin test scoring and interpretation by board-certified/board-eligible allergists. Ann Allergy Asthma Immunol 2002;89:368–71.
28. van Kampen V, de Blay F, Folletti I, et al. EAACI position paper: skin prick testing in the diagnosis of occupational type I allergies. Allergy 2013;68:580–4.
29. Cockcroft DW. Bronchial challenge testing. In: Adkinson NF Jr, Bochner BS, Burks AW, et al, editors. Middleton's allergy: principles and practice. 8th ed. Philadelphia, PA: Elsevier Saunders; 2014. p. 1042.
30. Crapo RO, Casaburi R, Coates AL, et al. Guidelines for methacholine and exercise challenge testing—1999. Am J Respir Crit Care Med 2000;161:309–29.
31. Scadding GW, Hansel TT, Durham SR. Nasal Provocation Testing. In: Adkinson NF Jr, Bochner BS, Burks AW, et al, editors. Middleton's allergy: principles and practice. 8th ed. Philadelphia, PA: Elsevier Saunders; 2014. p. 652.
32. Wood RA. Oral food challenge testing. In: Adkinson NF Jr, Bochner BS, Burks AW, et al, editors. Middleton's allergy: principles and practice. 8th ed. Philadelphia, PA: Elsevier Saunders; 2014. p. 1357.
33. Nowak-Wegrzyn A, Burks AW, Sampson HA. Reactions to foods. In: Adkinson NF Jr, Bochner BS, Burks AW, et al, editors. Middleton's allergy: principles and practice. 8th ed. Philadelphia, PA: Elsevier Saunders; 2014. p. 1310.
34. Stapel SO, Asero R, Ballmer-Weber BK, et al. Testing for IgG4 against foods is not recommended as a diagnostic tool: EAACI Task Force Report. Allergy 2008;63:793–6.
35. Sampson HA. Utility of food specific IgE concentrations in predicting symptomatic food allergy. J Allergy Clin Immunol 2001;107:891.
36. Hamilton RG, Mudd K, White MA, et al. Extension of food allergen-specific IgE ranges from the ImmunoCAP to the IMMULITE System. Ann Allergy Asthma Immunol 2011;107:139–44.
37. Participants summary report. College of American pathologists, diagnostic allergy (SE) survey, Cycle B. Northfield, IL: College of American Pathologists; 2015.
38. Chang TW, Davis FM, Sun NC, et al. Monoclonal antibodies specific for human IgE producing B cells: a potential therapeutic for IgE mediated allergic diseases. Biotechnology (NY) 1990;8:122–6.
39. Hamilton RG. Accuracy of US Food and Drug Administration cleared IgE antibody assays in the presence of anti-IgE (omalizumab). J Allergy Clin Immunol 2006;117:759–66.
40. Baker DL, Peng K, Cheu M, et al. Evaluation of two commercial omalizumab/free IgE immunoassays: implications for use during therapy. Curr Med Res Opin 2014;30:913–22.
41. Hamilton RG. Monitoring allergic patients on omalizumab with free and total serum IgE measurements. J Allergy Clin Immunol Pract 2016;4:366–8.
42. Radauer C, Bublin M, Wagner S, et al. Allergens are distributed into few protein families and possess a restricted number of biochemical functions. J Allergy Clin Immunol 2008;121:847–52.
43. Kleine-Tebbe J, Matricardi PM, Hamilton RG. Allergy work-up including component-resolved diagnosis: how to make allergen-specific immunotherapy more specific. Immunol Allergy Clin North Am 2016;36:191–203.
44. European Academy of Asthma Allergy and Clinical Immunology working group. Matricardi P, Kleine-Tebbe J, Ollert M, et al, editors. Handbook on molecular allergology. Pediatr Allergy Immunology May 27, 2016;23:1–250.
45. Thorpe S, Heath A, Fox B, et al. The 3rd international standard for serum IgE: international collaborative study to evaluate a candidate preparation. Clin Chem Lab Med 2014;52:1283–9.
46. Kober A, Perborn H. Quantitation of mouse-human chimeric allergen-specific IgE antibodies with ImmunoCAP technology. J Allergy Clin Immunol 2006;117:S219, [Abstract 845].
47. Merrett J, Merrett TG. Phadiatop—a novel antibody screening test. Clin Allergy 1987;17:409–16.
48. Szefler SJ, Wenzel S, Brown R, et al. Asthma outcomes: biomarker. J Allergy Clin Immunol 2012;129:S9–23.
49. Hiller R, Laffer S, Harwanegg C, et al. Microarrayed allergen molecules: diagnostic gatekeepers for allergy treatment. FASEB J 2002;16:414–16.
50. Martínez-Aranguren R, Lizaso MT, Goikoetxea MJ, et al. Is the determination of specific IgE against components using ISAC 112 a reproducible technique? PLoS ONE 2014;9:e88394.
51. Chapman MD, Wuenschmann S, King E, et al. Technological innovations for high-throughput approaches to in vitro allergy diagnosis. Curr Allergy Asthma Rep 2015;15:36.
52. Hamilton RG, Kleine-Tebbe J. Methods for IgE antibody testing: singleplex and multiplex assays. In: Matricardi P, Kleine-Tebbe J, Ollert M, et al, editors. Handbook on molecular allergology. Pediatr Allergy Immunology May 27, 2016;23:1–250.
53. Hamilton RG. Clinical laboratories worldwide need to report IgE antibody results on clinical specimens as analytical results and not use differential positive thresholds. J Allergy Clin Immunol 2015;136:811–12, 9.
54. Hamilton RG. Responsibility for quality IgE antibody results rests ultimately with the referring physician. Ann Allergy Asthma Immunol 2001;86:353–4.
55. Nordlee JA, Taylor SL, Jones RT, et al. Allergenicity of various peanut products as determined by RAST inhibition. J Allergy Clin Immunol 1981;68:376–82.
56. Charous BL, Schuenemann PJ, Swanson MC. Passive dispersion of latex aeroallergen in a healthcare facility. Ann Allergy Asthma Immunol 2000;85:285–90.
57. Hamilton RG, Wisenauer JA, Golden DB, et al. Selection of *Hymenoptera* venoms for immunotherapy based on patient's IgE antibody cross-reactivity. J Allergy Clin Immunol 1993;92:651.
58. Muller UR, Johansen N, Petersen AB, et al. *Hymenoptera* venom allergy: analysis of double positivity to honey bee and *Vespula* venom by estimation of IgE antibodies to species-specific major allergens Api m1 and Ves v5. Allergy 2009;64:543–8.
59. Ollert M, Blank S. Anaphylaxis to insect venom allergens: role of molecular diagnostics. Curr Allergy Asthma Rep 2015;15:26.
60. Caruso B, Bonadonna P, Bovo C, et al. Wasp venom allergy screening with recombinant allergen testing. Diagnostic performance of rPol d 5 and rVes v 5 for differentiating sensitization to *Vespula* and *Polistes* subspecies. Clin Chim Acta 2016;435:170–3.
61. Savi E, Peveri S, Makri E, et al. Comparing the ability of molecular diagnosis and CAP-inhibition in identifying the really causative venom in patients with positive tests to *Vespula* and *Polistes* species. Clin Mol Allergy 2016;14:3.
62. Sobotka AK, Valentine MD, Ishizaka K, et al. Measurement of IgG-blocking antibodies: development and application of a radioimmunoassay. J Immunol 1976;117:84–90.
63. James LK, Shamji MH, Walker SM, et al. Long-term tolerance after allergen immunotherapy is accompanied by selective persistence of blocking antibodies. J Allergy Clin Immunol 2011;127:509–16.
64. Golden DBK, Lawrence ID, Hamilton RG, et al. Clinical correlation of the venom-specific IgG antibody level during maintenance venom immunotherapy. J Allergy Clin Immunol 1992;90:386–91.
65. Craig CS, Schwartz LB. Tryptase and chymase: markers of distinct types of human mast cells. Immunol Res 1989;8:130.
66. Sprung J, Weingarten TN, Schwartz LB. Presence or absence of elevated acute total serum tryptase by itself is not a definitive marker for an allergic reaction. Anesthesiology 2015;122:713–14.

67. Yunginger JW, Nelson DR, Squillace DL, et al. laboratory investigation of deaths due to anaphylaxis. J Forensic Sci 1991;35:857–65.
68. Miller JS, Schwartz LB. Tryptase levels as an indication of mast cell activation in a patient with Hymenoptera anaphylaxis and mastocytosis. N Engl J Med 1987;316:1622.
69. Ruëff F, Przybilla B, Biló MB, et al. Predictors of severe systemic anaphylactic reactions in patients with Hymenoptera venom allergy: importance of baseline serum tryptase-a study of the EAACI Interest Group on Insect Venom Hypersensitivity. J Allergy Clin Immunol 2009;124:1047.
70. Lichtenstein LM, Osler AG. Studies on the mechanisms of hypersensitivity phenomenon. IX. Histamine release from human leukocytes by ragweed pollen. J Exp Med 1964;120:507.
71. Siraganian RP. Automated histamine analysis for in vitro allergy testing. II. Correlation of skin test results with in vitro whole blood histamine release in 82 patients. J Allergy Clin Immunol 1977;59:214.
72. Ito K, Sato S, Urisu A, et al. An evaluation of spontaneous histamine release and the low responders in a basophil histamine release test. Arerugi 2016;65:48–56.
73. Wegner F, Hockamp R, Rutschke A, et al. Superiority of the histamine release test above case history, prick test and radioallergosorbent test in predicting bronchial reactivity to the house dust mite in asthmatic children. Klin Wochenschr 1983;61:43.
74. de Weck AL. Cellular allergen stimulation test (CAST): a new dimension in allergy diagnostics. ACI News 1993;1:9–14.
75. Maly FE, Marti-Wyss S, Blumber S, et al. Mononuclear blood cell sulpholeukotriene generation in the presence of interleukin 3 and whole blood histamine release in honeybee and yellow jacket venom allergy. J Invest Allergy Clin Immunol 1997;7:217–24.
76. Moneret-Vautrin DA, Sainte-Laudy J, Kanny G, et al. Human basophil activation as measured by CD63 expression and LTC4 release in IgE-mediated food allergy. Ann Allergy Asthma Immunol 1999;82:33.
77. Bochner BS, Sterbinsky SA, Saini SA, et al. Studies of cell adhesion and flow cytometric analyses of degranulation, surface phenotype and viability using human eosinophils, basophils and mast cells. Methods 1997;13:61–8.
78. Sanz ML, Gamboa PM, Antepara I. Flow cytometric basophil activation test by detection of CD63 expression in patients with immediate-type reactions to beta lactam antibiotics. Clin Exp Allergy 2002;32:277–86.
79. Mangodt EA, Van Gasse AL, Decuyper I, et al. In vitro diagnosis of immediate drug hypersensitivity: should we go with the flow? [Review]. Int Arch Allergy Immunol 2015;168:3–12.
80. Buhring HJ, Sieffert M, Giesert C. The basophil activation marker defined by antigen 97A6 is identical to ectonucleotide pyrophosphate/phosphodiesterase 3. Blood 2001;97:3303–5.
81. Platz I, Binder M, Marxer A, et al. Hymenoptera venom induced up regulation of basophil activation marker ecto-nucleotide pyrophosphatase/phosphoesterase 3 (E-NNP3, CD203c) in sensitized individuals. Int Arch Allergy Immunol 2001;126:335–42.
82. Ebo DG, Hagendorens MM, Bridts CH, et al. In vitro allergy diagnosis: should we follow the flow? [Review]. Clin Exp Allergy 2004;34:332–9.
83. Hoffmann HJ, Santos AF, Mayorga C, et al. The clinical utility of basophil activation testing in diagnosis and monitoring of allergic disease. Allergy 2015;70:1393–405.

MULTIPLE-CHOICE QUESTIONS

1. In the diagnostic workup of a patient with a clinical history of a systemic anaphylactic reaction involving hypotension following a sting from a *Hymenoptera*, total and mature mast cell tryptase were requested to confirm mast cell involvement. What is the correct time to collect a whole blood specimen for a total and mature tryptase analysis?
   A. Immediately following reaction onset
   B. 0.5–4 hours following the reaction onset
   C. 4–12 hours following the reaction onset
   D. 15–24 hours following the reaction onset
   E. 24 hours following reaction onset

2. In the workup of a patient for a food allergy, several procedures are typically involved. Which is the optimal sequence of events for these procedures?
   A. Oral provocation tests are first performed, followed by the clinical history and then *in vivo* and/or *in vitro* tests for immunoglobulin E (IgE) antibody (sensitization).
   B. *In vitro* tests for sensitization are performed first, followed by collection of a clinical history and then oral provocation testing.
   C. Clinical history, followed by either *in vivo* or *in vitro* tests for sensitization and then an oral provocation test, if needed.
   D. Skin testing is performed first, followed by serology blood testing and then collection of the patient's clinical history; oral food challenge is considered the last resort.
   E. Performing a multiplex chip–based IgE antibody assay for identifying sensitization before collecting a clinical history from the patient.

3. Allergenic components are increasingly available to dissect the patient's IgE antibody sensitization profile. Exposure to which of the following principal allergen families poses the most serious risk for the induction of a systemic allergic reaction in an individual with peanut allergy?
   A. Lipocalins, which bind steroids, retinoids, and lipids
   B. Polcalcin, which is associated with calcium regulation
   C. Storage proteins (2S albumins, 7S globulins [vicilins], and 11S globulins [legumins]), which are involved in nutrient management
   D. Pathogenesis related protein 10 family (PR10), which are plant-related defense proteins
   E. Tropomyosins, which are integral components of actin filaments

96

Molecular Methods

John W. Belmont

Molecular analysis of genes and genomes can be central to arriving at a precision diagnosis. This chapter reviews the basic principles that underlie clinical molecular genetic testing, assesses representative standard methods that are widely employed, describes new DNA sequencing methods that are being rapidly introduced into clinical diagnostic laboratories, and suggests a multidisciplinary approach for implementation in immunological disorders.

BASIC PRINCIPLES

Genome Structure and Gene Expression

KEY CONCEPTS

Human Genomics

- The human genome encompasses approximately 20 000 protein-coding genes, but each cell expresses only a subset of those genes.
- Genetic and physical maps of the genome are essential to molecular diagnosis of immune system diseases.
- Genetic maps depend on the coinheritance of DNA segments—linkage—to associate DNA variants with disease.
- Physical maps of the genome describe the exact gene locations on a chromosome. The genome DNA sequence is the finest scale physical map of the genome.

The human genome is thought to have about 20 000 protein-coding genes distributed on 23 pairs of chromosomes (http://www.ncbi.nlm.nih.gov/entrez/). The total length of one copy of the genome is ≈3×10^9 nucleotide bases. The protein-coding segments (exons) are split by noncoding DNA sequences (introns). The aggregate protein coding sequences, referred to as the "exome," account for about 1–1.5% of the genome. Some of the remaining DNA contains regulatory elements that direct the expression of genes, control chromatin conformation, encode regulatory RNAs, act as origins for DNA replication, and participate in three-dimensional (3D) looping to produce the large-scale chromosome structure. About 40% of the total DNA is accounted for by families of repeated sequences. These repeat elements are generally silent, but they may be involved in some types of gene regulation and can become involved in mutational mechanisms of deletion, duplication, and insertion. Each cell expresses only a subset of the entire gene repertoire. "Housekeeping" genes are expressed in almost all tissues and cell types, where they perform basic metabolic and structural functions. Other genes are under very specific control, and their expression is restricted to one or a few cell types. Differential gene expression specifies the unique functions of cells (*e.g.,* immunoglobulin [Ig] in B cells, and T-cell receptor [TCR] in T cells). Genes that coordinate the expression of groups of tissue-specific genes primarily encode transcription factors that regulate the rate of messenger RNA (mRNA) transcription on their target genes. A few of these apparently act as "master" genes during particular developmental processes or in specific cell lineages. Some of these master genes (*e.g., PAX5*), may be involved in leukemias and lymphomas. So far, immunodeficiency genes appear to be in the category of genes that involve either innate or adaptive immunity controlling cell growth, differentiation, effector functions, or apoptosis.[1,2]

Polymorphic Variation and Linkage

A genetic map relates one gene to another based on how often they are inherited together. Within a specific region of the DNA, the maternal and paternal copies of the genome may be nonidentical, and the variants are called *alleles.* Protein and nucleotide differences are called *polymorphisms* when they are frequent enough to be found in 1–5% of the general population. Variations in single nucleotides occur in ≈1/100 bases when whole-genome sequencing is used to survey individuals.[3] Polymorphisms arise over time in a group of individuals because of mutations in the DNA (Fig. 96.1). Some polymorphisms involve simple sequence repeats so that there is variation in the number of repeat units. These are called *short tandem repeats* (STRs) and are widely used in forensic identification. In contrast to STRs, the most common polymorphisms involve a change in a single nucleotide position. Large databases of single nucleotide polymorphisms (SNPs, pronounced "snips") have been accumulated (http://www.ncbi.nlm.nih.gov/SNP/). It is thought that there could be >30×10^6 SNPs with minor allele frequency >5% available across the human genome. Many SNPs are found in noncoding DNA, and these have been shown to play a significant, albeit incompletely understood, role in common human diseases.

The composition of alleles within genetic loci in an individual is called the *genotype.* The individual's genotype interacts with the environment throughout life to create the *phenotype.* Some components of the phenotype, such as body weight, are simple to measure, whereas other clinically important phenotypes are based on complex laboratory evaluation (*e.g.,* in T-cell proliferation). A key distinction is drawn between discrete traits (normal vs abnormal) and quantitative traits (continuous range of values). Polymorphisms account for some of the variation observed at the phenotype level between healthy individuals or between populations, and the cumulative percentage of the variation explained by genetic variation is called *heritability.* In monogenic diseases (also called *mendelian diseases*), the presence of mutation(s) is usually considered necessary and sufficient to cause disease. The landscape of genetic variation in each individual

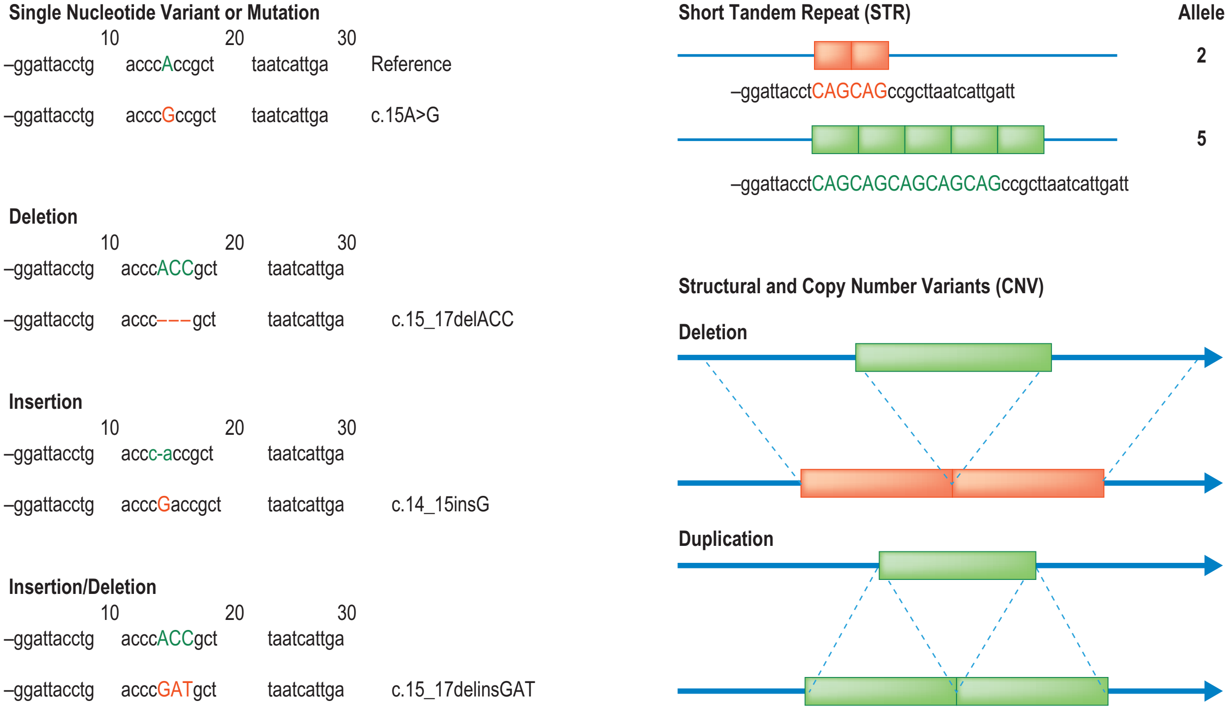

Single Nucleotide Variant or Mutation
10 20 30
–ggattacctg acccAccgct taatcattga Reference
–ggattacctg acccGccgct taatcattga c.15A>G
Deletion
10 20 30
–ggattacctg acccACCgct taatcattga
–ggattacctg accc–––gct taatcattga c.15_17delACC
Insertion
10 20 30
–ggattacctg accc-accgct taatcattga
–ggattacctg acccGaccgct taatcattga c.14_15insG
Insertion/Deletion
10 20 30
–ggattacctg acccACCgct taatcattga
–ggattacctg acccGATgct taatcattga c.15_17delinsGAT
Short Tandem Repeat (STR)
Allele
2
–ggattacctCAGCAGccgcttaatcattgatt
5
–ggattacctCAGCAGCAGCAGCAGccgcttaatcattgatt
Structural and Copy Number Variants (CNV)
Deletion
Duplication

FIG 96.1 Classes of DNA Variation Important for Genetic Testing and Human Genetic Diseases. DNA variations can be either benign or pathogenic, depending on whether they affect the underlying functions of a genetic locus and the encoded proteins. An international standard nomenclature is used to describe changes in the DNA. Structural variants have been found to be a common cause of human genetic disorders.

includes a continuum of gene effects ranging from weak (common SNPs detected in disease association studies) to strongly determinative (rare mutations detected in single gene disorders).

Genes can be mapped in relation to one another by a *linkage* map. Genes that are physically close can be shown to be predominantly coinherited. Linkage mapping played a large role in the identification of the X-linked immunodeficiency genes (Chapters 22, 33–36). Beyond small families and at the population level, the correlation of alleles among markers that are very close to each other in the genome is called *linkage disequilibrium* (LD). LD occurs because the mutation that creates a polymorphism occurs on a single chromosome with its whole complement of unique variants. This initial arrangement of alleles, however, is broken up by recombination over time. Hence, only markers that are relatively close to each other continue to have significant LD. A large international project (HapMap) created a dense map of SNP markers allowing a comprehensive view of LD in several reference populations.[4] This information has been efficiently exploited in genetic epidemiology projects called *genome-wide association studies* (GWAS; www.gwascentral.org), which have characterized several thousand common variants that contribute to many common immunological diseases, such as diabetes, rheumatoid arthritis (RA), and systemic lupus erythematosus (SLE).

The advent of projects that assess whole exomes and genomes in normal and disease populations, such as the 1000 Genomes Project (www.1000genomes.org/), has focused much more attention on rare genetic variation, particularly alleles with frequency of 0.01–1.0%. It is also apparent that the extent of individual genetic variation was underestimated from previous data and that each individual bears about 3.5 million simple nucleotide variants and about 1000 structural variants, many of which are unique to that person. Even given some technical reservations, the enormous extent of private variation has been clearly established.

Linkage and LD mapping have been much less useful for the identification of autosomal recessive genes responsible for many primary immunodeficiencies (PIDs) because the diseases are rare in the population and average families are not large enough to narrowly localize the causative genes. Whole-genome and exome sequencing technologies are being used to solve a great fraction of remaining single-gene disorders, including PIDs.[5,6] The biggest difficulty with these conditions is that similar clinical diseases can be caused by mutation in more than one gene locus (*locus heterogeneity*). Another challenge is that it may be difficult to ascertain multiple extremely rare families with mutations in the same gene. Criteria for rare disease gene identification have been proposed and have helped harmonize standards for future genome interpretation.[7] Successful identification of pathogenic variants, even those unique to single pedigrees, can be accomplished by combining DNA sequencing with complimentary functional technologies.

Physical Maps and DNA Copy Number Variation

Physical maps are different from the genetic map in that they describe how genes are arranged in the DNA on a scale as large as a whole chromosome and as fine as a single nucleotide. At the coarsest level of the physical map, genes are placed in chromosome segments corresponding to the Giemsa-stain banding pattern of metaphase chromosomes. A way to localize genes is to hybridize region-specific DNA probes directly to a chromosome spread. The technique, called *fluorescence in situ hybridization* (FISH), allows the detection of signals from the chromosome spread to directly localize genes.[8] FISH is used for cytogenetic examination of tumors, leukemias, and lymphomas but now has been largely replaced by array *comparative genome hybridization* (CGH), which examines many hundreds of thousands of positions using oligonucleotide probes.[9] High-density oligonucleotide arrays allow for both standardization of reagents between diagnostic laboratories and extensive customization for particular disease applications. These methods are detailed below. Pathogenic *copy number variants* (CNVs) are a major cause of birth defects and intellectual disability,[10] so evaluation for CNVs should be considered a first-line test for any infant or child with a complex presentation of immune deficiency. In DiGeorge syndrome (DGS), the characteristic deletion of chromosome 22q11 can be detected by microarray. FISH should not be used because microarray not only provides high-resolution determination of the deletion size but also leaves open the possibility of detecting other unsuspected DNA copy number abnormalities.

Mutation and Pathogenic Variants

KEY CONCEPTS

Mutation

- Mutation can occur by deletion, insertion, or duplication of short or long DNA segments.
- Single-base mutations can be caused by replication errors or by chemical deamination of methylcytosine.
- Single-base mutations can affect protein coding sequence, regulatory sequences, or the RNA splice signals within a gene.
- The parental origin of mutations (maternal or paternal) affects those mutational mechanisms most likely responsible.
- A disease can be caused by mutations in several alternative genes (locus heterogeneity).
- Most primary immunodeficiencies occur because of different kinds of mutations in particular genes (allelic heterogeneity).

DNA variants can involve single nucleotide substitutions, small or large deletions, insertions, inversions, duplications, or repeat expansions.[11] Some variants have no measurable or functionally significant phenotypic effect. They become one of the huge pool of neutral polymorphisms in the genome. Single-base substitutions in a triplet amino acid codon often disturb the normal function of proteins. These are described as *missense* (*i.e.,* causing an amino acid substitution) or *nonsense* (*i.e.,* terminating translation) mutations. Mutations affecting regulatory or splice signal sequences in RNA can also be deleterious. The primary mechanisms of mutation are misincorporation of nucleotides and faulty repair of chemically damaged nucleotides (*e.g.,* 8-oxodG) during DNA replication by DNA polymerases. Another important mechanism underlying single base or "point" mutations occurs at CpG dinucleotides. The cytosine is often methylated, and chemical deamination of the methylated cytosine gives the base thymidine. At the next round of replication the CG may be changed to TG or CA, depending on which DNA strand was altered. Small insertion and deletion (*indel*) mutations are also very common. Indels occurring in protein coding sequence are usually deleterious, as they can cause frameshifts (changing the reading frame, leading to translation termination within a few codons) and account for almost 25% of known human disease causing mutations.[12] They result from strand slippage in repetitive sequences and misaligned intermediates during DNA synthesis. An entirely different mechanism for mutation is expansion of short repeat segments (most often trinucleotides).[13] So far, no disorder that affects the immune system has been associated with this mechanism.

Heritable mutations can arise in either the male or the female germline, but the gender-related frequency is influenced by the mutational mechanism. Chromosomal nondisjunction, for example, occurs predominantly during female meiosis and has a strong maternal age effect.[14,15] The fact that spermatogonial proliferation occurs throughout life and involves many more cell replications than oogenesis increases the single base mutation frequency in the paternal germline.[16,17] Clinicians must cope with both locus and allele heterogeneity. *Locus heterogeneity* means that the same or a similar phenotype results from mutation in one of several different genes. For example, severe combined immunodeficiency (SCID) can result from mutation in the adenosine deaminase, interleukin-2 (IL-2) receptor γ-chain genes, *JAK3,* and so on. *Allelic heterogeneity* means that the disease is caused by different mutations in the same gene. In X-linked disorders, allelic heterogeneity is typically high because affected individuals have reduced reproduction (negative evolutionary selection), and most mutations are lost in the population after a few generations. At autosomal loci some mutations have reached appreciable frequency as a result of demographic processes (*e.g., founder effect*). In the heterozygous state, recessive mutations can be weakly deleterious, neutral, or actually confer a small advantage. Because the mutant alleles are common, it is possible to test affected individuals and potential carriers directly for those mutations. This is the basis for population screening for carrier status in cystic fibrosis. However, none of the immune deficiencies results from common mutations that would permit efficient screening. In the autosomal recessive immune deficiencies, most patients have two different, very rare mutations (*compound heterozygosity*) at the disease locus. If there is known parental consanguinity or the family comes from an isolated population, the affected individual may be homozygous for a rare mutation.

X-Chromosome Inactivation

Disturbances in the pattern of X inactivation are interesting phenomena in women who are carriers of X-linked disorders, including several immune deficiencies.[18] Measurement of X-chromosome activity in the blood cells of normal women shows that, on average, the contribution of maternal and paternal X chromosomes is approximately 50:50. Extreme skewing of X inactivation often results from abnormal proliferation of blood cells. X-inactivation analysis can be used to demonstrate clonal growth of premalignant (*e.g.,* myelodysplasia) and malignant cells. In a female carrying mutations at the X-linked severe combined immune deficiency (*XSCID*; OMIM 300400), agammaglobulinemia (*XLA*; OMIM 300300), or Wiskott-Aldrich syndrome (*WAS*; OMIM 301000) loci, cell competition and compensation mechanisms lead to a reduced contribution of cells expressing the mutant allele among affected cells. In XLA, the B-cell lineage shows selective use of the nonmutant active X chromosome. In XSCID, skewing of X inactivation is observed in B cells, T cells, and natural killer (NK) cells. In WAS, some degree of skewing of X inactivation results from defective hematopoietic stem cell (HSC) activity. Historically, skewed X inactivation was very helpful in linkage mapping of these conditions. As a clinical test, X inactivation now has reduced importance compared with direct mutation analyses.

DESCRIPTION AND PRINCIPLES OF DNA DIAGNOSTIC TECHNIQUES

The overriding theme of all DNA diagnostic techniques is the detection of variants in the DNA molecule that are associated with disease. This has historically involved a targeted search in one or a few specific genes or regions, but the field is evolving rapidly toward techniques that allow for simultaneous survey of the entire genome. DNA copy number analysis and DNA sequencing have come to dominate individual genetic diagnostic methods. Bioinformatics is a necessary diagnostic laboratory discipline that must be coupled to these methods, particularly when the whole genome is surveyed.

Detection of Disease-Causing Copy Number Variants: FISH and Microarrays

Many genetic diseases are caused by deletions or duplications in human DNA. These conditions typically have copy number abnormalities ranging from 500 kilobases (Kb) to 5 megabases (Mb) and are sometimes referred to as “genomic disorders.” The most common genomic disorder with a clinically important effect on the immune system is 22q11 deletion syndrome. This disorder affects about 1 per 4000 live-born infants and accounts for about 90% of DGS cases. The karyotype, in which a metaphase chromosome spread was stained and visualized under the microscope, was for many decades the only clinically useful test for the detection of copy number changes. The karyotype offered a genome-wide view, but at a very low resolution: alterations smaller than 5 Mb, including those that typically cause DGS, were not routinely detected. Two methods, described below, have replaced conventional karyotypes for routine genetic diagnosis. Whole-genome sequencing (WGS) combined with advanced bioinformatics methods promises to transform evaluation of CNVs yet again. In principle, the data collected in WGS can be used to resolve all CNVs regardless of length, thus capturing all forms of structural DNA variation from the smallest to gross abnormalities in chromosomes.

Fluorescence *in situ* Hybridization

FISH has been an important technology for detecting both constitutional and somatic abnormalities in chromosomes.[19] In its simplest form, a cloned gene segment (referred to as a *probe*) is labeled (Fig. 96.2) and allowed to hybridize to a preparation of early metaphase (condensed) chromosomes. The labeled probe hybridizes specifically to a locus and is then detected with secondary antibodies and high-sensitivity immunofluorescence techniques. The targeted locus appears as two pairs of dots corresponding to the sister chromatids from the two chromosomal homologues. Cells that are in interphase (when the DNA is largely unfolded in the nucleus) can also be probed. In the interphase cell, two single signals are observed corresponding to the two chromosomes. With the introduction of microarray-based methods for genome-wide copy number analysis, the practical application of FISH has been narrowed but does have a continuing role in defining balanced translocations and in cancer cytogenetics.

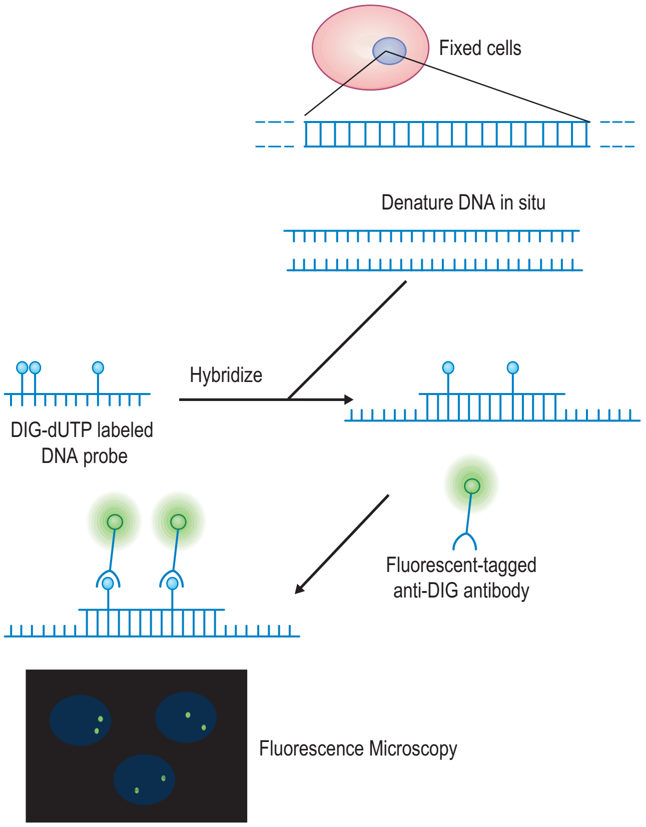


FIG 96.2 Fluorescence *in situ* Hybridization (FISH). Probes complementary to specific sites in the genome are labeled with a hapten (*e.g.,* digoxigenin) and then hybridized to denatured DNA within cells on microscope slides. The hybridized probes are then detected by immunofluorescence microscopy.

KEY CONCEPTS

Molecular Methods

- Most DNA diagnostics rely on DNA copy number analysis or DNA sequencing.
- Polymerase chain reaction (PCR) allows for exponential amplification of a DNA segment from the genome, allowing multiple analytical procedures on the single pure sequence.
- Array-based copy number analysis has become the first-line method for finding small abnormalities in the structure of chromosomes.
- New high-throughput sequencing methods have emerged as practical alternatives to older single analysis approaches.
- Bioinformatics is an emerging discipline relevant to molecular diagnosis. Integration of bioinformatics techniques plays a necessary role in data acquisition and interpretation.

Array-Based Copy Number Analysis

Oligonucleotide arrays have had a revolutionary impact on molecular diagnostics, by offering high-resolution, genome-wide detection of copy number changes. The basic principle involves synthetic oligonucleotides complementary to human nucleic acid sequences bound to a surface or bead, and hybridization of fragmented, labeled human nucleic acid samples (Fig. 96.3). Array printing of activated phosphoramidite nucleotide precursors allows for hundreds to thousands of oligonucleotides to be fabricated on activated glass surfaces (Agilent Technologies, Santa Clara, CA).[20] Light-activated oligonucleotide synthesis has also been used to efficiently customize oligonucleotide arrays (Affymetrix Inc., Santa Clara, CA; and Roche Nimblegen Inc., Madison, WI).[21] A third approach involves synthesis of oligonucleotides on microbeads that are randomly flowed over a patterned surface

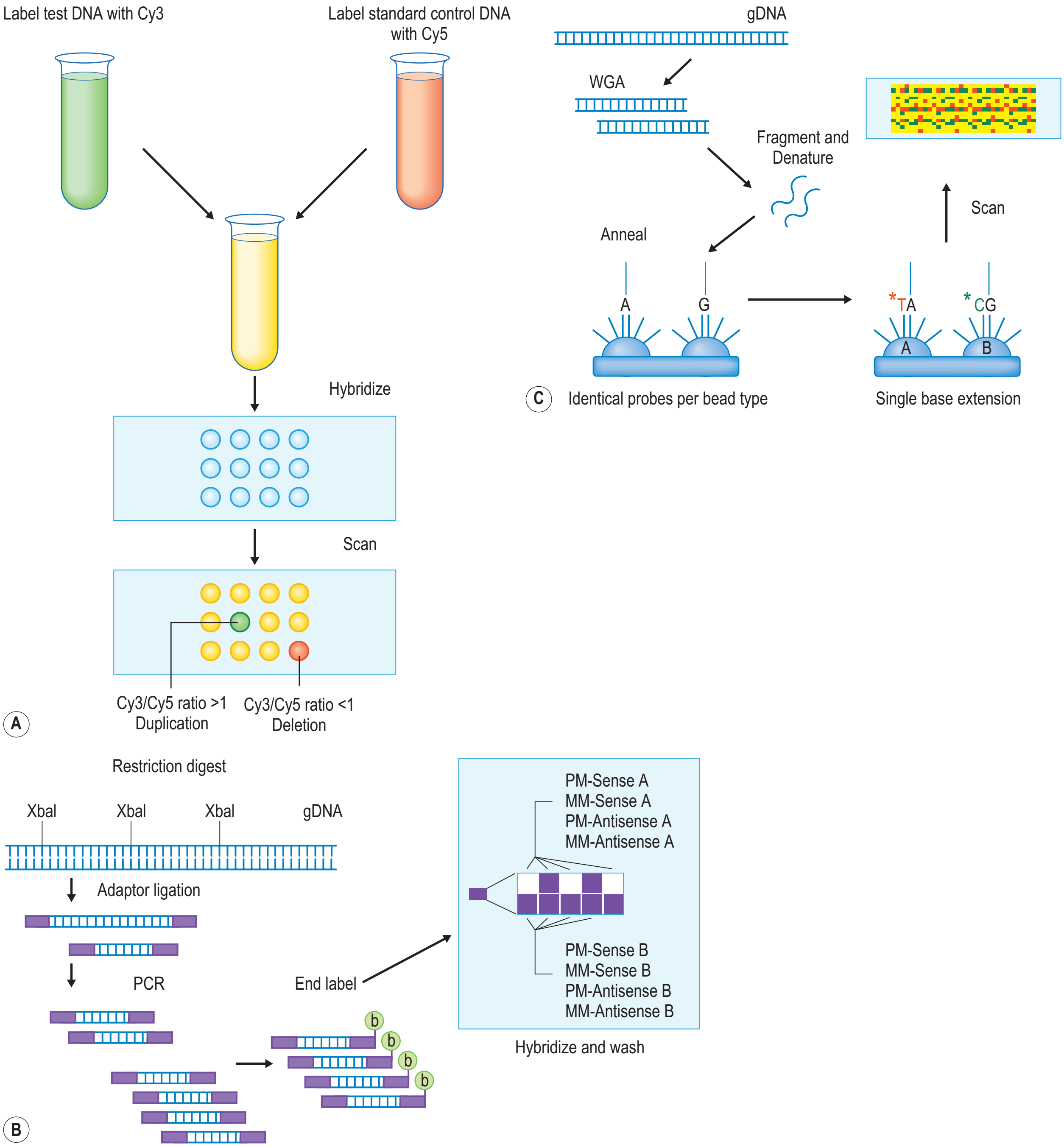

Label test DNA with Cy3
Label standard control DNA with Cy5
Hybridize
Scan
Cy3/Cy5 ratio >1
Duplication
Cy3/Cy5 ratio <1
Deletion
A
gDNA
WGA
Fragment and Denature
Anneal
A
G
*TA
*CG
A
B
Scan
C
Identical probes per bead type
Single base extension
Restriction digest
XbaI
XbaI
XbaI
gDNA
Adaptor ligation
PCR
End label
b
PM-Sense A
MM-Sense A
PM-Antisense A
MM-Antisense A
PM-Sense B
MM-Sense B
PM-Antisense B
MM-Antisense B
Hybridize and wash
B

FIG 96.3 Array-Based DNA Copy Number Analysis. (A) Array comparative genomic hybridization (aCGH) uses a reference DNA to compare with the test DNA. The labeled DNA is hybridized to an array of oligonucleotides that are designed to be uniquely complementary to hundreds of thousands of positions in the genome. The relative intensities of the fluorescence signal from each DNA probe allow assessment of copy number. Genotyping arrays are also widely used for copy number testing (B, C). (B) Whole genome sampling and amplification (WGSA). End-labeled polymerase chain reaction (PCR) products that constitute a sample from the genome are hybridized to oligo sets on standard chips. The oligos are organized into "probe quartets" consisting of perfect matches (PMs) and single-base mismatches (MMs) for each allele, A and B. Separate quartets for the sense and antisense strand of DNA help to ensure specificity, and in the actual assay these quartets are repeated three to five times on the chip. This type of assay has been developed to test a standard panel of >2 million DNA features per sample. (C) Single base extension (SBE; Illumina Infinium II). This assay uses whole genome amplification (WGA) to increase the molar concentration of target DNA. SBE on arrays of tagged microbeads allows allele discrimination, and the resulting products are detected by immunostaining (signal amplification). This assay has been used for assays of up to 5 million single nucleotide polymorphisms (SNPs) per sample.

(Illumina Inc., San Diego, CA).[22,23] These processes are very similar, in principle, to the fabrication of microelectronics devices, and so have given these arrays the nickname "gene chips."

Microarrays have been used in very-large-scale studies to genotype common polymorphisms. Robust chemistries that can analyze standard sets of up to 5 000 000 SNPs in a single assay are now routine (see Fig. 96.3). Nonamplified genomic DNA is too complex (*i.e.,* the concentration of any particular sequence is too low) to be analyzed directly. High-throughput genotyping assays require an array-based readout and scalable, multiplex assay chemistry. This is extremely challenging because of the complexity of the human genome (*i.e.,* any particular unique sequence in the mixture has a low molar concentration and because of the need for exquisite specificity at the single base level). An important general principle for achieving single-base specificity is the detection of physically coincident events.[24] This is exactly the route taken by the polymerase chain reaction (PCR), in which specific annealing of both primers is required for the reaction to take place (Fig. 96.4).

Whole genome sampling and amplification (WGSA) implemented in the Affymetrix SNP arrays uses parallel amplification of short DNA segments (200–1100 base pairs [bp]) that are then labeled and hybridized to oligonucleotide probe arrays.[25] The oligonucleotides are specific for each allele of SNPs, and the resulting hybridization intensities can be used to derive the genotypes. These assays are fixed in the sense that the investigator uses the information from a standardized array but cannot add SNPs for particular purposes. An alternative approach, incorporated in the Infinium assays (Illumina Inc., San Diego, CA), involves the detection of allele-specific primer extension or single base extension products on a self-assembling high-density microbead array.[26] In the Infinium assays, genomic DNA is first subjected to whole genome amplification (WGA), and those products are hybridized to an array of locus-specific 50-mer capture probes. Either primer extension or ligation reactions on the array surface are used to accomplish allele discrimination. Signal amplification with methods familiar to users of enzyme-linked immunosorbent assay (ELISA) is used to enhance the sensitivity. A distinct advantage of this class of assays is that the "assay conversion" (percent SNPs that actually allow inference of the genotype in the final assay mixture) is very high and thus allows for flexibility in designing tests customized to particular problems.

GENE EXPRESSION

Global gene expression analysis has been enabled by microarray technology giving rise to so called "transcriptomics." For RNA

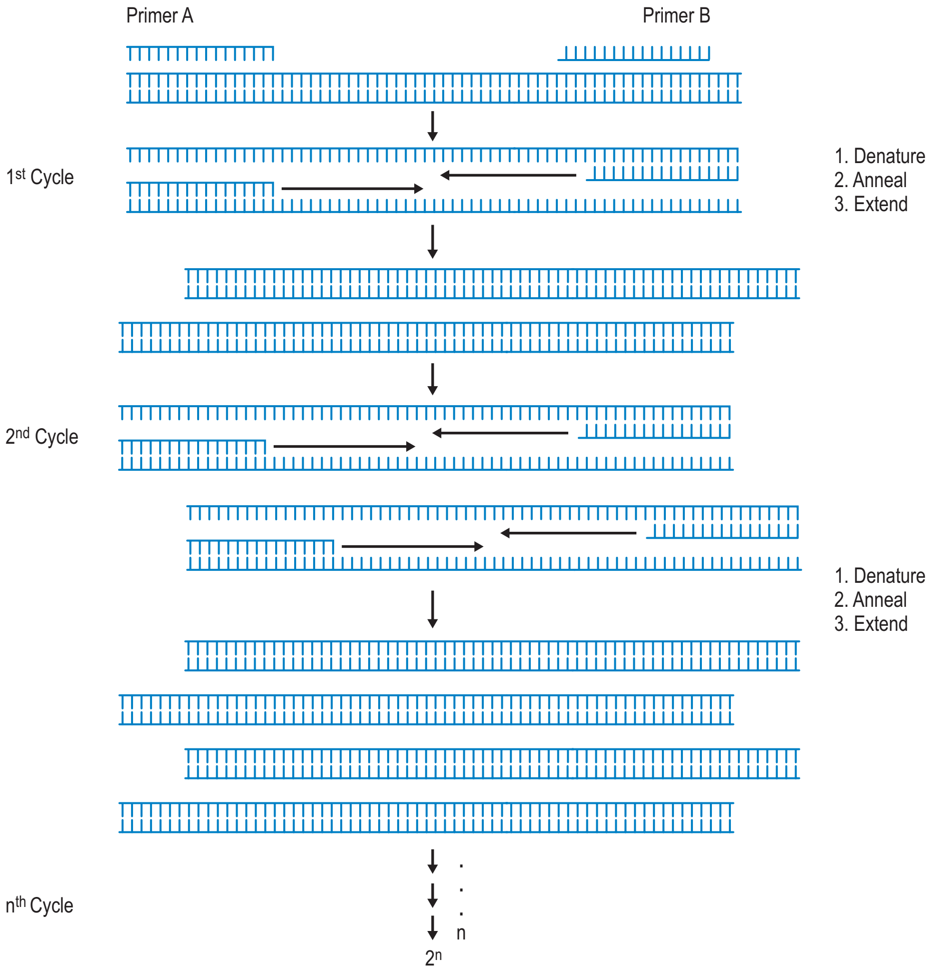


FIG 96.4 Polymerase Chain Reaction. Primers that hybridize to the target sequence are used to initiate multiple cycles of synthesis, melting, annealing, and synthesis. In practice, the potential geometrical increase in DNA tapers off as reaction components become limiting.

expression analysis, cDNA "target" derived from the tissue or cell source is labeled with fluorophores and used to determine complementary binding by massively parallel hybridization to the probes fixed on the array. The ability to massively sequence RNA (see below) from diverse biological samples has reduced use of arrays for research. RNA measurement, either by array or sequencing, has been slow to enter clinical practice. A major challenge lies in the variation between studies in sample selection, cell composition of samples, analysis platform, and statistical methods. Establishing criteria for analytical validity, clinical validity, and clinical utility are long-term challenges for use of gene expression biomarkers.

Detection of Disease Causing Mutations—Point Mutations, Insertions/Deletions, and Structural Variants

Methods for the detection of single-base substitutions in DNA are central to clinical genetic diagnostics.[27] Although complementary DNA (cDNA) synthesized from mRNA can be used, analysis of genomic DNA is far more common. Many different methods for mutation detection have been described, but virtually all clinical diagnostic laboratories now approach an unknown mutation with DNA sequencing. When Sanger sequencing is used, the first step is to design PCR primer sets that will allow amplification of each exon. The availability of the reference human genome sequence makes determination of gene structure and amplimer design for genomic DNA straightforward. Any mutation that directly affects amplification, such as deletion or mutation in the amplimer sequence, can produce misleading results. It is always desirable to have parental specimens available to serve as a reference for proving biallelic inheritance in the affected offspring. If there is a clear clinical diagnosis and only a single gene likely to be involved, then automated fluorescent sequencing is still the preferred method. For disorders with highly complex clinical and laboratory phenotypes, like primary immune deficiencies, the problem of locus heterogeneity is very severe. It is now routine practice to solve this problem by sequencing the complete exome or the complete genome in individual patients by using the next-generation technologies described below.

Sanger Sequencing

The direct determination of a DNA sequence is fundamental to mutation identification (Fig. 96.5). The oldest method still in common use (called *Sanger sequencing*) depends on termination of DNA synthesis by chemically modified nucleotides (dideoxyNTPs [ddNTPs]). A single-stranded DNA (ssDNA) template is hybridized to a primer that can recognize a short known segment of DNA. The ssDNA is usually produced by DNA amplification. The primer acts as a starting point for DNA synthesis by a DNA polymerase that is added to the reaction. Four separate reactions are set up, each "spiked" with one of the four possible ddNTPs (ddATP, ddTTP, ddCTP, or ddGTP). As synthesis proceeds, some of the strands incorporate a ddNTP, and no further extension can take place. In each reaction, a family of molecules is synthesized; their unit lengths are determined by whether the reaction had been terminated by incorporation of a specific ddNTP at a given position. Either the primer or the nucleotides may be labeled. Automated instrumentation for DNA sequencing using laser scanning of fluorophore-labeled reactions is standard. By using four different fluorescent labels the reaction products from one sample can be analyzed together, with many samples processed in parallel using renewable capillary electrophoresis arrays.

Next-Generation DNA Sequencing

New technologies that have dramatically increased the throughput and reduced the cost of DNA sequencing have been rapidly adopted in research and diagnostic laboratories. Unlike Sanger sequencing, which produces sequence data by controlled termination of the polymerase, the new techniques derive the sequence as nucleotides are sequentially added by the polymerase or interrogate the sequence by flow through nanopores. Although multiple chemistries and detection systems have been developed into sequencing instruments (Fig. 96.6), this discussion will review methods that have been adopted into clinically validated assays and also briefly describe two methods that, to date, have only been used in research but that illustrate the potential for very long reads. Such methods may have a role in better defining

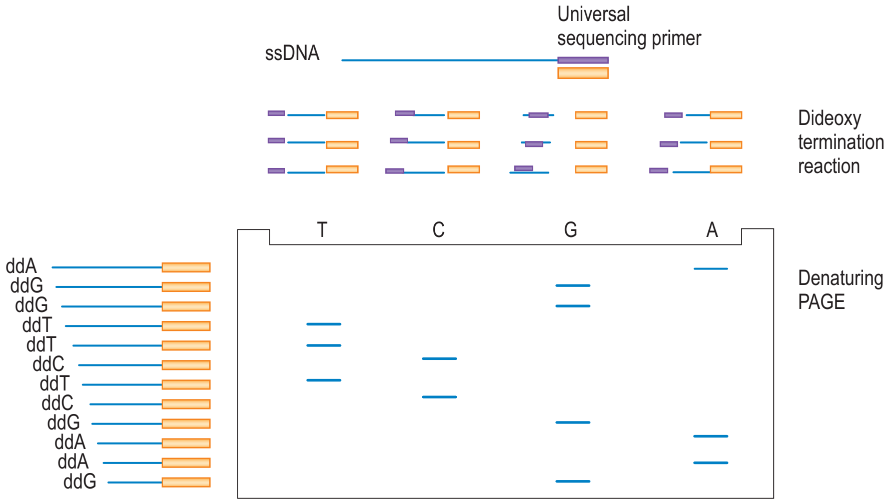


FIG 96.5 DNA Sequencing by Dideoxynucleotide (ddNTP) Chain Termination (Sanger). Copying of the DNA by the polymerase is terminated at specific positions when a ddNTP is incorporated. The ddNTP is mixed with deoxynucleotides (dNTPs) so that in each reaction, only some new strands terminate, whereas others continue through to the next complementary nucleotide. The sequencing products can be visualized by autoradiography or by laser scanning in an automated sequencer.

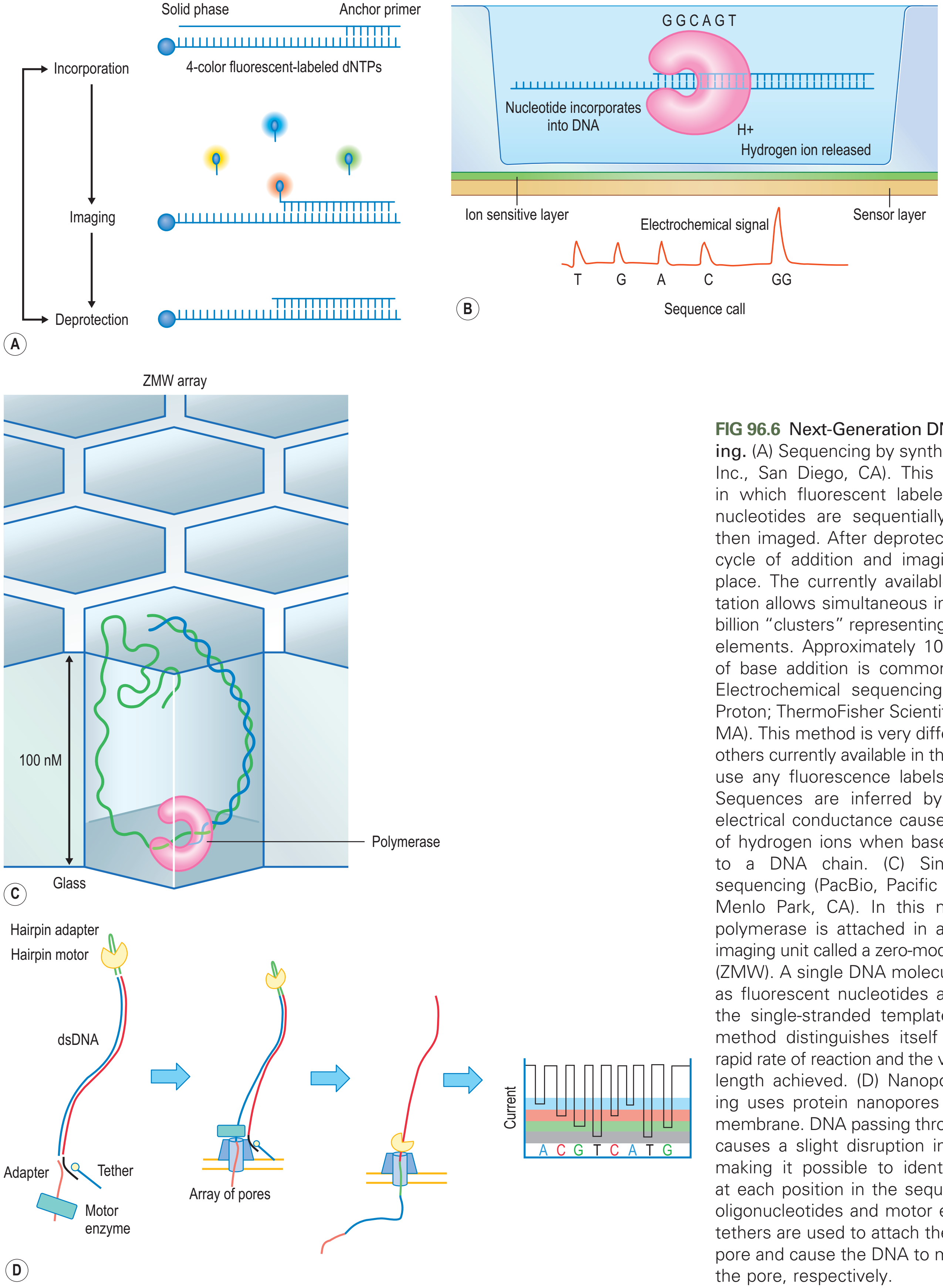

Solid phase
Anchor primer
Incorporation
4-color fluorescent-labeled dNTPs
Imaging
Deprotection
A
G G C A G T
Nucleotide incorporates into DNA
H+
Hydrogen ion released
Ion sensitive layer
Electrochemical signal
Sensor layer
T
G
A
C
GG
Sequence call
B
ZMW array
100 nM
Polymerase
Glass
C
Hairpin adapter
Hairpin motor
dsDNA
Adapter
Tether
Motor enzyme
Array of pores
Current
A C G T C A T G
D

FIG 96.6 **Next-Generation DNA Sequencing.** (A) Sequencing by synthesis (Illumina Inc., San Diego, CA). This is a method in which fluorescent labeled terminator nucleotides are sequentially added and then imaged. After deprotection, another cycle of addition and imaging can take place. The currently available instrumentation allows simultaneous imaging of >1 billion "clusters" representing unique DNA elements. Approximately 100-150 cycles of base addition is commonly used. (B) Electrochemical sequencing (IonTorrent/Proton; ThermoFisher Scientific, Waltham, MA). This method is very different from all others currently available in that it does not use any fluorescence labels or imaging. Sequences are inferred by changes in electrical conductance caused by release of hydrogen ions when bases are added to a DNA chain. (C) Single-molecule sequencing (PacBio, Pacific Biosciences, Menlo Park, CA). In this method DNA polymerase is attached in a microscope imaging unit called a zero-mode waveguide (ZMW). A single DNA molecule is imaged as fluorescent nucleotides are added on the single-stranded template DNA. This method distinguishes itself by the very rapid rate of reaction and the very long read length achieved. (D) Nanopore sequencing uses protein nanopores arrayed in a membrane. DNA passing through the pore causes a slight disruption in the current making it possible to identify the base at each position in the sequence. Tether oligonucleotides and motor enzymes and tethers are used to attach the DNA to the pore and cause the DNA to move through the pore, respectively.

structural variants in DNA. In addition, an important recent application of next-generation sequencing (NGS) will be discussed, circulating cell-free DNA (cfDNA). This technology has already revolutionized pregnancy screening for common chromosomal disorders and may be adapted to screening for low levels of solid tumor DNA in blood.

Sequencing by synthesis (Illumina). The most widely used system for NGS involves highly parallel imaging of single base addition on "clusters" of identical DNA molecules (see Fig. 96.6A). This technology has been commercialized by Illumina Inc. (San Diego, CA). In their approach, DNA is ligated to primers that allow it to be amplified on a surface called a "flow cell."[28] The DNA templates attach to the flow cell surface by hybridization to specific primers complementary to that used to prepare the DNA library. Solid-phase bridge amplification creates many identical copies of each single template molecule in a localized cluster. The density of the clusters is extremely high, and the specialized imaging system allows resolution of more than 10^9 clusters per flow cell. DNA sequence is determined in the flow cell by sequential addition of fluorescent-tagged nucleotides. A single-labeled nucleotide is added to nascent DNA in each sequencing cycle. Each nucleotide has a different fluorophor so that all four are added to the mix in each cycle of base extension. The nucleotide label also terminates polymerization (*i.e.,* only a single base is added at each cycle), and the identity of the incorporated nucleotide is determined by the fluorescent emission spectrum and intensity. Because the bases act as terminators, homopolymeric segments can be resolved relatively accurately. After each deoxynucleotide (dNTP) incorporation, the fluorescent dye is enzymatically cleaved to allow incorporation of the next nucleotide. This class of instruments is capable of sequencing complete human genomes in a single run.

Electrochemical sequencing. Electrochemical sequencing (IonTorrent/Proton; ThermoFisher Scientific, Waltham, MA) is distinctly different from systems that image fluorescent-tagged nucleotides because the sensor is a microelectronic device (see Fig. 96.6A).[29] DNA is first fragmented and then ligated to adapters, and the adaptor-ligated libraries are clonally amplified by emulsion PCR onto beads. Individual beads are then loaded into single-sensor wells. Nucleotides are provided in a stepwise fashion with incorporation increasing the length of the sequencing primer by one base when there is a complementary base on the template strand. Sequence determination relies on primer extension. Nucleotide incorporation into a nascent DNA strand by DNA polymerase results in hydrolysis of the nucleotide triphosphate. Hydrolysis causes production of a hydrogen ion for each nucleotide incorporated. The small shift in the pH of the surrounding solution is proportional to the number of nucleotides incorporated, which is then detected by the sensor on the bottom of each well, converted to a voltage, and digitized by off-chip electronics. The chip is automatically washed, and the cycle is repeated with the next nucleotide. Because of the small size of the wells, diffusion into and out of the well is very fast. The sensor layer is composed of tantalum oxide, which is sensitive to proton concentration (essentially pH) allowing rapid detection of the voltage transients that follow base incorporation in each well individually. The voltage change is roughly proportion to the length of a run of the same nucleotide in the DNA so that short homopolymers can be accurately called. Chip fabrication and design are very similar to methods used for microelectronic devices. This suggests that the method will have good potential to scale up to increase the number of molecules processed in parallel. Lacking optical components, the cost of the all-electronic detection system is far less than other sequencing instruments.

Long read sequencing—single-molecule sequencing and nanopores. A true single-molecule sequencing method has been developed by Pacific Biosciences (Menlo Park, CA).[30] In their method DNA polymerase is immobilized in a femtoliter-sized well with special optical properties called the zero-mode waveguide (ZMW) (see Fig. 96.6C). The small size of the ZMW "hole" prevents 600-nm wavelength laser light from passing entirely through, and only the bottom 30 nm where the polymerase is bound is subject to the fluorescence excitation. Labeled nucleotides are flowed into the chamber and complementary bases encountering the DNA polymerase are incorporated into the growing DNA chain. During incorporation, the DNA polymerase holds the nucleotide for tens of milliseconds, orders of magnitude longer than the average diffusing nucleotide. The transient light emission is then detected and the identity of the incorporated base recorded. Nucleotides with a fluorescent dye attached to the phosphate group of the nucleotide are cleaved when the nucleotide is incorporated into the DNA strand. The label diffuses away, leaving the DNA ready for the addition of the next base. The polymerase incorporates multiple bases per second, so the sequencing process is very fast with real-time observation of DNA synthesis.

Engineered nanopores are being evaluated for the detection of a broad array of biomolecules, including proteins and nucleic acids.[31] There are many potential variations in pore, sample preparation, and detection, which have given rise to active academic and commercial research programs. In one effective version, a protein channel (*e.g., Mycobacterium smegmatis* porin A, MspA) through which DNA may pass is held in a partition (membrane or other solid) that separates solutions containing charged ions. An applied voltage leading to movement of ions through the protein pore produces a measurable electrical current. DNA passing through the channel causes partial reduction of the current with the four DNA bases affecting the current by differing and characteristic amounts (see Fig. 96.6D). The changing electrical current is then used to infer the order of bases in the DNA molecule. Oxford Nanopore Technologies (Oxford, UK) released the first commercial devices based on this technology, and the uses are under active investigation. One of the most important strengths of this method is the very long single-molecule reads that can be obtained. High read depth can partially overcome problems with sequence accuracy (per base error rates of 10–30%). The fast run times and simple protocols for sample preparation have allowed an early stage instrument to be used to sequence Ebola virus on site in Africa. Because of the speed of analysis, "bedside" medical applications could be feasible (especially when base error rate is of less importance for the intended use) and appears especially promising in the arena of precision diagnosis of infectious diseases.

BIOINFORMATICS

Clinical Bioinformatics has emerged as a fundamental discipline in laboratory medicine as enormous DNA sequence, transcriptomics, and variant databases have been aggregated from individual laboratories and international genomics projects. The growing number of genetic loci that are already known to be important in immunological disease renders it critical that diagnostic laboratories make maximal use of automated processes in sample management, data acquisition, data analysis, and reporting. Laboratories using NGS instruments must establish an informatics pipeline for each clinical application. New assay development requires validation conforming to Clinical Laboratory Improvement Amendments of 1988 (CLIA 1988) regulations and laboratories must have an ongoing process to monitor data quality and ensure result accuracy. The US Food and Drug Administration (FDA) has released a draft guidance on standards for clinical NGS-based diagnostic tests (http://www.fda.gov/ucm/groups/fdagov-public/@fdagov-meddev-gen/documents/document/ucm509838.pdf).

Sample and Laboratory Process Management

Each diagnostic laboratory must deal with the generic operational problems of sample accession, tracking, and reporting. Clinical-grade laboratory information management systems (LIMS) are required to handle all these processes with associated regulatory compliance. DNA diagnostic laboratories have several unique problems and requirements that deserve comment. Automated data acquisition is an important component of DNA sequencing and genotyping requiring personnel specialized in information science and systems administration. Advanced statistical models are employed at many steps in the processes of base calling (*primary analysis*), alignment to the reference genome, and identification of positions that are different from the reference (together called *secondary analysis*). Once the raw data from arrays and NGS are produced, bioinformaticians develop, manage, and operate analysis pipelines that synthesize the results into forms comprehensible to the laboratory staff tasked with reporting the results. Bioinformaticians maintain or develop analysis information management systems (AIMS), which are also used to collect and monitor performance metrics and quality control. Specialized software is used to perform these functions and to report the metrics needed for quality control. The number of patient-specific data records and the complexity of relationships in family-based testing make it essentially impossible for manual processes to achieve the required reliability. Because of the broad intended use of genomic testing, there is an increasing importance to collection of patient phenotype data, which is needed for variant filtering and prioritization (see *Tertiary analysis* below).

Primary Data Analysis—Genotyping and Base Calling

Genotyping in the case of microarray and base calling in the case of sequencers are platform specific, and the required software is supplied by instrument vendors. Microarray data, whether array comparative genomic hybridization or SNP chip platforms, use signal hybridization intensity to estimate DNA copy number. Copy number calls are based on multiple adjacent assay positions with respect to the genome map (*i.e.,* the identification of clinically important CNVs is always supported by many independent data points and on-chip assay replicates). The resolution and reliability of the CNV calls depends on total number of positions on the array and their "responsiveness" to differences in DNA copy number. Laboratories using these methods must assess data quality with robust statistical procedures specifying in advance the minimum size and composition of called CNVs.

Sequencing data, especially when considering exome and whole genomes, presents much more challenging problems in bioinformatics. Base calling from the instrument raw data (primary analysis) typically takes place in local computers dedicated to the sequencer, but cloud-based methods can be used. Base calling generates sequence "reads" with their base quality scores. Some of the important measures of data quality at the primary analysis step are base quality score, number of reads per sample, length of reads, and fall-off of base quality with read position.

Secondary Data Analysis—Demultiplexing, Alignment, and Variant Calling

The next steps in next-generation sequence data analysis involve aligning reads to a reference sequence and generating variant calls. In many high-throughput applications, patient DNA samples are tagged with index sequences during preparation for sequencing (*library construction*). Molecular indexing or multiplexing allows pooling of the samples on the instrument and then sorting them out after sequencing. Demultiplexing of sequence reads is another step that is subject to quality monitoring.

After demultiplexing, the reads are mapped and aligned to the reference genome. Alignment of short-read sequences to the reference genome involves systematically matching read fragments to their correct location in the genome. The most widely used tools exploit the Burrow-Wheeler algorithm to carry out this process efficiently and precisely (bio-bwa.sourceforge.net/). Typically, only uniquely mapping reads are passed to the later steps of sequence analysis. This makes it difficult to analyze some segments of the genome that are important in health and disease. Some elements in the genome are composed of nearly identical sequences most often arranged in tandem on adjacent segments of chromosomes. Human leukocyte antigen (HLA) presents particular challenges: *(i)* certain HLA alleles may not be represented in the reference genomes; *(ii)* reads may align to more than one location in HLA leading to discard of the read or misalignment and false-positive variation; and *(iii)* identical reads may have origins in distinct haplotypes that cannot be easily recognized with short-read sequences.

Another important issue is that the reliability of variant calling is different with different classes of variation. Small insertion and deletion variants (indels) are clinically important because they often lead to frameshift and premature termination of proteins; but calling indels and automated application of consistent indel nomenclature are more difficult than single nucleotide variants (SNVs). The Genome Analysis Toolkit (http://www.broadinstitute.org/gatk/) is the most widely used software for variant calling.

Targeted resequencing and whole exome sequencing (WES) focus on protein-coding elements in the genome. Because of the complex and highly variable exon–intron structure of genes, there is considerable technical difficulty in using exon sequence data to call structural variants and CNVs. WGS, in contrast, surveys all the intron and intergenic sequence. New methods of PCR-free library construction enable the read count depth to be used as an accurate surrogate for copy number.[32] In addition, gaps in aligned reads (called "split reads") can be used to recognize deletions and other structural variants, including duplications and inversions. Although challenges still remain, it is possible that WGS data combined with standardized algorithms could allow a single test to be used for almost all classes of pathogenic alleles.

Variants are saved in a specified format called a *genomic variant call format (gVCF) file.* This format contains information not only about the positions that contain a nonreference genotype call but also about the quality of each site that is called with the reference homozygous base. This is important because it allows multiple samples to be aggregated (*e.g.,* to analyze mother, father, and their offspring jointly). Format standardization allows exchange and aggregation of data among laboratories around the world. Data aggregation is now widely recognized as a key step in the development of molecular diagnostics, not only to reduce errors in variant calling but also to enable sophisticated approaches to the problem of genotype-phenotype relationships in rare genetic diseases.

Tertiary Data Analysis—Variant Annotation, Interpretation, and Reporting

The next steps in the analysis involve annotation, filtering, and classification of variants as to their likely role in the disease or phenotype for which the test was submitted. The laboratory must ensure that the final report is accurate and complete without overinterpretation. Annotation of each variant detected in the sequencing means attaching information about population frequency (or lack of previous observation), differences in frequency between various ethnic groups, effect of the variant on protein coding sequence, effect on splicing, and so on. Diagnostic laboratories must actively use information contained in the public sequence databases to interpret the biological consequences of mutations and polymorphisms. There are databases that have an important role: *(i)* GenBank, which contains the reference human genome sequences (http://www.ncbi.nlm.nih.gov/entrez); *(ii)* dbSNP, which contains more than 150 000 000 common single-nucleotide polymorphisms; *(iii)* Exome Sequencing Project (https://esp.gs.washington.edu/drupal/), which has exome data on about 6500 individuals from several research projects; *(iv)* the 1000 Genomes Project (http://www.1000genomes.org/), which has exome and genome data from about 2500 multiethnic controls; and the *(v)* Exome Aggregation Consortium (http://exac.broadinstitute.org/), which has assembled variant data on more than 60 000 subjects from various disease and populations studies.

Interpretation of sequence and genotype information is the next step in carrying out a molecular diagnostic test. The laboratory must do the following: take into consideration the patient's presenting symptoms and other clinical data; examine the strength of the evidence that specific gene mutations can cause the suspected disease; look for variants in genes that might provide a causal explanation; determine whether a candidate variant has previously been reported to cause a disease; determine whether variants that appear in previous databases have a low frequency (*i.e.*, compatible with the observed frequency of the rare disease); determine whether loss-of-function, gain-of-function, dominant negative, haploinsufficiency, and so on are known underlying mechanisms of disease; determine whether variant occurs in an exon and/or protein domain compatible with alteration of function; and summarize whether the candidate variants occur in similarly affected family members (pattern of inheritance and cosegregation). If a variant has been observed in the proband and other family members are tested, then the expectation is that it will exactly segregate with the disease. An example would be testing for a mutation in X-linked SCID in two brothers. The causal variant will be present in both. When the proband is the only affected individual, then one may look to the literature; ClinVar (www.clinicalgenome.org/data-sharing/clinvar/), a database of variants with curated interpretations; and locus-specific public databases (bioinf.uta.fi/base_root/mutationdatabases.php) to determine whether the specific variant has previously been observed in another affected person. However, it is routine to find variants that have not previously been observed. Assessment of the functional consequence of a newly observed sequence variant is problematic, as there is no guarantee that it is anything more than a rare, but neutral, polymorphism. Loss-of-function mutations (*e.g.*, nonsense, frameshift, and conserved splice site mutations) have clear functional consequences. Missense substitutions are more difficult. Several methods based on sequence conservation and the chemical properties of amino acids have been developed. These are available in software, such as Sorting Intolerant From Tolerant (SIFT) (http://blocks.fhcrc.org/sift/SIFT.html), PolyPhen (http://coot.embl.de/PolyPhen/), and several others. CADD represents a class of software that creates a joint model based on integration of many of these methods.[33] All these criteria are put together in a process called *gene and variant curation.* Recommendations for scoring variants into five categories—pathogenic, likely pathogenic, variant of unknown significance, likely benign, and benign—have been given by the American College of Medical Genetics and Association for Molecular Pathology.[34]

The next step in the process is assembly of filtered and classified variants into the laboratory report. The report must be simple enough for physicians who are not experts in genomics to understand but, at the same time, must document the intermediate results required to demonstrate analytical accuracy and clinical validity. Clinical reports must contain clear nomenclature for the position and consequence of relevant variants along with appropriate indication of test and interpretive limitations. The final steps of interpretation and reporting may be facilitated by communication with the referring clinician so that there is active collaboration in arriving at a clinical diagnosis.

The final steps in bioinformatics management of NGS tests are aggregation and data sharing. Submission of deidentified cases to public-knowledge databases, such as ClinVar, will allow aggregation of variant data accompanied by sufficient clinical information to support variant classification. Once this information is accumulated in many clinical contexts, the value for patient care will be magnified beyond what could be inferred from the case data alone. Data sharing that is consistent, standardized, and collected with appropriate protections for confidentiality will be an invaluable resource for the future.

Clinical Performance of Genomics Assays

The development of specialized training programs in molecular genetics and molecular genetic pathology by the American Boards of Medical Genetics and Pathology[35] highlights the recognition by professional groups that this area of clinical testing is exceptionally complex. Although inferences made about disease diagnosis and carrier status based on direct detection of mutations are inherently categorical, data interpretation should incorporate known genotype–phenotype correlations, variable disease expression/severity, incomplete penetrance, gender-specific risk, and other available data. In the future, as DNA testing is employed in the assessment of more complex traits such as autoimmune and neoplastic diseases, genotype data will perhaps be expressed as "relative risk" and incorporate gene–gene and gene–environment interactions. The reliability of these estimates will require standard data acquisition algorithms and constant updating of population-based data.

RECOMMENDATIONS FOR USE

KEY CONCEPTS

Principles of Molecular Diagnostics

- DNA diagnostics can play an important role in the diagnosis of specific diseases in affected individuals and in genetic risk assessment for their family members.
- Genotyping can be used to conduct prenatal diagnoses.
- Prenatal diagnosis may be used not only for elective termination of pregnancy but also for treatment planning.
- Provision of genetic testing results to patients and their families carries responsibilities for sensitivity and confidentiality.

Clinically Important Applications of Next-Generation Sequencing

Molecular diagnostic techniques have a wide range of potential applications in clinical immunology. DNA diagnostic procedures are used to *(i)* perform HLA genotyping; *(ii)* analyze and monitor neoplastic disease; *(iii)* provide identification or DNA "fingerprinting"; *(iv)* monitor bone marrow engraftment; *(v)* establish a genetic diagnosis in a symptomatic individual; *(vi)* determine the risk of occurrence of a disease in offspring; and *(vii)* to establish a prenatal diagnosis. The use of DNA analysis for HLA typing is described in Chapter 5. DNA techniques are used in leukemias and lymphomas, primarily for investigation of cell lineage, proliferative clonality, and measurement of residual abnormal cells after therapy. Molecular analysis, and especially molecular cytogenetic analysis, is important in guiding initial and follow-up therapies (Chapters 78 and 79).

Targeted Panels

When the phenotype is sufficiently circumscribed that a limited list of genes is known to account for a high percentage of cases, it may be appropriate to use a targeted gene panel as the first-round genetic diagnostic test. The reasons include assay design strategy that puts maximum emphasis on complete gene coverage; lower unit cost of testing; and reduced occurrence of secondary findings (*i.e.,* apparently pathogenic variants in genes not suspected from the primary indication for testing). But efficient use of targeted panels requires high sensitivity in the sense that most of the important disease-causing genes must have been accounted for and give a conclusive diagnosis in a large fraction of clinical cases. This situation is often not met because of the lack of specificity in clinical presentation or the very large degree of locus heterogeneity. The latter is particularly important in PIDs.

Whole Exome/Whole-Genome Sequencing

In the past, clinical genetic testing would necessarily stop after the most likely genes had been investigated in a few targeted laboratory tests. That strategy changed in the last few years after it became possible to efficiently assess most known disease genes in a single test. Two approaches are in use today: WES and WGS. WES depends on "capture" of the 1–1.5% of the genome that contains the protein-coding segments of genes. Capture is typically achieved by liquid phase hybridization followed by PCR and sequencing library construction. The rationale for WES is that most interpretable disease causing variants occur in coding sequence, and thus the sequencing is focused on the interpretable fraction of the genome. The lower unit cost of WES also allows for greater read depth with concomitantly lower error rates. Some significant limitations of WES include incomplete coverage of disease-causing genes because of inefficient capture; uneven coverage across all targeted regions such that adding more read depth inefficiently adds to poorly covered regions; poor ability to resolve structural variants; limited ability to call copy number variants; and inability to call triplet repeat mutations. WGS, especially when implemented with PCR-free library construction techniques, overcomes most of these limitations. The falling cost of WGS, especially when coupled to automated library preparation systems, holds the promise of nearly complete analysis of the genome of individuals.

Cell-Free DNA and Liquid Biopsy

It has been known for many years that there is normally a small amount of genomic DNA in the cell-free plasma fraction of blood (called *cfDNA*). It is thought that this material comes from the normal processes of cell apoptosis, particularly from leukocytes. During pregnancy, there is a significant contribution of DNA from the placenta, and thus cfDNA can be used as a biological surrogate for the fetal genome. NGS enables sampling of this DNA at very high read depth, and common chromosomal imbalances in the fetus can be detected by analysis of DNA extracted from maternal plasma. The observation of a patient who had cancer during pregnancy led to the discovery that cancer cells may also contribute to cfDNA. New genomic assays for detection of cancer recurrence, changes in mutation spectrum, and clonal expansion are rapidly arriving in clinical practice. These assays are likely to be especially useful in lymphomas and other solid tumors that have required invasive biopsies. Serial testing by these so called liquid biopsy methods is very likely to have a large impact in immune-oncology practice.

LABORATORY STANDARDS AND REPORTING

All US laboratories providing any clinical data are subject to the regulatory requirements embodied in CLIA88.[36] CLIA88 mandates biannual laboratory inspections, quality control, quality assurance, proficiency testing, and personnel training standards. The American College of Medical Genetics has also produced guidelines for diagnostic laboratories conducting molecular genetics testing. A clear argument can be made for the clinical utility of early diagnosis for rare diseases, but demonstrating reduced cost of care or improved clinical outcome is difficult. Genomics tests, like all diagnostic testing, should be designed and developed such that the practiced test can meet the requirements of its intended use. The performance of genomic tests must be validated before they are offered for clinical diagnostic use. Critical measures of performance include sensitivity—how often the test is positive when a disease causing variant is present; specificity—how often the test is negative when a mutation is not present; "technical" positive predictive value (TPPV)—the fraction of true positives divided by the sum of the true positive and false-positive tests; and positive and negative percent agreement (PPA) with previous reference standards. These are all measures recommended by the FDA in its draft guidance for validating clinical NGS tests. New methods to test and validate bioinformatics software are also needed as the field moves forward (Fig. 96.7).

FUTURE DIRECTIONS

The next decade will bring further improvements in our ability to identify family-specific pathogenic variants for all genes that cause rare mendelian disorders, such as PIDs. Sequencing technology will continue to develop, further reducing the cost of genome analysis for individuals and families. As our understanding of the role of polymorphisms in disease risk increases, the importance of low-cost, high-throughput genotyping of standard variant panels will also increase.[37] The importance of bioinformatics in data analysis will become increasingly apparent as large amounts of individual sequence data are produced. International clinical variant databases will play an important role in both diagnosis and prognosis. New statistical approaches will be needed to exploit fully the potential of complete genome sequence for estimation

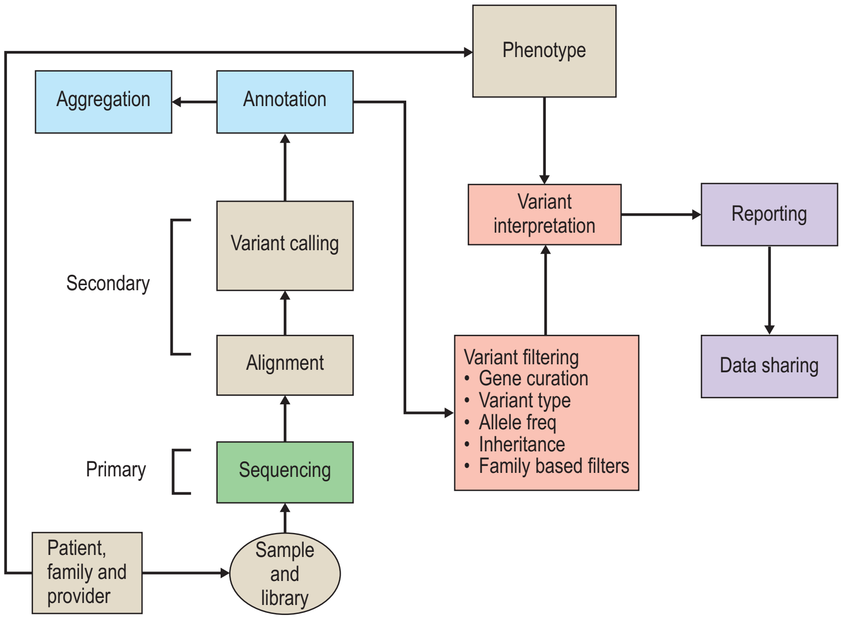


FIG 96.7 **Clinical Bioinformatics.** Primary, secondary, and tertiary phases of data analysis and reporting.

of disease risk. Prospective screening of newborns for the T-cell immune deficiencies is feasible.[38,39] Although rare, these are treatable disorders whose prognosis would be altered by early diagnosis before the onset of serious infection. In aggregate, the birth incidence is sufficient to merit screening if the techniques could be interfaced with current state screening programs. Screening for premalignant genotypic changes in peripheral blood might also be a reality one day. Careful clinical studies, including examination of the costs and benefits of mass screening programs, will be essential.

ON THE HORIZON

- Routine diagnosis of >300 primary immunodeficiencies by whole-genome sequencing
- Rapid discovery of new primary immune deficiency genes by whole genome sequencing in individual families
- Preconceptional screening to identify carrier status for thousands of genetic disorders
- Noninvasive prenatal screening of fetal DNA in maternal blood for copy number abnormalities and mutations

TRANSLATIONAL RESEARCH AND CLINICAL APPLICATION

DNA sequencing technology is evolving very rapidly, and we can anticipate additional technical platforms to become available over the next decade. Massive increases in sequence output, speed of sequencing, increases in read length, and more efficient bioinformatics methods will fuel further reductions in cost. Complete genome sequencing will likely become a first-line medical test for the diagnosis of most suspected genetic disorders, including the PIDs. Gene identification in single families with previously undescribed immune deficiencies will become routine. Personal complete genome sequencing for preconceptual genetic screening in the general population will become available and will impact risk counseling for thousands of patients with autosomal recessive and X-linked disorders. By analysis of either free DNA in maternal serum or circulating fetal cells, complete genome sequencing will provide a comprehensive testing platform for noninvasive prenatal diagnosis. Incorporation of other methods, such as RNA sequencing, will undoubtedly become part of routine functional assessment of suspected pathogenic variants. Single-cell sequencing technology promises to give much greater detail on the beautifully complicated problem of immune cell heterogeneity. Although single-cell sequencing has not yet been shown to have role in clinical diagnostics, it may present the next step in the long-term goal of understanding immunological diseases at the cellular level.

Please check your eBook at https://expertconsult.inkling.com/ for self-assessment questions. See inside cover for registration details.

REFERENCES

1. Raje N, Dinakar C. Overview of Immunodeficiency Disorders. Immunol Allergy Clin North Am 2015;35:599–623.
2. Picard C, Al-Herz W, Bousfiha A, et al. Primary Immunodeficiency diseases: an update on the classification from the International Union of Immunological Societies Expert Committee for Primary Immunodeficiency, 2015. J Clin Immunol 2015;35:696–726.
3. Genomes Project C, Auton A, Brooks LD, et al. A global reference for human genetic variation. Nature 2015;526:68–74.
4. International HapMap C, Altshuler DM, Gibbs RA, et al. Integrating common and rare genetic variation in diverse human populations. Nature 2010;467:52–8.
5. Picard C, Fischer A. Contribution of high-throughput DNA sequencing to the study of primary immunodeficiencies. Eur J Immunol 2014;44:2854–61.
6. Platt C, Geha RS, Chou J. Gene hunting in the genomic era: approaches to diagnostic dilemmas in patients with primary immunodeficiencies. J Allergy Clin Immunol 2014;134:262–8.
7. MacArthur DG, Manolio TA, Dimmock DP, et al. Guidelines for investigating causality of sequence variants in human disease. Nature 2014;508:469–76.

8. de Jong H. Visualizing DNA domains and sequences by microscopy: a fifty-year history of molecular cytogenetics. Genome 2003;46:943–6.
9. Coughlin CR 2nd, Scharer GH, Shaikh TH. Clinical impact of copy number variation analysis using high-resolution microarray technologies: advantages, limitations and concerns. Genome Med 2012;4:80.
10. Martin CL, Kirkpatrick BE, Ledbetter DH. Copy number variants, aneuploidies, and human disease. Clin Perinatol 2015;42:227–42, vii.
11. Shendure J, Akey JM. The origins, determinants, and consequences of human mutations. Science 2015;349:1478–83.
12. Mullaney JM, Mills RE, Pittard WS, et al. Small insertions and deletions (INDELs) in human genomes. Hum Mol Genet 2010;19: R131–6.
13. Budworth H, McMurray CT. A brief history of triplet repeat diseases. Methods Mol Biol 2013;1010:3–17.
14. Fragouli E, Wells D, Delhanty JD. Chromosome abnormalities in the human oocyte. Cytogenet Genome Res 2011;133:107–18.
15. Jones KT, Lane SI. Molecular causes of aneuploidy in mammalian eggs. Development 2013;140:3719–30.
16. Campbell CD, Eichler EE. Properties and rates of germline mutations in humans. Trends Genet 2013;29:575–84.
17. Segurel L, Wyman MJ, Przeworski M. Determinants of mutation rate variation in the human germline. Annu Rev Genomics Hum Genet 2014;15:47–70.
18. Belmont JW. Genetic control of X inactivation and processes leading to X-inactivation skewing. Am J Hum Genet 1996;58:1101–8.
19. Liehr T, Weise A, Hamid AB, et al. Multicolor FISH methods in current clinical diagnostics. Expert Rev Mol Diagn 2013;13:251–5.
20. Barrett MT, Scheffer A, Ben-Dor A, et al. Comparative genomic hybridization using oligonucleotide microarrays and total genomic DNA. Proc Natl Acad Sci USA 2004;101:17765–70.
21. Lipshutz RJ, Fodor SP, Gingeras TR, et al. High density synthetic oligonucleotide arrays. Nat Genet 1999;21:20–4.
22. Peiffer DA, Le JM, Steemers FJ, et al. High-resolution genomic profiling of chromosomal aberrations using Infinium whole-genome genotyping. Genome Res 2006;16(9):1136–48.
23. Cooper GM, Zerr T, Kidd JM, et al. Systematic assessment of copy number variant detection via genome-wide SNP genotyping. Nat Genet 2008;40:1199–203.
24. Fan JB, Chee MS, Gunderson KL. Highly parallel genomic assays. Nat Rev Genet 2006;7:632–44.
25. Kennedy GC, Matsuzaki H, Dong S, et al. Large-scale genotyping of complex DNA. Nat Biotechnol 2003;21:1233–7.
26. Gunderson KL, Steemers FJ, Lee G, et al. A genome-wide scalable SNP genotyping assay using microarray technology. Nat Genet 2005;37:549–54.
27. Strom CM. Mutation detection, interpretation, and applications in the clinical laboratory setting. Mutat Res 2005;573:160–7.
28. Bentley DR, Balasubramanian S, Swerdlow HP, et al. Accurate whole human genome sequencing using reversible terminator chemistry. Nature 2008;456:53–9.
29. Rothberg JM, Hinz W, Rearick TM, et al. An integrated semiconductor device enabling non-optical genome sequencing. Nature 2011;475:348–52.
30. Goodwin S, McPherson JD, McCombie WR. Coming of age: ten years of next-generation sequencing technologies. Nat Rev Genet 2016;17:333–51.
31. Stoloff DH, Wanunu M. Recent trends in nanopores for biotechnology. Curr Opin Biotechnol 2013;24:699–704.
32. Roller E, Ivakhno S, Lee S, et al. Canvas: versatile and scalable detection of copy number variants. Bioinformatics 2016;32:2375–7.
33. Kircher M, Witten DM, Jain P, et al. A general framework for estimating the relative pathogenicity of human genetic variants. Nat Genet 2014;46:310–15.
34. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med 2015;17:405–24.
35. Byers PH. Molecular genetic pathology. Coming of age in the molecular world. Am J Pathol 1999;155:673–4.
36. Rivers PA, Dobalian A, Germinario FA. A review and analysis of the clinical laboratory improvement amendment of 1988: compliance plans and enforcement policy. Health Care Manage Rev 2005;30:93–102.
37. Worthey EA, Mayer AN, Syverson GD, et al. Making a definitive diagnosis: successful clinical application of whole exome sequencing in a child with intractable inflammatory bowel disease. Genet Med 2011;13:255–62.
38. Routes JM, Grossman WJ, Verbsky J, et al. Statewide newborn screening for severe T-cell lymphopenia. JAMA 2009;302:2465–70.
39. Verbsky JW, Baker MW, Grossman WJ, et al. Newborn screening for severe combined immunodeficiency; the Wisconsin experience (2008-2011). J Clin Immunol 2012;32:82–8.

MULTIPLE-CHOICE QUESTIONS

1. Population screening for carrier status in genes that cause primary immunodeficiencies (PIDs) is made more difficult by:
   A. The prominent role of triplet repeat expansions among causative mutations
   B. Most primary immune deficiencies being inherited as autosomal dominant traits
   C. A large number of different genes (locus heterogeneity) and many very rare pathogenic variants (allelic heterogeneity)
   D. Most carriers coming from populations with a history of consanguinity

2. In the workup of a newborn with congenital heart disease, hypocalcemia, and absent thymus, the optimal testing approach is:
   A. Fluorescent *in situ* hybridization (FISH) for 22q11 microdeletions
   B. Array comparative genomic hybridization (aCGH)
   C. Sanger sequencing of the *TBX1* gene
   D. Next-generation sequencing whole exome sequencing

3. Next-generation sequencing (NGS) diagnostics used many bioinformatics methods, including:
   A. Base calling from raw instrument data (primary analysis)
   B. Demultiplexing, alignment to the reference genome, and variant calling (secondary analysis)
   C. Annotation of variants for their predicted effect on proteins and splicing
   D. Classification of variants for their likely role in causing disease
   E. All of the above

APPENDIX 1

Selected CD Molecules and Their Characteristics

Thomas A. Fleisher

CD Molecule	Predominant Distribution	Identity/Function
CD1a–e	Thymocytes, subset of lymphocytes, antigen-presenting cells	MHC class I–like molecules; presentation of nonpeptide antigens to T cells; thymic T-cell development
CD2	T cells	Binds to LFA-3; receptor for CD48; T-cell activation; adhesion
CD3	T cells	T-cell signaling complex; associated with TCR
CD4	T-cell subset	TCR coreceptor; interacts with MHC class II molecules on antigen-presenting cells; identifies T cells with helper function; signal transduction
CD5	Most T cells, thymocytes, B-cell subset	Binds to CD72; regulation of cell proliferation/activation; identifies B-1 B-cell subset
CD6	Thymocytes, T cells, B-cell subset	Binds CD166 (ALCAM); adhesion; mediates binding of developing thymocytes with thymic epithelial cells; thymic development; T-cell activation
CD7	Pluripotent hematopoietic cells, thymocytes, T cells	T-cell and NK-cell development
CD8	T-cell subset	TCR coreceptor; interacts with MHC class I molecules on antigen-presenting cells; identifies T cells with cytotoxic function
CD10	B cells	Neutral endopeptidase; enkephalinase; B-cell development; common acute lymphoblastic leukemia antigen (CALLA), neutrophils
CD11a	Leukocytes	α-chain of LFA-1; pairs with CD18; interacts with ICAM; adhesion and cellular migration
CD11b	Monocytes, granulocytes, NK cells	α-chain of complement receptor type 3 (CR3); pairs with CD18; adhesion molecule
CD11c	Monocytes, granulocytes, NK cells	α-chain of complement receptor type 4 (CR4); pairs with CD18; adhesion molecule
CD13	Hematopoietic stem cells, immature and mature myeloid and monocyte elements	Appears before CD33 during myeloid differentiation
CD14	Granulocytes, monocytes/macrophages	Receptor for LPS/LPB complex; myeloid differentiation antigen; cell activation
CD15	Neutrophils, eosinophils, monocytes, basophil subset	Sialyl Lewis X antigen, plays a role in cell adhesion, defective in leukocyte adhesion deficiency type 2
CD16a,b	NK cells, monocytes/macrophages, neutrophils	FcγRIIIA and FcγRIIIB (low-affinity IgG receptors-type III); phagocytosis; ADCC
CD18	Leukocytes	β-chain of β_2-integrin molecules, including LFA-1, CR3, and CR4; pairs with CD11a, b, and c
CD19	B cells	BCR coreceptor; signal transduction; complexes with CD21, expressed pre-B cells
CD20	B cells	Role in B-cell activation/differentiation
CD21	B cells; follicular dendritic cells	Complement receptor type 2 (CR2): C3d receptor; B-cell coreceptor subunit; EBV receptor
CD22	B cells	Associates with BCR; signaling; regulation of B-cell activation; adhesion
CD23	B cells, macrophages, eosinophils, platelets, follicular dendritic cells	FcεRII; (low-affinity IgE receptor), activated B cells
CD24	Leukocytes	Heat-stable antigen; costimulation; adhesion
CD25	Activated T cells and B cells	α-chain of IL-2 receptor, low-affinity IL-2 binding; signaling for cell proliferation/differentiation
CD26	Activated T cells and B cells; macrophages	Dipeptidyl peptidase; role in extracellular adhesion; cell activation
CD27	T cells; B-cell subset	Costimulation; T-cell proliferation, memory B cells
CD28	T cells	Binds B7-1 (CD80) and B7-2 (CD86); T-cell costimulation; signal transduction
CD29	Leukocytes	Integrin β_1-chain; pairs with CD49a-CD49f to form VLA-1-VLA-6 integrins, respectively; adhesion; signal transduction; development
CD30	Activated B cells and T cells	Binds to CD153; T-cell activation/regulation/differentiation, Reed–Sternberg cells
CD31	Monocytes, granulocytes, platelets, endothelial cells, B cells Recent thymic emigrants	PECAM-1; binds to CD38; adhesion; signal transduction
CD32	B cells, monocytes/macrophages, granulocytes, eosinophils	FcgRII (low-affinity IgG receptor-type II) phagocytosis; ADCC; B-cell regulation
CD33	Myeloid progenitors, granulocytes	Binds sialic acid
CD34	Hematopoietic progenitor cells; capillary endothelium	Mucosialin; binds to CD62L; adhesion

Continued

CD Molecule	Predominant Distribution	Identity/Function
CD35	Leukocytes; erythrocytes	Complement receptor 1 (CR1); C3b and C4b receptor; phagocytosis
CD36	Monocytes/macrophages; endothelium; platelets	Binds oxidized LDL; scavenger receptor; binds apoptotic cells; adhesion and endocytic receptor; GPIIIb; platelet adhesion and aggregation
CD38	NK cells; T-cell and B-cell subsets; monocytes	Binds CD31; cyclase; hydrolase; cell activation
CD40	B cells; antigen-presenting cells	Binds CD154 (CD40 ligand); B-cell proliferation, differentiation, and survival; T-cell costimulation
CD41	Megakaryocytes/platelets	Glycoprotein IIb; a IIb integrin chain; binds fibronectin, fibrinogen, von Willebrand factor, thrombospondin; extracellular adhesion; platelet aggregation
CD43	Leukocytes (except resting B cells)	Leukocyte sialoglycoprotein; may bind CD54; signal transduction; adhesion; antiadhesion
CD44	Leukocytes; memory T cells; erythrocytes	Binds hyaluronan (H-CAM), collagen, fibronectin, laminin, osteopontin; extra- and intercellular adhesion; T-cell costimulation; leukocyte homing
CD45	Leukocytes (pan-leukocyte marker)	Protein tyrosine phosphatase; cell differentiation; lymphocyte signal transduction and activation
CD45RA	Naive T-cell marker (in conjunction with CD62L); B cells, monocytes	CD45 isoform
CD45RB	B-cell and T-cell subsets, monocytes/macrophages, granulocytes	CD45 isoform
CD45RO	Memory and activated T cells; B cells, monocytes/macrophages	CD45 isoform
CD46	Hematopoietic cells	Membrane cofactor protein (MCP); binds to C3b and C4b and regulates complement pathway
CD47R	Leukocytes; endothelium	Integrin-associated protein (IAP); leukocyte migration, extravasation, and activation
CD48	Leukocytes (not neutrophils)	Binds CD2; adhesion; costimulation
CD49a–f	Various distributions	Integrin a1-6 chain; binds to CD29 to form VLA-1 to VLA-6; binds extracellular matrix components such as fibronectin, laminin, collagen (CD49D binds VCAM-1, fibronectin, MAdCAM-1, invasin); lymphocyte homing; extracellular adhesion; embryonic development
CD50	Thymocytes, B cells, T cells, monocytes, granulocytes	ICAM-3; adhesion
CD51	Platelets/megakaryocytes, granulocytes, monocytes, T cells	Integrin α-chain; associates with CD61; binds vitronectin, fibronectin, fibrinogen; extracellular adhesion; T-cell costimulation; epithelial inflammatory response
CD52	Leukocytes	GPI linked, signaling, defined by CAMPATH-1
CD54	Broad distribution; increased on activated leukocytes	ICAM-1; binds LFA-1; adhesion; leukocyte transendothelial migration, rhinovirus receptor
CD55	Hematopoietic cells and some nonhematopoietic cells	Decay accelerating factor (DAF); binds complement fragment C3b; regulation of complement activation
CD56	NK cells, NK-T cells	NKH-1; adhesion
CD57	NK cells, T-cell subset, B cells, monocytes	Oligosaccharide expressed on cell surface glycoproteins
CD58	Hematopoietic cells and non-hematopoietic cells	LFA-3; binds CD2; adhesion; lymphocyte coactivation
CD59	Hematopoietic cells and non-hematopoietic cells	Binds complement components C8 and C9 and regulates assembly of complement membrane attack complex
CD61	Platelets/megakaryocytes, macrophages	Integrin β_3 subunit; associates with CD41 or CD51
CD62E	Endothelium	ELAM-1 or E-selection; binds sialyl-Lewis X; adhesion; mediates rolling interaction of neutrophils on endothelium and neutrophil extravasation
CD62L	B cells, T cells, monocytes, NK cells	LECAM-1, LAM-1 (or L-selectin); binds CD34 and GlyCAM; adhesion; mediates rolling interactions with endothelium and call extravasation
CD62P	Platelets/megakaryocytes, endothelium	P-selectin; binds sialyl-Lewis X; mediates interaction of platelets with neutrophils and monocytes; mediates rolling interaction of neutrophils with endothelium
CD64	Monocytes/macrophages, mature neutrophils	FcγR1 (high-affinity IgG receptor)
CD65	Neutrophils, eosinophils, basophils monocyte subset, CD56 bright NK cells	During myeloid differentiation it appears after myeloperoxidase, appears to be the ligand for CD62L
CD66c	Differentiating myeloid cells peaks at promyelocyte stage, subpopulation of monocytes	Member of the carcinoeEDryonic antigen (CEA) family
CD68	Monocytes/macrophages	Macrosialin; early activation antigen; role in phagocytic activity
CD69	Activated T cells, B cells, macrophages, NK cells	Early activation antigen; costimulation
CD70	Activated B cells and T cells; macrophages	Binds to CD27; costimulation
CD71	Activated leukocytes, erythroid precursors	Transferrin receptor; cell activation
CD72	B cells	Ligand for CD5; B-cell activation and differentiation; costimulation

CD Molecule	Predominant Distribution	Identity/Function
CD73	B-cell and T-cell subsets	Ecto-5′ nucleotidase; allows nucleoside uptake
CD74	MHC class II expressing cells	MHC class II–associated invariant chain; involved in antigen processing and peptide presentation in antigen-presenting cells
CD77	Germinal center B cells	Entering apoptosis
CD79a,b	B cells	Iga, Igb; components of BCR complex that mediate signal transduction
CD80	Monocytes/macrophages, dendritic cells, activated B cells	B7-1; ligand for CD28 and CTLA-4; T-cell interaction with antigen-presenting cells; costimulation
CD81	Lymphocytes	Associates with CD19 and CD21 to form B-cell coreceptor complex; costimulation; adhesion
CD86	Monocytes, activated B cells	B7-2; ligand for CD28 and CTLA-4; T-cell interaction with antigen-presenting cells; costimulation
CD87	Granulocytes; monocytes/macrophages, activated T cells	Urokinase plasminogen activator receptor
CD88	Granulocytes, macrophages, mast cells	Complement component fragment C5a receptor
CD89	Monocytes/macrophages, granulocytes, B-cell and T-cell subsets	FcαR (IgA receptor)
CD91	Monocytes	α_2-macroglobulin receptor
CD94	NK cells, T-cell subset	Inhibits killing
CD95	Broad distribution	Fas or APO-1; induces apoptosis after being bound by Fas ligand
CD97	Activated B and T cells, monocytes, PMN	Activation antigen
CD102	Resting lymphocytes, monocytes, endothelial cells	ICAM-2; binds LFA-1 (CD11a/CD18); adhesion; T-cell costimulation
CD103	Intraepithelial lymphocytes, T-cell subset	αE integrin; T-cell development and costimulation
CD104	Epithelial cells, Schwann cells	β_4-integrin; epidermal adhesion to basement membrane
CD105	Endothelial cells, bone marrow cell subset, activated macrophages	Endoglin; adhesion
CD106	Endothelial cells	VCAM-1; ligand for VLA-4; lymphocyte adhesion; embryonic development
CD107a,b	Epithelial cells, subsets of monocytes, granulocytes, and lymphocytes	LAMP-1 and LAMP-2; adhesion
CD110	Hematopoietic stem cells, megakaryocytes, platelets	Thrombopoietin receptor; megakaryocyte proliferation and differentiation
CD114	Granulocytes	G-CSF receptor; regulates granulopoiesis
CD115	Monocytes/macrophages	M-CSF receptor; cell differentiation
CD116	Monocytes, neutrophils, eosinophils	GM-CSF a chain receptor; cell differentiation
CD117	Hematopoietic progenitors, mast cells	c-*kit*; stem cell factor (SCF) receptor; hematopoietic cell differentiation
CD118	Broad distribution	Type 1 interferon (interferon-α/β) receptor
CD119	Broad distribution	Interferon-γ receptor
CD120a	Hematopoietic cells, nonhematopoietic cells, myeloid cells	TNF receptor-type I signal transduction; apoptosis
CD120b	Hematopoietic cells, nonhematopoietic cells, myeloid cells	TNF receptor-type II; signal transduction; apoptosis
CD121a	Thymocytes, T-cell subset, fibroblasts, epithelial cells, and brain cells	IL-1 receptor-type I; signal transduction
CD121b	T-cell subset, myeloid cell subsets	IL-1 receptor-type II
CD122	NK cells, T-cell and B-cell subset	IL-2 and IL-15 receptor β chain; signal transduction; regulation of lymphocyte development, differentiation, activation, and proliferation
CD123	Bone marrow stem cells, granulocytes, monocytes, megakaryocytes	IL-3 receptor α chain; cell development and differentiation
CD124	Mature B cells and T cells, hematopoietic precursor cells	IL-4 receptor; signal transduction; lymphocyte development, activation, differentiation, and proliferation
CD125	Eosinophils, basophils, B-cell subset	IL-5 receptor; eosinophil and B-cell growth and differentiation
CD126	Activated B cells, plasma cells, T cells, granulocytes, monocytes/macrophages; also expressed on epithelial cells, fibroblasts, hepatocytes, and neural cells	IL-6 receptor α chain; regulation of B-cell and T-cell differentiation and function; hematopoiesis
CD127	Bone marrow lymphoid precursors, pro-B cells, T-cell precursors, T-cell subset, monocytes	IL-7 receptor α chain; signal transduction; B-cell and T-cell proliferation and differentiation
CD128	Neutrophils, basophils, T-cell subset	IL-8 receptor; neutrophil activation and migration
CD129	T cells	IL-9 receptor α chain; T-cell proliferation
CD130	Broad distribution	IL-6 receptor β chain (with CD126); signal transduction
CD131	Lymphocytes, granulocytes, monocytes	IL-3, IL-5, and GM-CSF receptor; common β chain; signal transduction; see CD123 and CD125
CD132	Lymphocytes	Common γ chain of high-affinity receptor for IL-2 (with CD25 and CD122), IL-4 (with CD124), IL-7 (with CD127), IL-9 (with CD129), and IL-15 (with CD122) receptors; signal transduction

Continued

CD Molecule	Predominant Distribution	Identity/Function
CD134	Activated T cells	OX-40 antigen of TNFR superfamily (binds OX-40 ligand); T-cell–B-cell interaction and T-cell costimulation
CD135	Lymphoid and myeloid cell progenitor subsets	Flt3 ligand receptor; development of myeloid and lymphoid progenitors
CD137	Activated T cells	4-1BB; binds 4-1BB ligand and extracellular matrix components; T-cell–B-cell interaction and T-cell costimulation; extracellular adhesion; signal transduction
CD138	B-cell subset, plasma cells, epithelial cells	Syndecan-1; binds interstitial matrix proteins; B cell–matrix interactions
CD140a,b	Endothelial cells	PDGF receptor α and β chain; embryonic development; signal transduction; chemotaxis
CD141	Endothelium	Thrombomodulin (binds thrombin); regulates coagulation
CD142	Endothelium	Tissue factor; binds plasma factors VII/VIIa; hemostasis, coagulation, and angiogenesis
CD143	Endothelium	Angiotensin-converting enzyme (ACE); binds angiotensin 1; regulates blood pressure
CD144	Endothelium	VE-cadherin; cell–cell adhesion; maintenance of endothelium integrity
CD146	Activated T-cell subset	Mel-CAM, adhesion molecule during development
CD150	T-cell and B-cell subsets	Surface lymphocyte activation marker (SLAM); B cell–T cell interaction; costimulation
CD151	Not defined	PETA-3; regulates platelet aggregation and mediator release
CD152	Activated T cells	CTLA-4; binds B7-1 (CD80) and B7-2 (CD86); T-cell costimulation–negative signal
CD153	T cells	CD30 ligand; T-cell activation, differentiation, and regulation
CD154	Activated T cells	CD40 ligand; T-cell costimulation
CD156a	Leukocytes, B cells	Transmembrane glycoprotein; disintegrin and metalloproteinase domain (ADAM) family member; leukocyte adhesion and protease function; infiltration of myelomonocytic cells
CD156b	Broad distribution	TNF-α-converting enzyme (TACE); disintegrin and metalloproteinase domain (ADAM) family member; cleaves TNF and transforming growth factor-α from cell surface, thereby releasing soluble form
CD158	NK cells	Killer immunoglobulin-like receptors (KIR); family of molecules that inhibit NK cytotoxic activity
CD159a	NK cells	NKG2A (killer cell lectin-like receptor)
CD161	NK cells	Natural killer cell receptor-P1; target cell recognition; NK cell activation
CD162	Granulocyte and T-cell subsets	P-selectin glycoprotein ligand-1 (PSGL-1); adhesion
CD166	Activated T cells, B cells	ALCAM; binds CD6; T-cell activation; thymocyte development
CD167a	Epithelial cells	Discoidin domain receptor 1 (DDD1); tyrosine kinase receptor; binds to collagen; cell–cell contact and adhesion
CD178	Activated T cells; various tissue cells	FAS ligand (ligand for CD95); binding to FAS triggers apoptosis
CD179a	Pro-B and pre-B cells	VpreB; forms surrogate light chain with CD179b; early B-cell differentiation
CD179b	Pro-B and pre-B cells	$\lambda 5$; forms surrogate light chain with CD179a; early B-cell differentiation
CD180	B cells	RP105; toll-like receptor family; regulates B-cell recognition and signaling of LPS
CD183	Effector/memory T cells, NK cells, eosinophils	CXCR3 receptor for interferon-inducible chemokines IP10, Mig, and I-TAC; chemotactic migration of effector T cells into areas of inflammation
CD184	Leukocytes; hematopoietic progenitors	CXCR4 receptor for chemokines such as stromal cell-derived factor 1 (SDF-1) (fusin); chemotaxis; HIV-1 coreceptor
CD195	Broad distribution; myeloid cells, lymphocytes, T lymphocytes, neurons, epithelium, endothelium	CCR5 receptor for chemokines such as macrophage inflammatory proteins, MIP-1a and MIP-1b, and RANTES; chemotaxis; HIV-1 coreceptor
CDw197	Lymphoid tissues, B cells, T-cell subset	CCR7 chemokine receptor; chemotaxis; T-cell homing and migration
CD201	Endothelial cells	Protein C receptor; coagulation
CD203c	Mast cells, basophils	Member of the ectonucleotide pyrophosphatase/phosphodiesterase enzymes
CD204	Myeloid cells, monocytes/macrophages	Macrophage scavenger receptor-1 (MSR1); mediates binding, internalization, and processing of various negatively charged macromolecules
CD206	Dendritic cells, macrophages, myeloid cells, endothelial cells	Mannose receptor, C-type 1; binds microorganisms; phagocytosis
CD207	Dendritic cells, Langerhans cells	Langerin; mannose receptor; phagocytosis and internalization of antigen for processing
CD208	Dendritic cells	DC-LAMP
CD209	Dendritic cells	DC-SIGN
CDw210	Broad distribution	I1-10 receptor α and β chain
CD212	T cells, NK cells	IL-12 receptor β_1 chain
CD213a1,a2	Lymphocytes, bronchial epithelial and smooth muscle cells	IL-13 receptor α_1 and α_2 chains
CDw217	Activated T-cell subset	IL-17; cytotoxic T-lymphocyte–associated serine esterase 8; stimulates cell activation; induces osteoclast differentiation factor (ODF)
CD220	Broad distribution	Insulin receptor; stimulates glucose uptake
CD molecule	Predominant distribution	Identity/function
CD221	Broad distribution	Insulin-like growth factor 1 receptor; cell signaling; activation, and differentiation
CD222	Broad distribution	Mannose-6-phosphate receptor; insulin-like growth factor 2 receptor
CD226	NK cells, platelets, monocytes, T-cell subset	Platelet and T-cell activation antigen 1 (PTA1); adhesion
CD233–241	Erythrocytes	Various erythrocyte membrane antigens, including blood group–associated glycoproteins

CD Molecule	Predominant Distribution	Identity/Function
CD242	Erythrocytes	ICAM-4
CD246	T cells	TCR or CD3 ζ chain; associated with TCR and CD3; couples TCR recognition with T-cell signaling
CD247	T cells	T cells ζ chain of the TcR
CD252	Activated B cells	OX40 ligand
CD253	Activated T cells	TRAIL, death receptor
CD254	Activated T cells, LN and BM stroma	RANK ligand
CD256	Monocytes, macrophages	APRIL, binds TACI and BCMA
CD257	Activated monocytes	BLyS, BAFF, binds TACI, BCMA, BAFFR, induces B-cell proliferation
CD261	Activated T cells, peripheral leukocytes	TRAIL-R2, DR5, death receptor
CD262	Peripheral lymphocytes	TRAIL-R1, DR4, death receptor
CD263	Peripheral lymphocytes	TRAIL-R3, DcR1, death receptor
CD264	Peripheral lymphocytes	TRAIL-R4, DcR2, death receptor
CD265	Broad distribution	RANK
CD267	B cells, activated T cells	TACI
CD268	B cells	BAFFR, binds BLys, mature B-cell survival
CD269	Mature B cells	BCMA, binds APRIL and BAFF, B-cell survival and proliferation
CD275	B cells, dendritic cells, monocytes	ICOSL, costimulation, cytokine production
CD278	Activated T cells	ICOS, T-cell costimulation

ADCC, antibody-dependent cellular cytotoxicity; *ALCAM,* activated leukocyte cell adhesion molecule; *APRIL,* a proliferation-inducing ligand; *BCMA,* B-cell maturation antigen; *BCR,* B-cell receptor for antigen; *DC,* dendritic cell; *CTLA,* cytotoxic T lymphocyte antigen; *EBV,* Epstein-Barr virus; *ELAM,* endothelial leukocyte adhesion molecule; *G-CSF,* granulocyte–colony-stimulating factor; *GM-CSF,* granulocyte macrophage–colony-stimulating factor; *GPI,* glycosyl-phosphatidylinositol; *HIV,* human immunodeficiency virus; *ICAM,* intercellular adhesion molecule; *ICOS,* inducible T-cell costimulator; *Ig,* immunoglobulin; IL, interleukin; *LAM,* leukocyte adhesion molecule; *LAMP,* latent membrane protein; *LDL,* low-density lipoproteins; *LECAM,* lymphocyte endothelial cell adhesion molecule; *LFA,* lymphocyte function antigen; *LPB,* lipopolysaccharide binding protein; *LPS,* lipopolysaccharide; *MAdCAM-1,* mucosal addressin cell adhesion molecule-1; *M-CSF,* macrophage colony-stimulating factor; *MHC,* major histocompatibility complex; *NK cells,* natural killer cells; *PDGF,* platelet-derived growth factor; *PECAM,* platelet endothelial cell adhesion molecule; *PETA,* platelet-endothelial cell tetraspan antigen; *PMN,* polymorphonuclear neutrophil; *SIGN,* specific intercellular adhesion molecule–grabbing nonintegrin; *TACI,* transmembrane activator and CAML interactor; *TCR,* T-cell receptor for antigen; *TNF,* tumor necrosis factor; *TNFR,* tumor necrosis factor receptor; *TRAIL,* TNF-related apoptosis inducing ligand; *VCAM,* vascular cellular adhesion molecule; *VLA,* very late antigen.

This list was adapted from the results of the Eighth International Workshop on Human Leukocyte Differentiation Antigens (HLDA8) held in Adelaide, Australia, in December 2004 (Proceedings of the 8th International Workshop on Human Leukocyte Differentiation Antigens. December 12–16, 2004. Adelaide, Australia. Cell Immunol 2005;236[1–2]:1–187). Engel P, Boumsell L, Balderas R, et al. CD nomenclature 2015: human leukocyte differentiation antigen workshops as a driving force in immunology. J Immunol 2015;195:4555–4563.

APPENDIX 2

Laboratory Reference Values

Thomas A. Fleisher

Immunoglobulin Levels (Age-Related Reference Ranges; see Chapter 93)

Age of Healthy Donors	IgG g/L	$\mathrm{IgG_1}$ g/L	$\mathrm{IgG_2}$ g/L	$\mathrm{IgG_3}$ g/L	$\mathrm{IgG_4}$ g/L	IgA g/L	$\mathrm{IgA_1}$ g/L	$\mathrm{IgA_2}$ g/L	IgM g/L
0–<5 months	1.0–1.34	0.56–2.15	≤0.82	0.07.6–8.23	≤0.198	0.07–0.37	0.10–0.34	0.004–0.055	0.26–1.22
5–<9 months	1.64–5.88	1.02–3.69	≤0.89	0.119–0.740	≤0.208	0.16–0.50	0.14–0.41	0.015–0.062	0.32–1.32
9–<15 months	2.46–9.04	1.60–5.62	0.24–0.98	0.173–0.637	≤0.220	0.27–0.66	0.20–0.50	0.028–0.070	0.40–1.43
15–<24 months	3.13–11.70	2.09–7.24	0.35–1.05	0.219–0.550	≤0.230	0.36–0.79	0.24–0.58	0.039–0.077	0.46–1.52
2–<4 years	2.95–11.56	1.58–7.21	0.39–1.76	0.170–0.847	0.004–0.491	0.27–2.46	0.16–1.62	0.013–0.311	0.37–1.84
4–<7 years	3.86–14.70	2.09–9.02	0.44–3.16	0.108–0.949	0.008–0.819	0.29–2.56	0.17–1.87	0.011–0.391	0.37–2.24
7–<10 years	4.62–16.82	2.53–10.19	0.54–4.35	0.085–10.26	0.010–1.087	0.34–2.74	0.21–2.21	0.014–0.480	0.38–2.51
10–<13 years	5.03–15.80	2.80–10.30	0.66–5.02	0.115–10.53	0.010–1.219	0.42–2.95	0.27–2.50	0.026–0.534	0.41–2.55
13–<16 years	5.09–15.80	2.89–9.34	0.82–5.16	0.200–10.32	0.007–1.217	0.52–3.19	0.36–2.75	0.047–0.551	0.45–2.44
16–<18 years	4.87–13.27	2.83–7.72	0.98–4.86	0.313–0.976	0.003–1.110	0.60–3.37	0.44–2.89	0.066–0.543	0.49–2.01
≥18 years	7.67–15.90	3.41–8.94	1.71–6.32	0.184–10.60	0.024–1.210	0.61–3.56	0.50–3.14	0.097–1.560	0.37–2.86

Immunoglobulin (Ig) levels were assessed in serum by nephelometry, and data were statistically analyzed for the mid-95% confidence interval. For total IgG and IgG subclass quantitation, data from 156 pediatric and 92 adult donors were used; for total IgA quantitation, data from 201 pediatric and 99 adult donors were used; for IgA subclasses, data from 119 pediatric and 99 adult donors were used; and for IgM quantitation, data from 212 pediatric and 401 adults were used at Mayo Medical Laboratories.

Total Serum IgE (IU/mL)

Age	Gender	Geometric Mean	Upper 95% Confidence Limit
6–14 years	M	42.7	527
	F	43.3	344
15–24 years	M	33.6	447
	F	18.6	262
25–34 years	M	16.8	275
	F	16.6	216
35–44 years	M	21.7	242
	F	19.3	206
45–54 years	M	19.2	254
	F	13.3	177
55–64 years	M	21.3	354
	F	11.7	148
65–74 years	M	21.2	248
	F	11.5	122
>75 years	M	18.4	219
	F	9.2	124
6–75 years	all M	22.9	317
	all F	14.7	189

Data were generated using individuals with negative skin prick tests (i.e., house dust mite, Bermuda grass, tree mix, weed mix, mold mix).
From Barbee RA, Halonen M, Lebowitz M, Burrows B. Distribution of IgE in a community population sample: correlations with age, sex, and allergen skin test reactivity. J Allergy Clin Immunol 1981;68(2):106–11, with permission.

Lymphocyte Immunophenotype: Adult Reference Range (95% Confidence Interval)

Surface Antigens	Percent Positive	Cells/mm^3
T Cells		
CD3	60–83.5%	714–2266
CD5	60–83.5%	723–2276
CD2	77.5–94%	817–2496
CD7	74–96.0%	775–2536
CD3/CD4	32–62%	359–1565
CD3/CD8	11–35%	178–853
CD4/CD45RO	10–44%	173–916
CD4/CD45RA	4–21%	55–593
CD8/CD45RO	1.5–11.5%	36–273
CD8/CD45RA	2.5–20.5%	44–405
CD3/CD8/CD28	9.5–26%	126–600
CD8/CD57	≤30.0%	≤521
CD3/HLA-DR	≤19.5%	≤450
CD3/CD25	11–40%	155–905
B Cells		
CD20	3–19%	59–329
CD19	3–19%	61–321
CD20/CD5	0.5–9.5%	14–159
CD20/CD23	1–14%	27–229
CD20/CD27	1–3.5%	12–68
NK Cells		
$\mathrm{CD3^-/CD16^+/CD56^+}$	6–35%	126–729
Lymphocytes	17–41	1173–2640

Data generated in the Flow Cytometry Section, Immunology Service, DLM, CC, NIH, Bethesda, MD. The 95% confidence interval for the WBC is 4300–9200/mm^3.

Age-Dependent Lymphocyte Immunophenotype Reference Range (80% Confidence Interval)

	T CELLS					
	CD3		CD4		CD8	
Age	**Percent positive**	**Cells/mm^3**	**Percent positive**	**Cells/mm^3**	**Percent positive**	**Cells/mm^3**
0–3 mo	53–84	2500–5500	35–64	1600–4000	12–28	560–1700
3–6 mo	51–77	2500–5600	35–56	1800–4000	12–23	590–1600
6–12 mo	49–76	1900–5900	31–56	1400–4300	12–24	500–1700
1–2 yr	53–75	2100–6200	32–51	1300–4300	14–30	620–2000
2–6 yr	56–75	1400–3700	28–47	700–2200	16–30	490–1300
6–12 yr	60–76	1200–2600	31–47	650–1500	18–35	370–1100
12–18 yr	56–84	1000–2200	31–52	530–1300	18–35	330–920

	CD4 T-CELL SUBPOPULATIONS					
	CD4/CD45RA		CD3/CD4/CD45RO		CD4/HLA-DR	
Age	**Percent CD4 positive**	**Cells/mm^3**	**Percent CD3/CD4 positive**	**Cells/mm^3**	**Percent CD4 positive**	**Cells/mm^3**
0–3 mo	64–95	1200–3700	2–22	60–900	2–6	40–180
3–6 mo	77–94	1300–3700	3–16	120–630	2–10	60–280
6–12 mo	64–93	1100–3700	5–18	160–800	2–11	50–260
1–2 yr	63–91	1000–2900	7–20	210–850	2–11	70–280
2–6 yr	53–86	430–1500	9–26	220–660	3–12	50–180
6–12 yr	46–77	320–1000	13–30	230–630	3–13	40–120
12–18 yr	33–66	230–770	18–38	240–700	4–11	30–100

	CD8 T-CELL SUBPOPULATIONS					
	CD8/CD45RA		CD3/CD4$^-$/CD45RO		CD8/HLA-DR	
Age	**Percent CD8 positive**	**Cells/mm^3**	**Percent CD3/CD4-positive**	**Cells/mm^3**	**Percent CD8 positive**	**Cells/mm^3**
0–3 mo	80–99	450–1500	1–9	30–330	2–20	20–160
3–6 mo	85–98	550–1400	1–7	30–290	3–17	30–170
6–12 mo	75–97	480–1500	1–8	40–330	4–27	40–290
1–2 yr	71–98	490–1700	2–12	60–570	6–33	60–600
2–6 yr	69–97	380–1100	4–16	90–440	7–37	70–420
6–12 yr	63–92	310–900	4–21	70–390	6–29	40–270
12–18 yr	61–91	240–710	4–23	60–310	5–25	30–180

	B CELLS AND NK CELLS			
	CD19		CD3$^-$/CD16$^-$56$^+$	
Age	**Percent positive**	**Cells/mm^3**	**Percent positive**	**Cells/mm^3**
0–3 mo	6–32	300–2000	4–18	170–1100
3–6 mo	11–41	430–3000	3–14	170–830
6–12 mo	14–37	610–2600	3–15	160–950
1–2 yr	16–35	720–2600	3–15	180–920
2–6 yr	14–33	390–1400	4–17	130–720
6–12 yr	13–27	270–860	4–17	100–480
12–18 yr	6–23	110–570	3–22	70–480

Data generated by Shearer WT, Rosenblatt HM, Gelman RS, Oyomopito R, Plaeger S, Stiehm ER, et al. Lymphocyte subsets in healthy children from birth through 18 years of age: the Pediatric AIDS Clinical Trials Group P1009 study. J Allergy Clin Immunol 2003;112(5):973–80.

Additional pediatric lymphocyte reference intervals can be found in the following references:

- Piatosa B, Wolska-Kusnierz B, Pac M, Siewiera K, Gałkowska E, Bernatowska E. B cell subsets in healthy children: reference values for evaluation of B cell maturation process in peripheral blood. Cytometry B Clin Cytom 2010;78(6):372–81.
- Schatorje EJH, Gemen EFA, Driessen GJA, Leuvenink J, van Hout RW, de Vries E. Paediatric reference values for the peripheral T cell compartment. Scan J Immunol 2012;75(4): 436–44.
- van Gent R, van Tilburg CM, Nibbelke EE, Otto SA, Gaiser JF, Janssens-Korpela PL, et al. Refined characterization and reference values of the pediatric T- and B-cell compartments. Clin Immunol 2009;133(1):95–107.